Masterarbeit

zur Erlangung des akademischen Grades

Diplom-Ingenieur

der Studienrichtung Angewandte Informatik
an der Paris-Lodron-Universität Salzburg

über das Thema

A Forensic Analysis of the CASIA-Iris V4 Database

eingereicht an der
Naturwissenschaftlichen Fakultät

von

Luca Debiasi, B.Eng.

Bergheimer Straße 6
5020 Salzburg

Begutachter: **Univ.-Prof. Dr. Andreas Uhl**

Salzburg, Juni 2015

This document is set in Palatino, compiled with pdfLaTeX2e and `Biber`.

The LaTeX template from Karl Voit is based on KOMA script and can be found online: https://github.com/novoid/LaTeX-KOMA-template

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Salzburg, _____ _____

Date Signature

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Salzburg, am _____ _____

Datum Unterschrift

Notice on Prior Publication

Parts of this thesis have already been published or submitted for publication. The work on "Generation of Iris Sensor PRNU Fingerprints from Uncorrelated Data" has been published in Proceedings of the 2nd International Workshop on Biometrics and Forensics (IWBF'14) [39]. The analysis of the CASIA-Iris V4 database using different forensic methods (without the application of additional PRNU enhancement techniques) has been published in Proceedings of the 3rd International Workshop on Biometrics and Forensics (IWBF'15) [38]. The work "Blind Biometric Source Sensor Recognition Using Advanced PRNU Fingerprints" is to appear in the Proceedings of the 23rd European Signal Processing Conference (EUSIPCO 2015).

Acknowledgement

This work is partially funded by the Austrian Science Fund (FWF) under Project No. P26630.

Portions of the research in this thesis use the CASIA-IrisV4 collected by the Chinese Academy of Sciences' Institute of Automation (CASIA) (`http://biometrics.idealtest.org/`).

Abstract

The photo response non-uniformity (PRNU) of a digital image sensor can be useful to enhance a biometric systems security by ensuring the authenticity and integrity of the images acquired with the biometric sensor, i.e. by identifying the image source or detecting a possible tampering of the images presented to the biometric system. Passive image forensic techniques have shown to be suited for this tasks for digital consumer cameras.

Previous studies regarding the feasibility of this application using biometric sensors have been conducted for iris and fingerprint sensors by studying the differentiability of the sensors PRNU fingerprints. The results obtained on the CASIA-Iris V4 database showed a high variation among the various subsets of the database. The researchers assumed that this high variation could either be caused by the highly correlated data or come from the usage of different sensors for the acquisition of the subsets.

To investigate the latter issue a forensic investigation on the CASIA-Iris V4 database has been performed in this thesis, where an existing forensic technique has been applied and additionally several novel forensic techniques have been proposed to detect the presence of images from multiple sensors in the image data sets. Furthermore, the forensic techniques have been applied on a test data set with a known number of sensors first to evaluate their performance.

Since there is no specific documentation on the number of sensors used for the acquisition, the investigation on the CASIA-Iris V4 database has been conducted in a blind manner and without any a priori knowledge about the sensors. In addition different PRNU enhancement approaches have been applied for the investigation to reduce the contamination of the PRNU fingerprints by the image content, which is an issue in this scenario due to the strong correlation of the content among all images under investigation.

Zusammenfassung

Die sogenannte „Photo Response Non-Uniformity“ (PRNU) eines digitalen Bildsensors kann zur Erhöhung der Sicherheit in einem biometrischen Systems beitragen, indem die Authentizität und Integrität der mit dem Sensor aufgenommenen Daten sichergestellt werden. Dies kann unter anderem durch eine Identifizierung der Quelle oder durch die Erkennung von möglichen Veränderungen des Inhaltes der Bilder, welche dem biometrischen System präsentiert werden, geschehen. Bei der Verwendung von digitalen Verbraucher-Kameras wurden hierfür bereits erfolgreich passive Methoden aus der digitalen Bildforensik angewendet.

Machbarkeitsstudien über die Anwendung dieser passiven Methoden bei biometrischen Sensoren wurden bereits mit Iris- und Fingerabdruck-Daten durchgeführt, wobei die Unterscheidbarkeit der „PRNU fingerprints“ der Sensoren untersucht wurde. Eine Studie, welche auf der CASIA-Iris V4 Datenbank durchgeführt wurde, ergab hierbei hohe Schwankungen in Hinsicht auf die Unterscheidbarkeit der einzelnen Datensets. Die Forscher vermuteten die Herkunft dieser Schwankungen in der hohen Korrelation der Bilder untereinander oder dass die Datensets mit Hilfe von mehreren Sensoren aufgenommen worden sein könnten.

Um die letztere Vermutung zu analysieren wurde in dieser Masterarbeit eine forensische Untersuchung der CASIA-Iris V4 Datenbank unter Zuhilfenahme eines bestehenden und neu entwickelter forensischer Verfahren durchgeführt, um das Vorkommen von Bildern, die von unterschiedlichen Sensoren aufgenommen wurden, in einem Datenset zu erkennen. Weiteres wurden die forensischen Verfahren zuvor an einem Test-Datenset, für welches die Anzahl der Sensoren bekannt ist, angewendet, um ihr Leistungsverhalten zu evaluieren.

Da es keine genaue Dokumentation bezüglich der Anzahl der für jedes Datenset verwendeten Sensoren gibt, wurde die forensische Untersuchung der CASIA-Iris V4 Datenbank blind und ohne a-priori Wissen durchgeführt. Hierbei wurden außerdem verschiedene Ansätze, die der Kontaminierung der „PRNU fingerprints“ durch die Bildinhalte entgegenwirken welche in diesem Szenario durch die hohe Korrelation der untersuchten Bilder ein Problem darstellt, angewendet.

Contents

1 Introduction

The imaging sensor is an essential component of every electronic device capable of producing digital pictures. It consists of a large number of photo sensitive detectors made of silicon, commonly called pixels. They have the ability to convert photons into electrons by exploiting the photoelectric effect [30, 34]. The charge accumulated in every pixel is first amplified and afterwards converted into a digital signal, which is further processed and then stored onto a storage device like a memory card.

Biometric systems are not new, for example they are deployed as authentication systems in industrial settings and high-security areas such as laboratories, banks and border control for several years, but their use is becoming more and more widespread in everyday life as well, i.e. fingerprint recognition in the case of smartphones [14] or Microsoft planning to introduce system support for biometric authentication in Windows 10 [5]. Another important application is the Indian UIDAI program "Aadhaar" [1] issuing a unique identification number to each Indian resident and using different biometric modalities to identify and distinguish the enrolled subjects.

The sensors used in biometric systems are often specialized devices and are usually adapted to deliver the desired data for a special modality. For many modalities, such as iris, fingerprint, face, gait and others, digital images of the biometric trait are captured and the specific features are then extracted from these images. The security of the whole system should not be compromised by the data presented to the system, thus it has to be assured that the data has been acquired with the proper sensor (sensor authenticity) and additionally that the data has not been tampered or altered (data integrity). Further details on the security of biometric systems are given in section 1.3.

[1] https://uidai.gov.in

To ensure the authenticity of the biometric sensor, first the discriminative power of the biometric sensors has to be evaluated, as it has been done in [64] and [4] using digital image forensic techniques, which are further explained later in section 1.1 and 1.2. The results from Höller *et al.* [64], where the discriminative power of five iris sensors from the *CASIA-Iris V4* database has been evaluated, vary highly: Some sensors showed satisfying results while others did not and it opened the question what might have caused this variability in the results.

This thesis aims at resolving the open questions from Höller *et al.* [64] and explain the variability in the results by performing a forensic analysis of the *CASIA-Iris V4* database. This chapter gives a further introduction to Digital Multimedia forensics, Digital Image Forensics and Security in Biometric Systems. Chapter 2 explains the theoretical background of the forensic techniques applied in this thesis and shows the definitions of important concepts used for this work.

The motivation for this thesis is further discussed in chapter 3 and related work from literature dealing with a similar scenario is presented in chapter 4. The techniques applied for the forensic analysis of the *CASIA-Iris V4* database are proposed and explained in chapter 5, while the database itself and its data sets are described in chapter 6 together with a test data set used to evaluate the forensic techniques beforehand.

The experimental set-up and the results for the *Test* and *CASIA-Iris V4* data sets are shown in chapter 7 and 8. Finally, chapter 9 concludes the thesis and an outlook on future work is given.

1.1 Digital Multimedia Forensics

Digital Multimedia describes various forms of how to represent any content in a digital manner, thus different from traditional representation forms such as printed material or hand-crafted objects. These forms include still images, audio, video, text, animations or even combinations of any of those.

With the increasing diffusion of digital audio and visual contents in everyday life, the investigation on multimedia objects is acquiring more and more

interest within the framework of digital investigations, that considers all the aspects including digital data and digital devices.

As the presence of digital content in our life is rising, also gathering informations on multimedia content is becoming more important for law enforcement authorities, protection of data privacy, security of computer systems, digital rights management and others.

As outlined in [13], Multimedia Forensics aims at acquiring knowledge on the history of audio-visual contents by evaluating the traces that are left on the data by each processing step. Many algorithms have been proposed that extract features from the audio-visual content and, based on these features, try to infer some information about the acquisition device, the editing/coding processing undergone by the digital content, possible inconsistencies revealing tampering operations, etc.

In digital image and video forensics various techniques exist, like identification of the source sensor of a image or video, or verification of the integrity of the data (if it has been modified or not) without any prior knowledge about the data. These tools work with intrinsic signatures added to the data during the acquisition and in the subsequent processing. There also exist forensic techniques for digital audio data like audio tampering detection, as described in [11].

1.2 Digital Image Forensics

Digital image forensics covers the part of Digital Multimedia Forensics dealing with still images and analyzing traces in still image data. Two major tasks in this field are establishing an images origin and its integrity. An images origin can be determined by the level of analysis and can either be the device type (i.e. camera or scanner), the manufacturer, the model or even a specific sensor among many models from the same manufacturer. The verification of the integrity of an image analyses whether the image has been geometrically transformed (e.g. cropped, rotated, turned, flipped etc.) or if parts of the image have been tampered (e.g. deleted, copied, replaced,

(a)Original (b)Tampered (c)Mask

Figure 1.1: Tampered image (*malawi.png*) from the "Image Manipulation Dataset" with the original image showing a person on a boat (a), the tampered image where the person and the boat have been removed (b) and the tampering mask showing the altered image areas in white (c).

altered). An example for such a image tampering taken from the "Image Manipulation Dataset" [8] is given in Figure 1.1.

In contrast to digital watermarking as authenticity technique, as mentioned in [23], digital image forensics does not require any active embedding step at the time of creation or publication. Evidence is extracted merely from structural analysis of image files and statistical analysis of the image data (i. e. the two-dimensional array of pixel intensities).

An important tool used to perform this forensic task is the photo-response non-uniformity (PRNU) of imaging sensors as described by [13, 19]. It can be used for a variety of important digital forensic tasks, such as device identification, device linking, recovery of processing history, and detection of digital forgeries. The PRNU is an intrinsic property of all digital imaging sensors due to slight variations among individual pixels in their ability to convert photons to electrons. Consequently, every sensor casts a weak noise-like pattern onto every image it takes. This pattern, which plays the role of a "sensor fingerprint", is essentially an unintentional stochastic spread-spectrum watermark that survives processing, such as lossy compression or filtering. This fingerprint can be estimated from images taken by the camera and later detected in a given image to establish image origin and integrity. There exist two types of imaging sensors commonly found in digital cameras, camcorders, and scanners: charge-coupled device (CCD) and complementary metal-oxide semiconductor (CMOS). Both consist of a large number of photo detectors, also called pixels. Pixels are made of

silicon and capture light by converting photons into electrons using the photoelectric effect. The accumulated charge is transferred out of the sensor, amplified, and then converted to a digital signal in an A/D converter and further processed before the data is stored in an image format, such as JPEG.

Even though the PRNU is stochastic in nature, it is a relatively stable component of the sensor over its life span and is therefore a very useful forensic quantity, responsible for a unique sensor fingerprint with the following important properties [19]:

1. **Dimensionality**: The fingerprint is stochastic in nature and has a large information content, which makes it unique to each sensor.
2. **Universality**: All imaging sensors exhibit PRNU.
3. **Generality**: The fingerprint is present in every picture independently of the camera optics, camera settings, or scene content, with the exception of completely dark images.
4. **Stability**: It is stable in time and under a wide range of environmental conditions (temperature, humidity, etc.).
5. **Robustness**: It survives lossy compression, filtering, gamma correction, and many other typical processing procedures.

As mentioned before, the PRNU fingerprint can be used for various forensic tasks [19]:

- By testing the presence of a specific fingerprint in the image, one can achieve reliable device identification (e.g., prove that a certain camera took a given image) or prove that two images were taken by the same device (device linking). The presence of camera fingerprint in an image is also indicative of the fact that the image under investigation is natural and not a computer rendering.
- By establishing the absence of the fingerprint in individual image regions, it is possible to discover replaced parts of the image (integrity verification).

- By detecting the strength or form of the fingerprint, it is possible to reconstruct some of the processing history. For example, one can use the fingerprint as a template to estimate geometrical processing, such as scaling, cropping, or rotation. Non-geometrical operations will also influence the strength of the fingerprint in the image and thus can potentially be detected.
- The spectral and spatial characteristics of the fingerprint can be used to identify the camera model or distinguish between a scan and a digital camera image (the scan will exhibit spatial anisotropy).

1.3 Security in Biometric Systems

Biometrics-based authentication is becoming more and more popular as an alternative to knowledge and possession based authentication techniques like passwords, smart-cards, PINs etc. because it offers various advantages over the latter: They can be lost or forgotten and are not reliant on the presence of a human being. Biometric authentication is able to resolve these issues, since biometric features are diverse among different human beings and they can not be lost or forgotten, but they can eventually be stolen or forged and compromised biometric templates cannot be revoked and reissued. Therefore such biometrics-based authentication systems need to be designed to resist attacks when employed in a security-critical environment. Jain *et al.* [33] state that public acceptance of biometrics technology will depend on the ability of system designers to demonstrate that these systems are robust, have low error rates, and are tamper proof.

Ratha *et al.* [57] identified eight stages in a generic biometric system where attacks may occur. These stages are shown in Figure 1.2, where the numbers in the figure correspond to the items in the following list:

1. **Presenting fake biometrics at the sensor**: In this mode of attack, a possible reproduction of the biometric feature is presented as input to the system. Examples include a fake finger, a copy of a signature, or a face mask.
2. **Resubmitting previously stored digitized biometrics signals**: In this mode of attack, a recorded signal is replayed to the system, bypassing

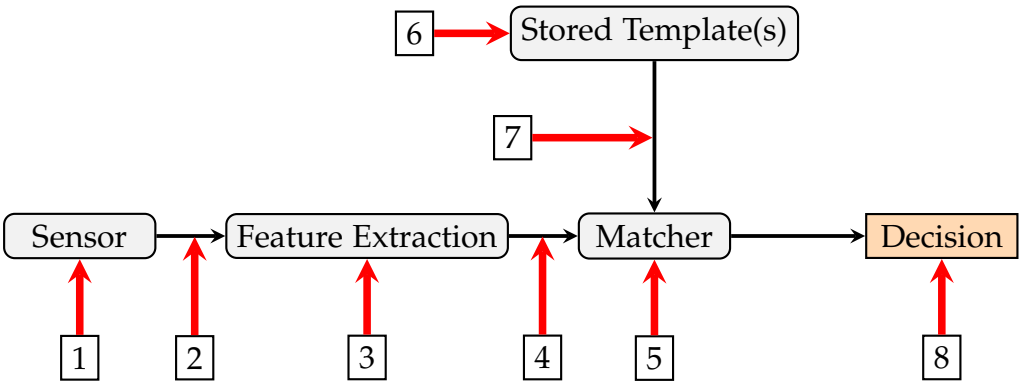


Figure 1.2: Attack vectors in a generic biometric system.

the sensor. Examples include the presentation of an old copy of a fingerprint image or the presentation of a previously recorded audio signal.

3. **Overriding the feature extraction process**: The feature extractor is attacked using a Trojan horse, so that it produces feature sets preselected by the intruder.
4. **Tampering with the biometric feature representation**: The features extracted from the input signal are replaced with a different, fraudulent feature set (assuming the representation method is known). Often the two stages of feature extraction and matcher are inseparable and this mode of attack is extremely difficult. However, if minutiae are transmitted to a remote matcher (say, over the Internet) this threat is very real. One could "snoop" on the TCP/IP (Transmission Control Protocol/Internet Protocol) stack and alter certain packets.
5. **Corrupting the matcher**: The matcher is attacked and corrupted so that it produces preselected match scores.
6. **Tampering with stored templates**: The database of stored templates could be either local or remote. The data might be distributed over several servers. Here the attacker could try to modify one or more templates in the database, which could result either in authorizing a fraudulent individual or denying service to the persons associated with the corrupted template. A smart-card based authentication system, where the template is stored in the smart-card and presented to the

authentication system, is particularly vulnerable to this type of attack.

7. **Attacking the channel between the stored templates and the matcher**: The stored templates are sent to the matcher through a communication channel. The data travelling through this channel could be intercepted and modified.
8. **Overriding the final decision**: If the final match decision can be overridden by the hacker, then the authentication system has been disabled. Even if the actual pattern recognition framework has excellent performance characteristics, it has been rendered useless by the simple exercise of overriding the match result.

Some biometric traits, such as fingerprints, irises, palm-prints, faces, etc. are inevitably presented to the open public and can easily be collected by the attacker, e.g. lifting fingerprints from a glass or taking a picture of the face of desired subject with a telephoto lens where eventually also the iris can be extracted [27, 41]. The goal in this scenario is to spoof the system at sensor level (stage 1 in the previous list) by presenting a reproduction of the biometric traits. If no further security procedures are established at this level, like a liveness detection, the acquired image data could be presented to the sensor or or could be inserted into the data transmission from sensor to the feature extractor (stage 2 in the previous list).

Encryption and other classical authentication techniques like digital signatures or data-hiding have been suggested to secure the previously mentioned transmission channel (stage 2) by verifying the senders (i.e. sensor and feature extractor) authenticity, as well as the integrity of the entire authentication mechanism. The proposed approaches can be divided into active and passive-blind approaches.

Active methods consist of data hiding approaches [62, 66] and the digital signature approaches [61, 63, 50, 49]. Höller *et al.* [64] describe the pros and cons of these active methods as follows:

- **Classical digital signatures** work by adding additional data to verify the original data, whereas watermarks become an integral part of the sample data, and moreover, spatial locations of eventual tampering can be identified [29].
- **Fragile watermarks** (as proposed for these tasks in e.g. [55, 35, 58]) cannot provide any form of robustness against channel errors and

unintentional signal processing "attacks" like compression, which is the same as with classical digital signatures.

- **Semi-fragile watermarks** have been designed to differentiate between allowed signal processing operations and malicious attacks and have also been suggested for employment in biometric systems [37, 48, 15, 1].

Höller *et al.* [64] also mention that a general drawback of watermarks is the representation of additional data which is inserted into the sample data, where an impact on recognition accuracy may be expected. In fact, literature reports on corresponding effects in case of iris recognition [28], speech recognition [40], and fingerprint recognition [31].

Passive-blind approaches, in contrast to active methods, do not need any prior information about the image. As stated in [56], passive-blind approaches are mostly based on the fact that forgeries can bring specific detectable changes into the image (e.g., statistical changes). In high quality forgeries, these changes cannot be found by visual inspection. Höller *et al.* [64] propose a suitable passive approach to secure the transmission channel between the sensor and the feature extractor, making use of sensor fingerprints based on a sensors PRNU [7]. Besides image integrity, this technique can also provide authenticity by identifying the source sensor uniquely and important properties as required in a biometric scenario have been demonstrated: suitability to manage large datasets [26, 54], robustness against common signal processing operations like compression and malicious signal processing [59, 2], and finally methodology to reveal forged PRNU fingerprints has been established [25].

2 Theoretical Background

As described by [32], in theory, the amount of electrons (charge) outputted by a pixel should depend solely on the intensity of the incident light. In reality, however, there are many factors that introduce both systematic and random deviations. These persistent fluctuations, which are independent from scene to scene, can be exploited for forensic analysis of the images acquired by a digital imaging sensor. Consequently, every sensor casts a weak noise-like pattern onto every image it takes. This pattern, the photo-response non-uniformity (PRNU), is essentially an unintentional stochastic spread-spectrum watermark [20].

The PRNU of a imaging sensor emerged as an important component to perform various forensic tasks such as device identification, device linking, recovery of processing history and the detection of digital forgeries. This chapter provides a specific explanation of how a PRNU noise estimate is obtained, a sensor fingerprint is generated from multiple noise estimates and the matching of sensor fingerprints against noise residuals or other sensor fingerprints is performed.

Section 2.1 describes the sensor output model by Fridrich [19]. This model is used to extract the PRNU noise residual of an image, as shown in section 2.2. The concept of a PRNU fingerprint and the procedure of how to estimate the fingerprint from different images of the same sensor is explained in section 2.3 and section 2.4 explains the sensor fingerprint matching. Section 2.5 concludes this chapter by describing different post processing techniques to reduce contaminations in the PRNU estimates.

2.1 Sensor Output Model

As described by Fridrich [19], pixels are usually rectangular and have a size of several microns across. The amount of electrons generated by the incident light at each pixel depends on the physical dimensions of the pixels photosensitive area. Every pixels physical dimension varies slightly due to imperfections in the manufacturing process and furthermore, the inhomogeneity of silicon contributes to variations in the quantum efficiency among pixels (the ability to convert photons to electrons). Essentially all types of imaging sensors (CCD, CMOS, JFET, or CMOS-Foveon-X3) are built from semiconductors, and their manufacturing techniques are similar. Therefore, these sensors will likely exhibit fingerprints with similar properties.

According to Fridrich [19], the raw output of a sensor with $w \times h$ pixels can be modeled as:

$$\begin{aligned} Y &= I + I \circ K + \tau D + C + \Theta \\ \textit{with} \quad Y, I, K, D, C, \Theta &\in \mathbb{R}^{w \times h}; \tau \in \mathbb{R} \end{aligned} \tag{2.1}$$

where Y is the sensor output, commonly denoted as image. I represents the incoming light or rather the scene, $I \circ K$ the photo-response non-uniformity PRNU, τD the dark current (with τ being a multiplicative-factor representing exposure settings, sensor temperature, etc.). The matrix C is a light-independent offset and Θ some modeling noise, which is a collection of all other noise sources mostly random in nature and thus difficult to use for forensic purposes (readout noise, shot noise or photonic noise, quantization noise, etc.).. Since all pixels are independent [19] and all operations element-wise, the matrix-elements $y_{x,y} \in Y$ are denoted as $y \in Y$ for simplicity reasons. The same applies to $i \in I$, $k \in K$, $d \in D$, $c \in C$ and $\theta \in \Theta$.

Equation 2.1 also shows that, besides the PRNU, the sensor output essentially contains all systematic defects of the sensor, including stuck (pixels that consistently produce a constant output independently of illumination) and partially stuck pixels (pixels that add a constant offset to the incident light) [17, 16, 43, 42]. It should be stressed that the defects represented by matrices K, D, and c usually represent quite small deviations with the

exception of the pixel defects and that the three matrices can be estimated from multiple images taken by the camera [20].

2.2 PRNU Noise Residual

The differences among pixels are captured with the matrix K in Equation 2.1 in the previous section, which has the same dimension as the sensor. When the imaging sensor is illuminated with ideally uniform light intensity I and other additional noise sources are non present, the sensor captures a signal $I + IK$. This signal, $I + IK$, represents the scene I overlaid with the noise pattern IK, which is usually referred to as the PRNU.

To extract the noise pattern from an image, a denoising filter F has to be applied to the sensor output Y. The denoised output $F(Y)$ is then subtracted from the sensor output Y [19] to obtain the PRNU noise residual W, also known as sensor pattern noise (SPN) in literature. The extraction can be seen in Equation 2.2.

$$\begin{aligned} W &= Y - F(Y) \\ &= IK + \tau D + C + I - F(Y) + \Theta \\ &= IK + \tau D + C + \Xi \end{aligned} \tag{2.2}$$

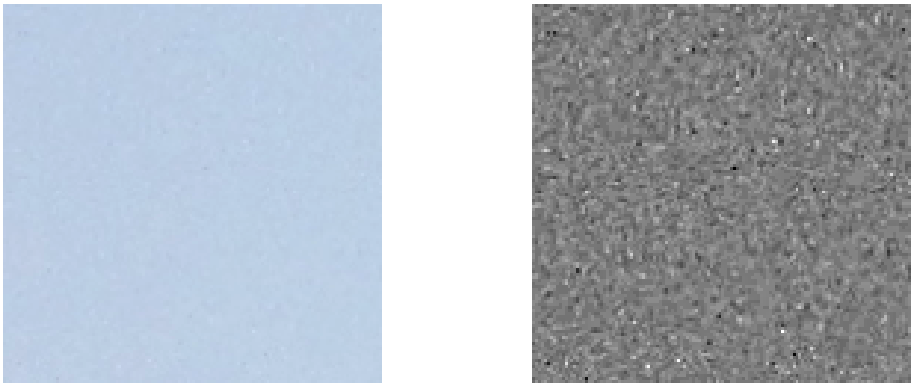
The variable Ξ stands for the sum of the modeling noise and the remnant of the content $I - F(Y)$ present due to the inability of the denoising filter to separate content from noise. The term $I - F(Y)$ is especially large in textured regions and around edges and leads to a less accurate estimation of the PRNU.

The denoising function F filters out the sensor pattern noise. In this work the denoising filter as described in Appendix A of [52] has been used, because it is producing good results in filtering out the PRNU. The filter is constructed in the wavelet domain, where the high-frequency wavelet coefficients of the noisy image are modeled as an additive mixture of a locally stationary i.i.d. (independent and identically distributed) signal with zero mean (the

noise-free image) and a stationary white Gaussian noise $N(0, \sigma_0^2)$ (the noise component). For further details please refer to Appendix A of [52].

The matrix K is responsible for a major part of what is commonly denoted as the camera fingerprint, as it can be seen Equation 2.2. Its energy depends on the intensity of the incident light, therefore the PRNU term IK is only weakly present in dark images respectively dark image regions. Saturated images or image regions, where the pixel value is at its maximum, do not contain any traces of the PRNU or other noise signals.

Furthermore the energy of the PRNU noise residual strongly varies among camera models, but in general it is a quite weak signal [20, 64] which resembles white noise with an attenuated high-frequency band.



(a)*Magnified portion of a sky image* (b)Extracted PRNU

Figure 2.1: Patch with size of 128x128 pixels extracted from the upper left corner of a sky image acquired with a Casio EX-Z150 camera and the extracted PRNU noise residual.

Figure 2.1 shows a magnified portion of a sky image with uniform lighting from a *Casio EX-Z150* in the Dresden Image Database [23] and the extracted PRNU noise residual. The brighter pixels generate more electrons, while the darker ones have a lower response.

2.3 PRNU Fingerprint

Using only one image for estimating a sensors PRNU is usually not sufficient, because that specific image may contain various kinds of disturbances such as the shot noise (random variations in the number of photons reaching the pixel caused by quantum properties of light), the readout noise (random noise introduced during the sensor readout), high frequency image content (edges or textured regions), and also unknown noise sources [20]. To suppress these undesirable random noise components, multiple images from the same sensor are averaged to isolate the systematic components of all images. The averaged noise is denoted as PRNU fingerprint or reference pattern noise (RPN) in literature.

Fridich [32] states that, for example, the fingerprint can be estimated experimentally by taking many images of a uniformly illuminated surface. The best images for estimating the fingerprint are those with high luminance (but not saturated) and small σ^2 (images with a smooth content). If the camera under investigation is available to the analyst, unsaturated out-of-focus images of bright cloudy sky would be the best. In practice, good estimates of the fingerprint may be obtained from as few as 50 natural images depending on the camera. If sky images are used instead of natural images, only approximately one half of them would suffice to obtain an estimate with a comparable accuracy.

To obtain a high quality estimate for the PRNU of a sensor, a series of N images i has to be taken for the estimation, where I_i denotes the i-th image and $i = 1...N$. For each image I_i, the noise residual estimate W_I^i has to be extracted as explained in Equation 2.2 in the previous section. The PRNU fingerprint $\hat{K}$ of a sensor is then estimated using a maximum likelihood estimator as shown in Equation 2.3.

$$\hat{K} = \frac{\sum_{i=1}^{N} W_I^i I^i}{\sum_{i=1}^{N} (I^i)^2} \tag{2.3}$$

After the estimation of the PRNU fingerprint several post-processing techniques can be to remove non-unique artifacts (NUA), since they are not

unique and occur among images from various sensors and hence can lower the differentiability among different sensors. Some examples for NUA and a more detailed explanation are given in section 2.5.

2.4 Fingerprint Matching

To be able to determine whether an image has been acquired with a specific sensor or not, the presence of the sensors PRNU fingerprint can be evaluated in the image under investigation. A method to do so for images that have not been geometrically transformed is the normalized cross correlation (NCC), which is defined in Equation 2.4.

$$NCC(A,B) = \frac{\sum_{w=1}^{W}\sum_{h=1}^{H}(A(w,h)-\overline{A})(B(w,h)-\overline{B}))}{\sqrt{(\sum_{w=1}^{W}\sum_{h=1}^{H}(A(w,h)-\overline{A})^2)(\sum_{w=1}^{W}\sum_{h=1}^{H}(B(w,h)-\overline{B})^2)}} \tag{2.4}$$

A and B are two matrices of the same size $w \times h$ and $\overline{X}$ is the average or mean of the matrix X with size $W \times H$:

$$\overline{X} = \frac{1}{WH}\sum_{w=1}^{W}\sum_{h=1}^{H}X(w,h) \tag{2.5}$$

The normalized cross correlation (NCC) is used to detect the presence of a PRNU fingerprint $\hat{K}$ in an Image I with

$$\rho_{[I,\hat{K}]} = NCC(W_I, I\hat{K}) \tag{2.6}$$

where ρ indicates the correlation between the noise residual W_I of the image I and the PRNU fingerprint $\hat{K}$ weighted by the image content of I.

On the other hand, the NCC is also used to measure the similarity of two sensor fingerprints $\hat{K}_I$ and $\hat{K}_J$ from two sensors S_i and S_j, as shown in Equation 2.7.

$$\rho_{[\hat{K}_I,\hat{K}_J]} = NCC(\hat{K}_I, \hat{K}_J) \tag{2.7}$$

To measure the the discriminative performance of two distinct sensors S_i and S_j, the correlation score ρ, as described in Equation 2.6, is calculated between every image from a sensor S_i and the PRNU fingerprint $\hat{K}_i$ of the sensor S_i, where only images are used that have not been part of the PRNU fingerprint estimation to obtain the matching scores. Additionally the correlation ρ between all images from the other sensors S_j, where $i \neq j$, and the PRNU fingerprint $\hat{K}_i$ of the sensor S_i is also calculated to obtain the non-matching scores.

Once the matching and non-matching scores have been calculated, the equal error rate (EER) is calculated to evaluate the discriminability between different sensors. It relies on 2 different types of errors, the false match rate (FMR) and the false-non-match rate ($FNMR$) and it determines the value where FMR and $FNMR$ are equal, i.e., where $FMR = FMNR$. Let I be an image captured with sensor S_I, J is an image captured with sensor S_J and $\hat{K}$ is the PRNU fingerprint estimate for the sensor S_I, where $S_I \neq S_J$.

A false match (FM) is given, if

$$\rho_{[J,\hat{K}]} \geq min(\rho_{[I,\hat{K}]}) \tag{2.8}$$

Equation 2.8 implies that the PRNU fingerprint of sensor S_I is more likely present in the image J although it has been acquired with the sensor S_J and the NCC score $\rho_{[J,\hat{K}]}$ is higher than the minimal matching NCC score $\rho_{[I,\hat{K}]}$. This leads to an identification error because no NCC score threshold can be found to exclude the NCC score of image J from the matches without also excluding the NCC score of image I. The more images J from sensor S_J have a higher correlation score with the PRNU fingerprint $\hat{K}_I$ than images I taken with sensor S_I, the higher the FMR becomes.

This brings us to the other error type, the false non match (FNM), which is given if

$$\rho_{[I,\hat{K}]} \leq max(\rho_{[J,\hat{K}]}) \tag{2.9}$$

Equation 2.9 implies that the PRNU fingerprint of sensor S_I is less likely present in the image I although it has been acquired with the same sensor S_I and the NCC score $\rho_{[I,\hat{K}]}$ is lower than the maximal non-matching NCC score$\rho_{[J,\hat{K}]}$. This leads to an identification error because no NCC score threshold can be found to exclude the NCC score of image I from the non-matches without also excluding the NCC score of image J. The more images I from sensor S_I have a lower correlation score with the PRNU fingerprint $\hat{K}_I$ than images J taken with sensor S_J, the higher the FNMR becomes. The EER rate describes the point where the FMR and FNMR are equal and can therefore be used to measure the discriminability between two different sensors, where a lower EER implies better discrimination performance.

The EER is calculated by comparing different sensors pairwise. To estimate the real variability of all the correlation scores ρ, the interval of confidence at α% is estimated, because only a finite set of values was used to calculate the EER. To do so, the calculation of EER and threshold was repeated 1000 times on the respective set of m matching and n non-matching ρ values, by drawing m correlation values from the matching-data set and n correlation values from the non-matching data set, making use of sampling with replacement. As a result, we obtained 1000 EERs and the range containing α% of these values is the interval of confidence. This range excludes the $\frac{100-\alpha}{2}$% lowest and $\frac{100-\alpha}{2}$% highest measures, hence the most extreme values.

2.5 PRNU post-processing

The PRNU respectively the PRNU fingerprint may be subject of undesired contaminations that may cause a decrease of the discriminative power between different sensors. These artifacts are caused by two sources: non-unique artifacts (NUA) and the interference of image content. NUA are usually spatially dependent similarities shared among sensors of the same manufacturer or model, but can also be caused by specific steps during the in-camera processing that are usually performed on every camera, e.g. JPEG compression or color filter array (CFA) interpolation. On the other hand, high frequency components of the image content can interfere with

the PRNU because it consists of high-frequency components itself, hence it is rather difficult to separate these two properties.

To deal with both of these artifacts, various enhancement methods are proposed in the literature. Some methods aim at removing the NUA while others aim at the separation of the PRNU from the image content or try to suppress the contamination by the image content.

NUA may lead to false accusations because of an increased similarity between the noise pattern of an image and the PRNU fingerprint of a different device with similar characteristics [22]. NUA with regular grid structure, such as JPEG compression and CFA artifacts, can be suppressed by applying a zero mean operation to each row and each column of the estimated reference noise pattern (PRNU fingerprint) [7]. Arbitrary periodic structures in noise estimates can be attenuated by Wiener filtering in the frequency domain. If periodic artifacts are not removed properly, cross-correlation peaks may not only result from the alignment of the estimated noise of an image and the corresponding reference noise pattern, but also from non-unique artifacts shifted by multiples of a full spatial period length [22].

Gloe *et al.* [22] report a number of model-specific artifacts that may cause problems with sensor noise estimates from images in the *Dresden Image Database* [23]. These artifacts are of non-trivial nature and concern models of well-known camera manufacturers, namely Nikon, Fujifilm and Casio, respectively and consist of diagonal artifacts, exposure time dependent post processing artifacts and focal length dependent distortion correction artefacts observable for various camera models.

Goljan *et al.* [24] report non-linear artifacts due to correction of radial lens distortion. Although a relatively simple polynomial model was proposed to suppress these artifacts, it is also acknowledged that it generally remains an open question how more and more advanced in-camera processing procedures affect the reliability of PRNU-based camera identification.

Li [44] points out that the SPN of an image can be contaminated by the image content. High frequency image content such as edges and textures can be recognized in the SPN and therefore attenuate the quality of the PRNU: The larger the magnitude of a component in the PRNU, the more likely

it is contaminated by the image content. Therefore Li proposed various attenuation models and evaluated their effect in [44]. To enhance an SPN, first a discrete wavelet transformation (DWT) is performed, followed by a low-pass filtering in the DWT. Then the PRNU is extracted using the denoising filter from Appendix A of [52] and finally the enhancement model is applied to the unenhanced PRNU in the DWT domain. Li reported that the models 3, 4, and 5 delivered the best results. The equation for the attenuation model 3 is given in Equation 2.10.

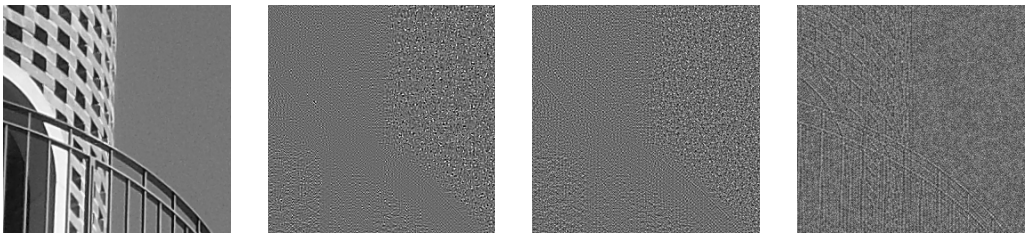
$$n_e(i,j) = \begin{cases} 1 - e^{-n(i,j)}, & \text{if } 0 \leq n(i,j) \leq \alpha \\ (1 - e^{-\alpha}) \cdot e^{\alpha - n(i,j)}, & \text{if } n(i,j) > \alpha \\ -1 + e^{n(i,j)}, & \text{if } -\alpha \leq n(i,j) < 0 \\ (-1 + e^{-\alpha}) \cdot e^{\alpha + n(i,j)}, & \text{if } n(i,j) < -\alpha \end{cases} \tag{2.10}$$

Caldelli *et al.* [6] picked up this idea and proposed an improved attenuation model, which is given by Equation 2.11. In both attenuation functions (Li model 3 and Caldelli) $n(i,j)$ is the initial coefficient in the DWT domain, $n_e(i,j)$ is the enhanced coefficient and α is the threshold at which the content is separated from the PRNU components.

$$n_e(i,j) = \begin{cases} 0, & \text{if } n(i,j) < -\alpha \\ -cos(\frac{n(i,j)\pi}{2\alpha}), & \text{if } -\alpha \leq n(i,j) \leq 0 \\ cos(\frac{n(i,j)\pi}{2\alpha}), & \text{if } 0 < n(i,j) \leq \alpha \\ 0, & \text{if } n(i,j) > \alpha \end{cases} \tag{2.11}$$

The enhancement effects on the PRNU of both functions is illustrated in Figure 2.2, where the contrast has been enhanced for better visibility.

Kang *et al.* [36] proposed a source camera identification scheme based on an eight-neighbor context-adaptive PRNU predictor to enhance the receiver operating characteristic (ROC) performance of camera source identification (CSI). The PRNU predictor is able to further suppress the contaminations from image content and leads to a more accurate PRNU estimation because of its adaptability of different image edge regions. The proposed method



(a)Original (b)Unenhanced (c)Li Model 3 (d)Caldelli

Figure 2.2: Examples of different enhancement approaches suppressing image content contaminations in the PRNU.

needs at least 100 original images to create a camera fingerprint; the advantage of the proposed method decreases when the camera fingerprint is created with less original images.

In [9], Cooper proposed a simplified filtering strategy which at its base level may be realized using a combination of adaptive and median filtering applied in the spatial domain. The proposed filtering method is interlinked with a further two stage enhancement strategy where only pixels in the image having high probabilities of significant PRNU bias are retained. The proposed base adaptive spatial filter coupled with a median filter significantly out performs the commonly applied wavelet denoising filtering method and is further improved by the proposed two stage data optimization process. The system described produces higher discriminability between matching and non-matching PRNU data, leading to an improvement in robustness to image data sets that have been made under more challenging conditions (such as increased levels of compression).

3 Iris-Sensor Authentication using Photo-Response Non-Uniformity

Integrity and authenticity are important issues in biometrics security and digital image forensics have shown that the these issues can be evidenced by a specific sensor fingerprint. This chapter shows the results of previously conducted studies that evaluated the application of PRNU fingerprints as an authentication method for biometric systems. The results of the previous studies form the basis for the motivation of this thesis, which is explained at the end of this chapter.

3.1 Iris-Sensor Authentication using Camera PRNU Fingerprints

Previous feasibility studies by Höller *et al.* [64] have been conducted on the *CASIA-Iris V4* database. The differentiability of the sensors in the CASIA Iris V4 database using sensor fingerprints has been tested with the conclusion, that the EERs and respective thresholds vary highly. Some sensors showed satisfying results while others showed EERs of over 20%, as it can be seen in Table 3.1.

The question raised, that if the PRNU fingerprint is going to be applied as an authentication measure for iris databases, several open questions need to be answered. Firstly, it is not clear if the poor EER values for some sensor combinations come from the image's special content with low variance between the images, or from the properties of the sensors. In the latter case, the PRNU would not be suited for sample data authentication. However, this case was determined as implausible since for some sensors the EERs

Data set	Interval	Lamp	Twins	Distance
Thousand	20.68	11.74	1.68	1.13
Distance	21.21	7.42	0.37	
Twins	18.89	14.39		
Lamp	22.42			

Table 3.1: EERs showing the pairwise discriminability of all possible sensor combinations in the CASIA Iris V4 database.

Data set	Sensor
Casia-Iris-Interval	CASIA close-up iris camera
Casia-Iris-Lamp	OKI IRISPASS-h
Casia-Iris-Twins	OKI IRISPASS-h
Casia-Iris-Distance	CASIA long-range iris camera
Casia-Iris-Thousand	Irisking IKEMB-100

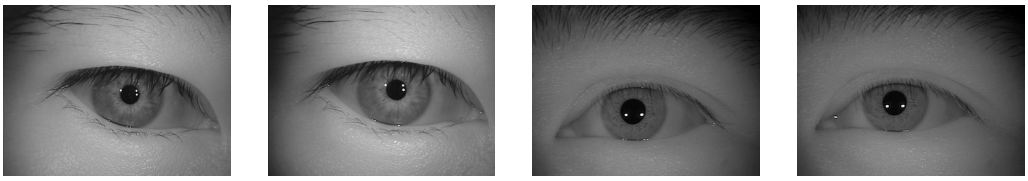
Table 3.2: CASIA Iris-V4 subsets with respective sensors

were very low. In the first case the results would bear some implications for future sensor construction and/or sensor use in establishing an iris database. It was assumed that this high variation could be caused by correlated data used to generate the sensors PRNU fingerprint. On the other hand Höller *et al.* [64] suspected that multiple sensors may have been used for the acquisition of the *CASIA Iris-V4* sub sets, wich are listed in Table 3.2 with the corresponding sensors used for the image acquisition.

3.2 Generation of Iris-Sensor PRNU Fingerprints From Uncorrelated Data

As mentioned in the previous section, Höller *et al.* [64] came to the conclusion that low variance between images used to generate the PRNU fingerprint of the imaging sensors might result in a poor fingerprint. To clarify this

question, uncorrelated data has to be used to generate the PRNU fingerprints for the specific sensors. Since the data sets do not contain such data, it had to be additionally acquired with the same sensors as used for the *CASIA-Iris V4* database [39]. Exemplary images from the *Lamp* and *Thousand* data set showing the correlated image content are can be seen in Figure 3.1.



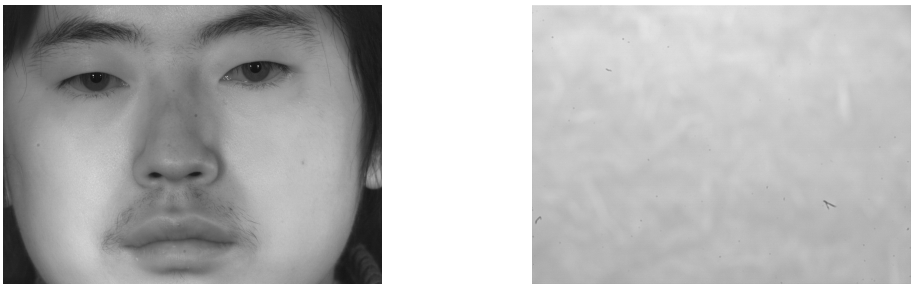
(a) (b) (c) (d)

Figure 3.1: Four exemplary images taken from the *Lamp* (a,b) and *Thousand* (c,d) data sets showing the high image content correlation in the *CASIA-Iris V4* database.

To obtain high-quality PRNU fingerprints, images with uncorrelated content and high saturation are needed. Generating the desired data with the biometric sensors is not trivial, since these sensors are not like cameras in a common way. The configuration options are often limited either by the sensor itself or by the software used to acquire the data. Usually these sensors are optimized to only take specific types of images and have a built-in quality assessment, which ensures that the acquired image satisfies defined constraints. To capture the uncorrelated data different materials, like paper sheets and plastic foil, were used to obtain uncorrelated out-of-focus images with high luminance. An examples for the acquired uncorrelated data for the *CASIA long-range iris camera* sensor is given in Figure 3.2 and additional information on the acquisition process for the different sensors can be found in [39].

As determined by Fridrich [19], the best images for estimating the fingerprint are those with high luminance (but not saturated) and small σ^2 (images with a smooth content). If the camera under investigation is available to the analyst, unsaturated out-of-focus images of bright cloudy sky would be the best.

The sensors used for the acquisition of uncorrelated images in [39] are: *CASIA long-range iris camera*, *OKI IRISPASS-h*, *Irisguard AD100* and *Irisguard*



(a)Image from *Distance* (b)Uncorrelated Data

Figure 3.2: Image taken from the *Distance* data set (a) and uncorrelated data acquired with the *CASIA long-range iris camera* (b).

H100 IRT. The acquired uncorrelated images have been used to generate a PRNU fingerprint for each of the previously mentioned sensors and two different device identification (DI) experiments have been conducted (similar to the DI experiment in [64]): the first one using correlated images (iris images from the various data sets) and the second one using the uncorrelated images to generate the PRNU fingerprints for the iris sensors that were available for the additional acquisition mentioned before.

The two DI experiments have been conducted for the following sensors: *CASIA close-up iris camera*, *OKI IRISPASS-h (1)*, *OKI IRISPASS-h (2)*, *CASIA long-range iris camera*, *Irisking IKEMB-100*, and *Irisguard H100 IRT*. Additionally, the two DI experiments have been divided into two separate groups: the first group contains only sensors also present in the CASIA-Iris V4 database and is subsequently denoted as "CASIA-Iris V4 Experiment", while the second group contains the other sensors not present in the "CASIA Iris V4 experiment", namely the *OKI IRISPASS-h (1)* and *Irisguard H100 IRT*. The latter group subsequently denoted as "2013 Iris Data Sets Experiment". The association between the sensors and data sets for the two groups of experiments is illustrated in Table 3.3.

The detailed set up for the experiments can be found in the paper [39], while the results are presented in the following section.

Data set	Sensor
Casia-Iris-Interval	CASIA close-up iris camera
Casia-Iris-Lamp	OKI IRISPASS-h (1)
Casia-Iris-Twins	OKI IRISPASS-h (2)
Casia-Iris-Distance	CASIA long-range iris camera
Casia-Iris-Thousand	Irisking IKEMB-100

(a)CASIA-Iris V4 Experiment

Data set	Sensor
Irispass-2013	OKI IRISPASS-h (1)
H100-2013	Irisguard H100 IRT

(b)2013 Iris Data Sets Experiment

Table 3.3: Association of the various data sets with the corresponding sensors for the CASIA-Iris V4 Experiment (3.3a) and the 2013 Iris Data Sets Experiment (3.3b).

3.3 Device Identification Results

The following section shows the results from [39] for the "CASIA-Iris V4 Experiment" in the subsection 3.3.1 and thereafter, the "2013 Iris Data Sets Experiment" in the subsection 3.3.2. Both subsections show the results using the correlated data to generate the PRNU fingerprints first, and the results with PRNU fingerprints generated from uncorrelated data afterwards.

3.3.1 CASIA-Iris V4 Experiment

Table 3.4 shows the first iteration of the CASIA Iris V4 experiment, where the images used to calculate the PRNU fingerprint for each sensor are taken from the data set of the given sensor, which contains correlated data. In comparison to the experiment from Höller *et al.* [64] the results are not exactly the same, but since the variations are small they are still comparable. It can be seen that for some sensor combinations, like *Twins* $\leftrightarrow$ *Distance* or *Twins* $\leftrightarrow$ *Thousand* the results in respect to the differentiability of the two

Data set	Interval	Lamp	Twins	Distance
Thousand	24.67	14.67	4.33	5.33
Distance	17.67	11.33	1.33	
Twins	17.67	15.33		
Lamp	25.33			

(a)EERs

Table 3.4: Equal error rates (a) for all possible sensor combinations of the CASIA-Iris V4 subsets with PRNU fingerprints generated from correlated data.

Data set	Interval	Lamp	Twins	Distance
Thousand	24.67	23.67	4.33	49.17
Distance	52.33	58.17	46.00	
Twins	17.67	24.83		
Lamp	30.83			

(a)EERs

Table 3.5: Equal error rates (a) for all possible sensor combinations of the CASIA-Iris V4 subsets with PRNU fingerprints generated from uncorrelated data.

sensors look very promising, while other combinations do not show such satisfactory results , e.g. *Interval* $\leftrightarrow$ *Thousand* or *Interval* $\leftrightarrow$ *Lamp*.

On the other hand, Table 3.5 shows the second iteration of the CASIA Iris V4 experiment, where the additionally acquired uncorrelated images have been used to calculate the PRNU fingerprint for the sensors *OKI IRISPASS-h (1)* and *CASIA long-range iris camera*. For all other sensors, the PRNU fingerprint was calculated as before from images of the respective data set.

For a better comparison between the two previous iterations, Table 3.6 shows the the absolute difference for the EER values between the first experiment iteration and the second. The values were calculated by subtracting the EER result of a specific sensor pair (S_i, S_j) , $i \neq j$, from the first iteration from the EER value of same sensor pair (S_i, S_j) of the second iteration. Therefore positive values indicate that the EER has increased and negative values that it has decreased.

Data set	Interval	Lamp	Twins	Distance
Thousand	0.00	+9.00	0.00	+43.84
Distance	+34.66	+46.84	+44.67	
Twins	0.00	+9.50		
Lamp	+5.50			

Table 3.6: Absolute differences of the obtained EER values between PRNU fingerprints generated from correlated (from Table 3.4) and uncorrelated data (from Table 3.5). Positive values indicate that the EER for the specific sensor pair has increased with the usage of uncorrelated data for the PRNU fingerprints, while negative values that the EER has decreased.

The results show that the generation of the PRNU fingerprints from the uncorrelated data has brought a general increase in the EER between all the different sensor combinations. The increase varies a lot between the different sensors. The sensor pairs having "0.00" as result used the same fingerprints as in the first iteration.

For the *CASIA long-range iris camera* respectively the *Distance* data set there has been an increase of up to almost 50%. Looking at the correlation scores from this sensor using the PRNU fingerprint estimated by the uncorrelated data, it can be seen that they are very low. This indicates that the images used to generate the fingerprint and the images in the *Distance* data set have been acquired with a different sensor. Another explanation for this behavior is that the two generated PRNU fingerprints do not correlate at all, which could also indicate that the images might be from two distinct sensors. Taking a closer look at the images from the *Distance* data set and the additionally acquired uncorrelated data shows that both contain pixel defects, but the positions of these defects do not match as illustrated in Figure 3.3. Furthermore the pixel defects from an arbitrary image in the *Distance* data set (acquired in 2009) cannot be found in any image with uncorrelated content (acquired in 2013), which excludes ageing effects since pixel defects do not recover over time. All these hints lead to the assumption, that the sensor used to acquire the uncorrelated data was not the same as the one used to acquire the *Distance* data set and therefore explains the very high EERs.

(a)*CASIA long-range iris camera* (b)*Distance* data set

Figure 3.3: Patch with size of 128x128 pixels extracted from the upper left corner of: (a) image from uncorrelated data acquired with CASIA long-range iris camera; (b) image from the *Distance* data set. The position of the pixel defects (marked with a circle) is varying between both images.

The EER for all comparisons including the *OKI IRISPASS-h (1)* sensors data used for the *Lamp* data sets PRNU fingerprint showed an increase of about 9-10%. Because the EER is already relatively high for the *Lamp* data set (see Table 3.4), Höller *et al.* [64] assumed that the data set could have possibly been acquired with different sensors from the same model, which could be further endorsed by the increase of the EER. On the other hand the uncorrelated data acquired with this sensor could have been insufficient to improve the quality of the sensors PRNU fingerprint. Using the fingerprint generated from the uncorrelated data, where it is verified that the exact same sensor has been used to acquire all the images, it could be investigated if the Lamp data set was acquired with different sensors.

3.3.2 2013 Iris Data Sets Experiment

The results of the 2013 Iris Data Sets Experiment in Table 3.7 show that using correlated data to calculate the PRNU fingerprints already leads to a EER of 0% for the two data sets under investigation. Using uncorrelated data to generate the PRNU fingerprints for both sensors does neither improve nor reduce the discriminability of the two sensors, since the EER does not change. Therefore the general increase in the *CASIA-Iris V4 Experiment* in 3.3.1 is not confirmed, which leads to the assumption that the increase of

Data set pair	Irispass-2013, H100-2013
EER corellated data	0.00
EER uncorellated data	0.00
CI corellated data	(0.00-0.00)
CI uncorellated data	(0.00-0.00)
EER absolute difference	0.00

Table 3.7: EERs, CI and absolute EER difference for the 2013 Iris Data Sets Experiment performed with the *OKI IRISPASS-h (1)* and emphIrisguard H100 IRT sensors. The *Irispass-2013* has been acquired with the *OKI IRISPASS-h (1)* and the *H100-2013* data set with the *Irisguard H100 IRT* sensor.

the EER in the *CASIA-Iris V4 Experiment* might come from the usage of different sensors during the acquisition of the images in the data sets and that the fingerprints generated from uncorrelated data could further reveal this by showing increasing EER results. To verify this assumption further investigation on the *CASIA Iris V4* data set is needed.

3.4 Summary

The results from [39] showed that using uncorrelated data to generate the PRNU fingerprint does not improve the discriminability between the various sensors of the *CASIA-Iris V4* database. Hence other factors must be causing the high variations in the results and the acquisition and usage of uncorrelated data for the PRNU fingerprint generation was not helpful in addressing the issue of the varying discriminative power. The use of uncorrelated data yielded to an increase of the EER for the respective sensors, varying from a negligible increase of 0-1% to an increase of up to 50%. Because it is not verified whether each data set has been acquired using only a single sensor or multiple instances, it is difficult to interpret the results, but hints have been found that the multiple sensors could have been used. It is possible that the usage of uncorrelated data does not bring any benefit in this case because the estimated fingerprints have already been accurate enough.

Further investigation on the *CASIA-Iris V4* database is needed, to be able to explain the variations in the discriminability results for the various sensors. Because of the hints found in [64] and [39] suggesting that multiple sensors of the same model might have been used for the acquisition of the datasets and the fact that the documentation of the *CASIA-Iris V4* database does not contain any information on the number of sensors, but only the model used for each data set. Because for some sensors the acquisition of uncorrelated data has been very challenging and less effective, as further described in [39], other techniques have to be investigated to be able to deal with the correlated content of the images in this context.

The motivation of this thesis is to use forensic investigation techniques to detect the presence of images from multiple sensors in the data sets, or to confirm that all the images in each data set have been acquired using a single sensor. In addition, in the case that images from multiple sensors are detected, regrouping of images coming from the same sensor and hence establishing a ground truth for the *CASIA-Iris V4* database is intended in this work. To address the issue with the highly correlated image content in the data sets, the application of PRNU enhancing techniques described in section 2.5 are included in this work and the impact is evaluated.

4 Related Work for Classification of Images from Unknown Source

The forensic investigation in this work is dealing with a data set, where no information on the number of source devices used to acquire the images is known and can therefore be seen as a blind classification of the image source in an open set scenario.

Related work regarding this scenario proposes various clustering methods, which are summarized in the following sections. Further details can be found in the corresponding papers.

4.1 Unsupervised Classification of Digital Images Using Enhanced Sensor Pattern Noise

In his work Li [45] proposes an unsupervised image classifier, which is capable of clustering images taken by an unknown number of unknown digital cameras into a number of classes, each corresponding to one camera. The classifier training algorithm, which is used to perform this task, is given as follows:

- **Step 0.** Initialization:
  Prepare the dataset, specify the sizes of image blocks and the training set.
- **Step 1.** Extract and enhance the fingerprint of each block cropped from the images
- **Step 2.** Establish the $M \times M$ similarity matrix ρ for the training set
- **Step 3.** Execute the classifier trainer based on the similarity matrix ρ. For each fingerprint
  - **3.1.** Assign a unique random class label
  - **3.2.** Calculate a reference similarity
  - **3.3.** Establish a membership committee
  - **3.4.** Update the class label iteratively until the stop criteria are met
- **Step 4.** Classify the rest of the dataset using the trained classifier

In the *Initialization Step (Step 0)* two sets of images, S_1 and S_2, are created. The set S_1 contains all the images under investigation, where horizontal images are left intact and vertical images are rotated by 90 degrees in clockwise direction. S_2 is then created by rotating each image in S_1 by 180 degrees. Furthermore in this step different parameters for the PRNU extraction and the size of the training set are defined.

Step 1 extracts the sensor pattern noise (SPN) from the center of each image in S_1 and S_2 according to Fridrich [19] with the size defined in Step 0. The extracted SPN is then enhanced according to equation 4.1 proposed in [44], where $n(i,j)$ is the initial coefficient in the DWT domain, $n_e(i,j)$ is the enhanced coefficient and α is the threshold at which n is attenuated.

$$n_e(i,j) = \begin{cases} e^{-0.5\frac{n^2(i,j)}{\alpha^2}}, & \text{if } 0 \leq n(i,j) \\ -e^{-0.5\frac{n^2(i,j)}{\alpha^2}}, & \text{otherwise} \end{cases} \tag{4.1}$$

In *Step 2* a training set with M images and a corresponding $M \times M$ similarity matrix ρ is established, with $\rho(i,j)$ indicating the similarity between the

SPNs i and j. The similarity between any two enhanced SPNs i and j is calculated using

$$\rho(i,j) = \frac{(n_i - \overline{n_i}) * (n_j - \overline{n_j})}{\|n_i - \overline{n_i}\| * \|n_j - \overline{n_j}\|}, i, j \in [1, M] \tag{4.2}$$

To calculate the similarity between two SPNs i and j, four combinations resulting from the two data sets S_1 and S_2 have to be taken into account. The maximum of the four combinations is then selected as the $\rho i, j$ and $\rho j, i$.

During the classifier training in *Step* 3 each SPN is treated as a random variable and an optimal class label d_k is assigned in an iterative manner until the stop criteria are met. For each SPN i (i.e., each row of the similarity matrix ρ), a simple k-Means clustering method is used to cluster the $M-1$ similarity values between i and the rest of the training set into two groups (one as intra-class and the other inter-class) and the average of the centroids of the two clusters is taken as a reference similarity, r_i. To determine the class label for each SPN i, a membership committee C_i is established with c SPN members from the training set that are most similar to i. The trainer updates the label of each SPN iteratively based on the cost (similarity) of SPN i in respect to the members of the membership committee C_i. This optimization problem is solved using a Markov random fields approach. When no changes of class labels occur after x iterations, the stop criterion is met.

Step 4 uses the centroids of the image classes determined by the classifier trained in *Step* 3 to cluster the remaining images not present in the training set. To classify an image i, its similarity to each centroid is calculated and the class label is assigned according to the centroid closest to the image i. During the image clustering process, the centroids of the classifier can either be fixed throughout the entire process or updated when new images are assigned the corresponding classes. The update is performed by recalculating the average SPN when a new image is assigned to a class. Another option is that if the similarity between a SPN and the most dissimilar centroid is less than a threshold set by the user, a new class with the SPN as the founding member can be created and allowed to attract new members just like the classes identified by the trainer.

To demonstrate the performance of the proposed classification system, it has been tested on a dataset consisting of 1200 photos of 1536×2048 pixels taken by six cameras, each responsible for 200. The proposed classifier has been able to perform at the quality level as accurate as 1.222% error rate without the user providing any additional information about the dataset.

4.2 Fast Image Clustering of Unknown Source Images

In their paper Caldelli *et al.* [6] propose a new technique which aims at blindly clustering a given set of N digital images. The technique is based on the technique by Li [45], presented in the previous Section 4.1, and improves it both in terms of error probability and of computational efficiency. The system is able, in an unsupervised and fast manner, to group images without any initial information about their membership. The sensor pattern noise (SPN) is extracted for each image as reference and the following classification is performed by means of a hierarchical clustering procedure. Hierarchical clustering generates a hierarchy of clusters which may be represented by a tree-like two-dimensional structure known as dendrogram, which illustrates the fusions (agglomerative clustering or bottom-up) or divisions (divisive clustering or top-down) made at each stage of analysis: The root of the dendrogram is a single cluster containing all the elements, and the leaves correspond to the individual elements. The proposed agglomerative hierarchical clustering procedure over a set of N elements produces a series of partitions of the elements $P_0, \ldots, P_{N-1}$, where the first partition P_0 consists of N single object clusters, while the last partition P_{N-1} consists of a single group containing all the N elements. The procedure merges pairs of clusters at each step until all clusters have been merged into a single one that contains all the elements.

The SPN for each image is extracted according to the procedure proposed by Fridrich [19] and, similar to the approach of Li [45] from the previous Section 4.1, the SPN is enhanced after the extraction. The enhancing function gives larger weighting factors to the weak components of the SPN in the DWT domain, and smaller weighting factors to large components, and it is

described in equation 4.3, where $n(i,j)$ is the initial coefficient in the DWT domain, $n_e(i,j)$ is the enhanced coefficient and α is the threshold at which n is attenuated.

$$n_e(i,j) = \begin{cases} 0, & \text{if } n(i,j) < -\alpha \\ -cos(\frac{n(i,j)\pi}{2\alpha}), & \text{if } -\alpha \leq n(i,j) \leq 0 \\ cos(\frac{n(i,j)\pi}{2\alpha}), & \text{if } 0 < n(i,j) \leq \alpha \\ 0, & \text{if } n(i,j) > \alpha \end{cases} \tag{4.3}$$

Hierarchical clustering does not require a pre-specified number of clusters. However, a partition of disjoint clusters just as in flat clustering is desired: In this case, the hierarchy needs to be cut at some point. The criterion based on the silhouette coefficient has been used to determine the cutting point. The use of the silhouette coefficient combines both the measures of cohesion (similarity inside clusters) and separation (similarity among disjoint clusters). For each SPN n_i, the coefficient s_i is simply calculated as shown in equation 4.4. The cohesion, a_i, is the average correlation of n_i to all other SPNs in the same cluster, while the separation, b_i, is the correlation of n_i to all other SPNs in each of the other clusters, taking the average over all clusters.

$$s_i = b_i - a_i \tag{4.4}$$

This calculation is performed at each loop of the algorithm and for each SPN. At iteration q a global measure of the silhouette coefficient SC_q is calculated, as shown in equation 4.5, by averaging the coefficients related to each SPN that belongs to a certain cluster and taking the average value with respect to all the current K-clusters, which range from N to 1.

$$SC_q = \frac{1}{N} \sum_{i=1}^{N} s_i \tag{4.5}$$

The minimum coefficient SC_q is found over the $N-1$ iterations, resulting in q^* being the last iteration that has to be executed. The partition of the images P_q is saved for each iteration, hence the optimal clustering P_{q^*} can be selected for the iteration q^*. The pseudo-code for the hierarchical clustering algorithm is given as follows:

1. Initialization: $K \Leftarrow N$, calculate similarity matrix $H \in \mathbb{R}^{N \times N}$
2. Loop over $q \Leftarrow 1$ to $N - 1$
   a) Search for the pair of clusters $\langle U, V \rangle$ that have the largest similarity
   b) Delete the rows and columns of H referred to the clusters $\langle U, V \rangle$
   c) Update H by calculating the new similarity values between the new cluster $Z \Leftarrow \langle U, V \rangle$ and the remaining clusters
   d) $K \Leftarrow K - 1$
   e) Calculate the silhouette coefficient SC_q
   f) Save the current partition P_q
3. Calculate the minimum value of the silhouette coefficients: $q^* \Leftarrow min_q(SC_q)$
4. Get the optimal partition by selecting the one relative to the iteration q^*, which is the partition P_{q^*}

At the end of the clustering procedure, the number of clusters M is obtained, that is supposed to be exactly the real number of devices which generated the given N images of the training set. For each of the obtained M clusters, a reference SPN is calculated (as the centroid of the cluster) simply by averaging all the SPNs belonging to that cluster. The centroids of the clusters provided by the mentioned procedure are then used as the trained classifier to group the images belonging to the test set. The final classification of the images consists of comparing the similarity of the current image (taken from the data set under investigation) to each of the centroids, and then the image is assigned to the cluster whose centroid exhibits the largest similarity.

The experimental results obtained on a dataset, containing 1200 images at different resolutions (from 3MP to 12MP) taken by six cameras in different periods of time, confirm that the proposed hierarchical clustering technique permits both to reduce computational complexity and to improve the true positive rate (TPR) for image grouping, because it works with with larger PRNU patch sizes as compared to the method [45] presented in Section 4.1 and with a feasible computational time. It also showed satisfactory results

for non-uniform datasets, where each cluster contains a different number of images, demonstrating the adaptability of the method to a real world scenario.

4.3 Blind image clustering based on the Normalized Cuts criterion for camera identification

Amerini *et al.* [3] propose a methodology based on Normalized Cuts (NC) criterion and evaluate it in comparison with other state-of-the-art techniques, such as Multi-Class Spectral Clustering (MCSC) and Hierarchical Agglomerative Clustering (HAC).

Spectral clustering techniques make use of the spectrum (eigenvalues) of the similarity matrix of the data to perform a dimensionality reduction before the clustering is performed with the lower dimensional features. Multi-Class Spectral Clustering techniques incorporate these properties, but also have some drawbacks: Due to the random initialization the performance of the clustering may vary and a criterion for selecting the best number of K clusters is needed. To overcome these inconveniences, Amerini *et al.* [3] proposed another spectral clustering method, named Normalized Cuts method, which does not require any knowledge of the number of expected clusters and does not present any randomness in performance. The proposed method fits the problem of blind image clustering well because it does not require a priori knowledge of the amount of classes in which the dataset has to be divided, it solely needs a stop threshold.

Given a graph $G = (V, E)$, the edges E connecting each pair of nodes/images V are weighted by means of a chosen similarity function $w(i, j)$, with i and j being two nodes of the graph. The graph G is partitioned into two disjoint graphs A and B ($A \cup B = V, A \cap B = \emptyset$) by simply removing edges connecting the two parts. The total weight of the removed edges in this partitioning step gives a computation of the degree of dissimilarity between these two disjoint partitions and is called the *cut*. Its value can be determined as shown

in equation 4.6.

$$cut(A,B) = \sum_{u \in A, v \in B} w(u,v) \tag{4.6}$$

The optimal bipartition of a graph is obtained by means of the minimization of the *cut* value. The authors proposed the *normalized cut (Ncut)*, defined in equation 4.7, as dissimilarity measure between two disjoint graphs A and B.

$$Ncut(A,B) = \frac{cut(A,B)}{assoc(A,V)} + \frac{cut(A,B)}{assoc(B,V)} \tag{4.7}$$

The association measure $assoc(A,V)$, representing the total connections from nodes in A to all nodes in the graph (V), is defined as:

$$assoc(A,V) = \sum_{u \in A, t \in V} w(u,t) \tag{4.8}$$

The minimization of the *normalized cut* is NP-complete, therefore the authors propose a discrete approximation to solve this problem as follows, where the graph $G = (V,E)$ is to be partitioned into two sets A and B.

Let x be an indicator vector of dimension $N = |V|$, where $x_i = 1$ if the node i belongs to A and -1 if it does not belong to A and W is the similarity matrix. Let $d(i) = \sum_{j \in V} w(i,j)$ be the total connection from node i to all other nodes and D an $N \times N$ matrix with d on its diagonal. On the basis of these assumptions the problem of minimizing the *normalized cut* can be rewritten as:

$$\begin{aligned} \min_x Ncut(x) &= \min_y \frac{y^T(D-W)y}{y^T D y} \\ \text{where } y = (1+x) - b(1-x) \text{ and } b &= \frac{\sum_{x_i>0} d_i}{\sum_{x_i<0} d_i} \end{aligned} \tag{4.9}$$

If y is relaxed to take on real values, equation 4.9 can be minimized by solving the generalized eigenvalue system in equation 4.10.

$$(D-W)y = \lambda D y \tag{4.10}$$

Summarizing, the partitioning method can be described as follows:

1. Given a set of images, set up a weighted graph $G = (V, E)$, compute the weight on each edge which measures the similarity between the SPNs of two images.
2. Solve equation 4.10 for eigenvectors with the smallest eigenvalues.
3. Use the eigenvector with the second smallest eigenvalue to bipartition the graph by finding the splitting point such that *Ncut* is minimized.
4. Decide if the current partition should be subdivided recursively by checking the stability of the cut.

The decision in the last step (step 4), wether the current cluster should be further subdivided or not, is taken by comparing the *Ncut* value to a prespecified threshold. For this purpose an *aggregation coefficient (AC)*, which is computed for each cluster, has been defined (see equation 4.13). The cluster is then further subdivided if the AC value is lower then the pre-defined threshold T_h.

$$AC(k) = \frac{1}{N_k} \sum_{i,j} w(i,j) \tag{4.11}$$

The *aggregation coefficient (AC)* simply calculates the mean value over all weights among nodes N_k of a cluster. The pre-calculated threshold T_h has been experimentally determined considering five diverse sets of images whose acquisition cameras were known to the authors, where a threshold value of $T_h = 0.037$ has been obtained.

According to the authors the proposed technique has been able to outperform other state-of-the-art techniques, as the technique described in Section 4.2, in terms of accuracy and computational effort.

4.4 Silhouette Coefficient Based Approach on Cell-Phone Classification for Unknown Source Images

Luan *et al.* [51] propose a silhouette coefficient based algorithm for source cell-phone classification, similar to the approach of Caldelli *et al.* [6] presented in Section 4.2.

A graph based approach is chosen by the authors, which changes the source cell-phones identification into a graph partition issue where the sensor pattern noise SPN of N images are considered as nodes V and the edges E represent the similarity between two SPNs in a graph $G = (V, E)$. The edges E are weighted according to the similarity value between the SPNs of two images. The weight for each edge $\omega_{i,j}$ in the similarity matrix W with $N \times N$ for two SPNs r_i and r_j is defined as:

$$\omega_{i,j} = \begin{cases} 0, \text{ if } corr(r_i, r_j) < 0 \\ corr(r_i, r_j), \text{ otherwise} \end{cases} \tag{4.12}$$

The normalized cross correlation (NCC) $corr(r_i, r_j)$ of the two SPNs r_i and r_j, used as similarity measure, is defined as:

$$corr(r_i, r_j) = \frac{(r_i - \overline{r_i}) * (r_j - \overline{r_j})}{\|r_i - \overline{r_i}\| * \|r_j - \overline{r_j}\|} \tag{4.13}$$

The authors use the multi-class spectral clustering algorithm of Amerini *et al.* [3] to calculate the eigenvalues and eigenvectors of the affinity matrix W to obtain a $N \times L$ indication matrix X. The membership of each node in the L subsets v_j is defined by:

$$X_{i,j} = \begin{cases} 1, \text{ if } V_i \in v_j \\ 0, \text{ otherwise} \end{cases} \tag{4.14}$$

The clustering algorithm follows an iterative approach where the partition number L and partition P_L are recorded for each iteration, and the optimal

partition of the graph P_L^* is selected by calculating the silhouette coefficient SC_L according to the approach of Caldelli *et al.* [6] as described in Section 4.2.

The experiments performed with the proposed approach have been conducted using images from 5 different cell-phones with average accuracies between 77% and 100%.

4.5 Blind Camera Fingerprinting and Image Clustering

Bloy [21] proposes a Blind Camera Fingerprinting and Image Clustering (BCFAIC) technique, which performs an agglomerative clustering to construct PRNU fingerprints from a mixed set of images, enabling identification of each images source camera without any prior knowledge of the source. This technique does not rely on a known training set, test set or ground truth. It solely depends on a pre-calculated threshold function.

The author applied the PRNU extraction approach of Fridrich [19] to see if a fingerprint for a given camera could be developed from a large set of uncontrolled and unknown images, some of which come from the given camera and some of which do not. The goal is, given a large database of images, which contains 50 or more shots from a number of cameras, to group together at least 50 images from each camera so that a PRNU fingerprint can be developed.

The clustering algorithm makes use of a threshold function t, which is estimated making use of the inter sensor distances. The inter sensor distances are determined by generating n-image PRNU fingerprints from images of a given sensor and calculating the correlation with single images from a different camera, for $n = 1 \ldots 50$. From these correlation results for each n, the standard deviation is calculated and three times the standard deviation is added to the maximum observed inter sensor correlation value for each n. Finally, a quadratic curve is fitted to the resulting values for each n, which is increasing as a function of n and results in a very conservative threshold function t, which is given in equation 4.15. The author determined the

function t using images from four different *Canon PowerShot A530* consumer cameras.

$$t = -0.00002438n^2 + 0.0002889n + 0.009 \tag{4.15}$$

Using the previously calculated threshold function t an automatic clustering algorithm to separate a mixed set of images into subsets of images that are acquired with the same camera performs the following steps:

1. Randomly select pairs of images until a pair is found whose noise correlation exceeds $t(1)$; average the PRNU of this pair to form a fingerprint.
2. Perform the first pass: for each remaining image, correlate the PRNU with the fingerprint. When the correlation value exceeds t(# of images in fingerprint cluster), average (cluster) it into the fingerprint. When $n = 50$ images have been averaged into the fingerprint or all images have been tried, stop and go to Step 3.
3. Perform the second pass: loop over all the unclustered images a second time, correlating with the current fingerprint and adding those that exceed the threshold. (Do not average more than 50 images into the fingerprint but allow more than 50 to be associated with the fingerprint.)
4. Repeat Step 1. Give up when Step 1 has tried 1000 pairs without success.

The result of this algorithm is a list of partitions, containing a PRNU fingerprint and associated images that are matching to it. Every partition is representing a different sensor. If an image did not match with any of the fingerprints in the list, it is labeled as unassociated image. Images that have been associated with the wrong sensor respectively cluster are labeled as mismatched, which applies when a partition contains only a few images from a certain sensor while the majority of the images belong to another sensor. Clearly such mismatched images can only be categorized if the association of each image to a specific sensor is known and not in a “blind” scenario.

Four identical instances of a typical consumer camera were tested, and only the default JPEG output was used in order to make the task more difficult. In order to increase the difficulty further, a highly textured set of outdoor imagery was collected for testing in hopes that the texture would reduce the effectiveness of the denoising filter. The goal was to use automatic clustering to generate a fingerprint by associating at least 50 images from each of the four cameras out of a combined set of images that included not only the four cameras in question but several other makes and models as well.

The clustering technique proposed by the author performs almost perfectly on a 200 image indoor test set in that no image is included in an incorrect fingerprint and almost all images from the same camera are grouped together. Also performing a more challenging test with highly textured outdoor imagery no images were mistakenly associated with the wrong fingerprint.

This algorithm has been reimplemented in this work because of the promising results on the open set scenario without any need of a priori knowledge about the number of distinct cameras respectively sensors and any other constraints.

5 Novel Forensic Techniques for Unknown Source Sensor Detection

This chapter proposes various novel forensic techniques which can be used to detect the presence of images from multiple sensors in a data set without any a priori knowledge on number of the sensors. All of the following techniques are based on the sensors PRNU, which is extracted from the images under investigation. There are no special requirements to the data set under investigation, except that the images must have been acquired with a digital camera and hence contain a PRNU signal. If the images are computer generated renderings, this signal would not be present in the images. Since the PRNU extraction works with grayscale images, color images need to be converted into grayscale images first.

The novel forensic techniques consist of the four proposed techniques in [38]:

- K-Means Clustering (KM)
- PCA K-Means Clustering (PCAKM)
- Sliding Window Fingerprinting (SWFP)
- Device Identification On Dataset Partitions (DIODP)

A detailed description of each technique is given in the following sections 5.1 to 5.4.

5.1 K-Means Clustering

The second technique, the K-Means Clustering (KM), uses the *k-means++* algorithm by Arthur and Vassilvitskii [65] as implemented in Matlab.

The idea behind K-Means Clustering is to partition the data into k distinct clusters, where it finds a partition of the data in which objects within each cluster are as close as possible to each other, and as far from objects in other clusters as possible. The distance measure decides how close or rather how far two objects are. Each cluster in the partition is defined by its member objects and by its centroid, or center. The centroid for each cluster is the point to which the sum of distances from all objects in that cluster is minimized. The minimization of the overall sum of distances is performed by a two-phase iterative algorithm.

The two phases of the iterative algorithm are:

1. **Batch updates**: Each iteration reassigns all objects at once to their nearest cluster centroid, followed by a recalculation of the cluster centroids. This phase mostly does not converge to a local minimum. A local minimum would be the partition of the data, where moving any object to another cluster would increase the total sum of distances. Hence this phase is only approximating a solution in a fast manner and acts as a starting point for the second phase.
2. **Online updates**: Each iteration, which passes through all the objects, reassigns individual objects to other clusters if they reduce the overall sum of distances. After each reassignment, a recalculation of the cluster centroids is performed like in the first phase. In contrast to the first phase, this phase is able to converge to a local minimum, although there might be other solutions with a lower total sum of distances. Finding this global minimum is usually done by using several replicates and choosing random starting points for each replicate.

The result of this algorithm is a set of clusters that are as compact and well-separated as possible. In this work the PRNU noise residuals of the images in the investigated data set are defined as the n objects to cluster, while the clusters k represent the different sensors. Since the number of

sensors is unknown, the clustering is repeated with a varying value for the number of clusters k.

In order to qualitatively evaluate the outcome of the clustering the Mean Silhouette Value (MSV) by Rousseeuw [60] has been chosen. The silhouette value for each point is a measure of how similar that point is to points in its own cluster, when compared to points in other clusters. The result for $k = 1$ has been determined by calculating the pairwise NCC between all point combinations i and j, where $i \neq j$, and then calculating the mean correlation over all points. For all $k >= 2$ the Mean Silhouette Value for the i-th point, S_i, is defined as

$$MSV = \frac{1}{N} \sum_{n=1}^{N} \frac{b_i - a_i}{max(a_i, b_i)} \tag{5.1}$$

where N is the number of noise residuals, a_i is the average distance from the i-th point to the other points in the same cluster as i, and b_i is the minimum average distance from the i-th point to points in a different cluster, minimized over clusters. The silhouette value ranges from -1 to $+1$. A high silhouette value indicates that i is well-matched to its own cluster, and poorly-matched to neighboring clusters. If most points have a high silhouette value, then the clustering solution is appropriate. If many points have a low or negative silhouette value, then the clustering solution may have either too many or too few clusters.

5.2 PCA K-Means Clustering

The PCA K-Means Clustering (PCAKM) uses the same clustering algorithm as previously mentioned in 5.1, the only difference is that in this technique the clustering objects n consist of principal components generated by a Principal Component Analysis (PCA) [18] on the PRNU noise residual of each image.

The PCA is a powerful tool to identify patterns in high dimensional data and therefore it helps for analyzing such data. It is a statistical procedure

that performs an orthogonal transformation on a set of possibly correlated observed variables, which results in a new set of variables that are orthogonal to each other, so that redundant information is excluded. These newly generated variables are called principal components and their number is less or equal to the number of original variables and each principal component is a linear combination of the original variables. The principal components resulting from the transformation have a key characteristic: The first principal component covers the largest variance in the original data and each following component also covers the largest possible variance under the constraint that it is orthogonal to all the previous components. Therefore, the principal components as a whole form an orthogonal basis for the space of the data. Because of this special characteristic the variance in the original data can be represented by a smaller number of principal components, without losing much information.

The PCA is used in this technique to reduce the dimensionality of the PRNU noise residual (which can be seen as a high dimensional variable) by representing it with its first 5 principal components. This reduction has a huge impact on the performance of the K-Means clustering, because both memory usage and computational effort are reduced drastically, since the objects to cluster have a significantly lower dimension.

5.3 Sliding Window Fingerprinting

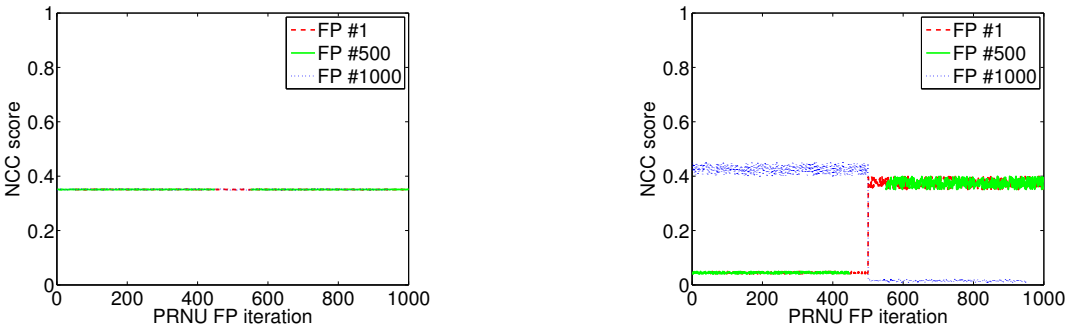
The Sliding Window Fingerprinting (SWFP) technique consists of a so called "sliding window" with an arbitrary but fixed size n that moves over a data set image by image. This novel forensic technique uses an iterative algorithm which performs the following steps:

1. Start at image with index $i = 0$.
2. Gather images inside the sliding window with size n, hence the images with index $i \dots i + n$.
3. Extract the PRNU noise residual for each image.
4. Compute a PRNU fingerprint using the images inside the window.
5. Increment the index i by 1.
6. Repeat step 2 until all the images have been used to calculate a PRNU fingerprint.

Moving the window over the whole data set yields a list of PRNU fingerprints, which have been computed using sequential overlapping windows. For a data set containing m images, $m - n$ PRNU fingerprints are generated. After generating the fingerprints, the similarity of a PRNU fingerprint FP_i from the iteration i with all other fingerprints FP_j where $i \neq j$ is computed by calculating the NCC score of each fingerprint pair. This leads to a similarity matrix with size $(m - n) \times (m - n)$ containing all the NCC scores.

The idea behind this method is to find "jumps" in the resulting similarity values between the PRNU fingerprints, which are referred to as *transitions* in the data. If all images in the data set have been acquired using a single sensor, the NCC scores should all be quite similar and there should be no recognizable transitions in the data. On the other hand, if images from multiple sensors are present in the database and these images occur in short sequences, the correlation between PRNU fingerprints generated with data from a single sensor and data from an other sensor (or even mixed data from both sensors) should decrease and therefore a transition between both sensors should be recognizable.

Figure 5.1 shows the output of this method on exemplary data, which in the first case contains images from a single sensor and in the second case images from two distinct sensors. The plot shows three different curves, that are representing the NCC correlation scores of three different fingerprint iterations with all other fingerprints. The gaps in the curves have been introduced because the correlated PRNU fingerprints contain common



(a)1 sensor case (b)2 sensor case

Figure 5.1: Sample Output for the SWFP technique; (a) shows the output for images from one sensor, while (b) shows the output for images from two sensors. Subfigure (b) shows a well recognizable transition at iteration 500: The NCC scores increase for the PRNU fingerprint with iteration 1, while they drop for iteration 1000. This transition shows the alteration from the first sensor to the second sensor in the sample data.

images used for their generation, which yields to a much higher NCC score in depending on the number of common images. Hence, the plotted results show solely NCC scores of PRNU fingerprints that have been generated using distinct images. The dashed (red) line shows the NCC scores for the fingerprint with iteration 1, the solid (green) line for iteration 500 and the dotted (blue) line for iteration 1000. For the case of a single sensor, no clear transitions can be recognized as described before. For the two sensor example, on the other hand, the output shows clear transitions for the various PRNU fingerprint iterations.

This technique has the disadvantage that it is not able to detect single images from other sensors, but is limited to sequences of images where the number (of images) is preferably larger as n (size of the sliding window). To bypass this limitation, this technique has been applied using various window sizes. Again this is only possible to a certain extent since using too few images to generate the PRNU fingerprint lead to a weak estimate according to Fridich [19].

5.4 Device Identification on Dataset Partitions

A device identification experiment, as performed by Höller *et al.* in [64], identifies whether different data sets have been acquired using different cameras. This is done by assuming that each data set has been acquired with a different sensor and calculating the intra- and inter-sensor similarities for each data set combination. The resulting similarity scores are used to calculate the EER, which is low when the two data sets under investigation are acquired with different sensors and high when they are acquired with the same sensor.

In this work it is investigated whether a single data set contains images from one or more sensors, hence a method to apply the device identification approach to a set of images of unknown source has to be found. The Device Identification on Dataset Partitions (DIODP) solves this issue by dividing the data set into n disjoint partitions with the same size and assumes that each partition has been acquired with a different sensor.

The first step for the device identification is to calculate a PRNU fingerprint for each of the n partitions. The images inside each partition are randomly shuffled, so that the PRNU fingerprint is not just computed from the first images in the partition. We use half of the images (up to a maximum of 50) to calculate the PRNU fingerprint and the remaining images in the partition to calculate the NCC scores (inter- and intra-partition scores). From these scores the pairwise EER for two partitions P_i and Pj, where $i \neq j$, is calculated as illustrated in Figure 5.2. If the resulting EER score is low (e.g. 0%), the extracted PRNU and respectively the PRNU fingerprint is different for both partitions, hence they must have been acquired with different sensors. On the other hand, if the resulting EER score is high (e.g. 50%), the extracted PRNU is very similar for both partitions and their images have likely been acquired with the same sensor.

To determine if images from different sensors are present in the data set, the distribution of the EER scores is analyzed. If it contains mostly or solely high EER values (above 40%), the data set is likely to contain images from a single sensor. Occurrences of very low EER values (below 10%) indicate that the data set might contain data from multiple sensors, because the assumption that all partitions are acquired with different sensors holds for

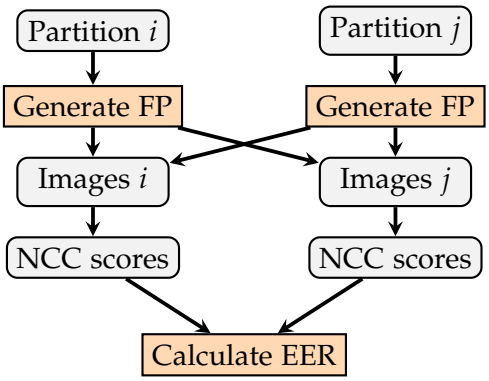
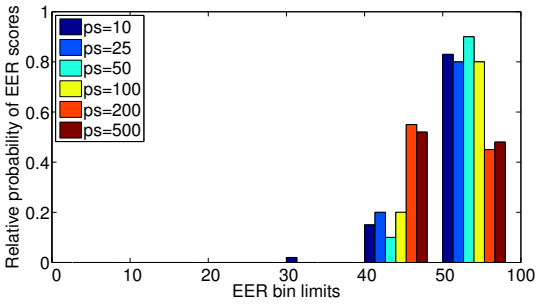


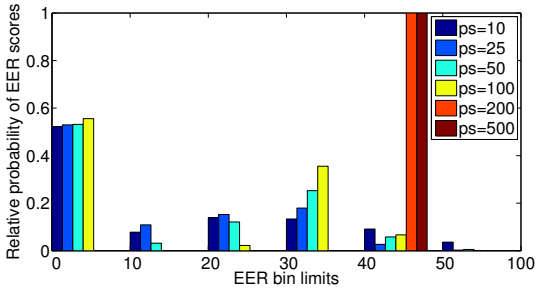
Figure 5.2: Calculation of the EER score of two disjoint partitions i and j.

at least some of the partitions. The more such low EER values show up, the more likely the dataset is containing images from multiple sensors. The resulting EER scores have been pooled into several bins in the output plot for better visibility. Six bins with the following limits have been defined for the EER scores: below 10%, between 10% and 19%, between 20% and 29%, between 30% and 39%, between 40% and 49% and above 50%.

An example of the distribution for different partition sizes in the case of a single sensor and two sensors is given in Figure 5.3. As mentioned before, the single sensor case contains mostly EER scores above 40% for this example, while the two sensor case shows a significant amount of low EER scores below 10% for most partition sizes. This technique is capable of pointing out whether a single sensor or multiple sensors have been used to acquire a data set, but it is difficult to estimate the exact number of sensors from the results. To overcome the problem with varying sequences of images from multiple sensors, the technique has been applied using different partition sizes. If any of the EER score distributions for the varying partition sizes shows a significant number of very low EER scores, the data set is suspected to contain images from multiple sensors.



(a)1 sensor case



(b)2 sensor case

Figure 5.3: Sample Output for the DIODP technique; (a) shows the output for one sensor, while (b) shows the output for two sensors. Both EER score distributions are clearly distinguishable, where (a) shows mostly very high EER scores and (b) shows a significant amount of very low EER scores.

6 Data Sets

In this work a forensic investigation on the CASIA-Iris V4 Database [1] has been conducted. The assumption, that all images in each sub set have been acquired with a distinct single sensor unit, is derived from the documentation of the database, but it is not explicitly noted that only one sensor instance of the same model has been used to acquire the images for the various subsets. Because of different hints on the possible presence of images from multiple sensors in the sub sets, as described in Chapter 3, this work addresses this question by applying the forensic techniques presented in Chapter 5. Additionally, to be able to evaluate the implementations and algorithms of the forensic techniques under controlled conditions, a test data set with images from known sensors has been generated.

A description of all data sets and their respective sub sets is given in the following sections.

6.1 CASIA Iris-V4

The CASIA-Iris V4 Database contains a total of 54601 iris images from more than 1800 genuine and 1000 virtual subjects. All iris images are 8 bit grey-level JPEG files, collected under near infrared illumination. For this work we used images from five different sub sets, which are shown in Table 6.1. The sixth sub set of the CASIA-Iris V4 Database, *CASIA-Iris-Syn*, has been discarded because it contains automatically synthesized iris images, which have not been acquired using a digital image sensor and do not contain a genuine PRNU signal.

[1] CASIA Iris Image Database, `http://biometrics.idealtest.org/`

Subset name	Short name	Sensor	Resolution
CASIA-Iris-Interval	intv	CASIA close-up iris camera	320×280
CASIA-Iris-Lamp	lamp	OKI IRISPASS-h	640×480
CASIA-Iris-Twins	twin	OKI IRISPASS-h	640×480
CASIA-Iris-Distance	dist	CASIA long-range iris camera	2352×1728
CASIA-Iris-Thousand	thou	Irisking IKEMB-100	640×480

Table 6.1: Subsets of the CASIA-Iris V4 database used in this work.

(a) (b) (c) (d) (e)

Figure 6.1: Sample images of the CASIA-Iris V4 database: (a) *intv*, (b) *lamp*, (c) *thou*, (d) *dist* and (e) *twin*.

For the CASIA Iris V4 data sets it is not clear, whether the single data sets have been acquired with a single sensor or if multiple instances of the same sensor model have been used. The documentation does not explicitly denote the number of sensors, but only the sensors manufacturer and model. Furthermore the same sensor model was used for the acquisition of two different data sets (OKI IRISPASS-h sensor for *lamp* and *twin*), which also confirms the assumption of the use of multiple sensors.

6.2 Test Data Set

The sensors used to generate the Test data set are an *OKI IRISPASS-h* and an *Irisguard H100 IRT* iris sensor. The images have been acquired under controlled conditions during a Short Term Scientific Mission (COST-STSM-IC1106-12803) funded by the COST Action IC1106[2] "Integrating Biometrics and Forensics for the Digital Age" at the NLPR-CASIA in Beijing (CN) in

[2] http://www.cost.eu/COST_Actions/ict/Actions/IC1106

2013 in collaboration with the local researchers, hence it is known that all images inside a sub set have been acquired using a single sensor. The 1000 images used for the following data sets have reliably been acquired with the two mentioned sensors, 500 with each one. All images are 8 bit grey-level JPEG files which have been acquired under near infrared illumination.

(a) (b)

Figure 6.2: Sample images from the *OKI IRISPASS-h* (a) *Irisguard H100 IRT* (b) iris sensors.

Two sub sets have been generated using those images for the evaluation of the forensic techniques, each containing images from both sensors and with a total of 1000 images per sub set. In the *test-sequential* data set the first 500 images come from the *OKI IRISPASS-h* and the latter 500 from the *Irisguard H100 IRT* sensor, as shown in Figure 6.3.

OKI	H100

Figure 6.3: Structure of the *test-sequential* data set: the first 500 images are from the *OKI IRISPASS-h* (OKI), while the latter 500 are from the *Irisguard H100 IRT* (H100) sensor.

Figure 6.4 shows the *test-mixed* data set containing alternating images from the *Irisguard H100 IRT* and the *OKI IRISPASS-h* sensor in blocks of 100 images.

H100	OKI	H100	OKI	H100	OKI	H100	OKI	H100	OKI

Figure 6.4: Structure of the *test-mixed* data set: 10 alternating blocks of 100 images from the *Irisguard H100 IRT* (H100) and *OKI IRISPASS-h* (OKI) sensors.

7 Experiments

The purpose of this work is to investigate the presence of images from multiple sensors inside each subset of the CASIA-Iris V4 Database, which has been described in Chapter 6. This is achieved by applying the forensic techniques from Chapter 5 on the subsets of the CASIA-Iris V4 Database. The following set-up describes mutual parameters for all the applied forensic techniques, which have been used for all the following experiments.

All the forensic investigation techniques used in this work are based on the sensors PRNU. The extraction of the PRNU and the PRNU fingerprint generation have been performed according to Section 2.2 and 2.3. Hence, for an estimation of each sensors PRNU fingerprint, the algorithm described by Fridrich [32] was used to extract the PRNU using the wavelet-based denoising filter from Appendix A in [52], where color images have been converted into grayscale images before the extraction. For the proposed techniques used for the forensic investigation the PRNU noise residual of every image is extracted from 4 patches located in the corners with a size of 128×128 pixels each, resulting in a total noise residual size of 256×256 pixels. This is done because the image size is varying between the data sets and because the corners mostly are less textured and contain less correlated content than the center of the images. The PRNU noise residuals have been normalized in respect to the L_2-norm because its embedding strength is varying between different sensors [64].

Since the images in the CASIA Iris-V4 database have not been geometrically transformed, the normalized cross correlation (NCC) has been used for similarity measures of two PRNU noise residuals or PRNU fingerprints respectively according to Section 2.4.

In addition, two different PRNU enhancement approaches are applied in this work and their impact is evaluated. Both aim to filter out scene details

by the following idea: Scene details contribute to the very strong signal components in the wavelet domain, so the stronger a signal component in the wavelet domain, the more it should be attenuated. The enhancement of the PRNU is performed in the discrete wavelet transform (DWT) domain, where an enhancement function is applied to the coefficients. Two different enhancement functions are used to filter out the contaminations related to the image content: *EnhLi3* that corresponds to the Model 3 from [47] and *EnhCald* that is proposed in [6]. Since the parameters for α proposed in the respective papers were quite different and yielded to unsatisfactory results, the threshold has been set to $\alpha = 2$. This value has been determined by analyzing the distribution of intra- and inter-sensor NCC scores of the sensors in the *Test* data set in Section 6.2. After the application of the enhancement function, the resulting coefficients are transformed back into the spatial domain by performing an inverse DWT (IDWT).

The PRNU extraction algorithm of [32] has been reimplemented by Höller *et al.* in [64], whose code has been used as a basis for this work. The code has been supplemented with the possibility of using different techniques to enhance the PRNU and to be able to make use of multi-core machines, several parts have been refactored to be executed in parallel. The BFAIC technique by Bloy [21] and the other forensic techniques have been implemented by the author of this work. As additional post processing step a zero mean operation has been applied to each extracted PRNU noise residual.

In the following Section 7.1 the parameter set-up for the experiments conducted with the various forensic techniques are described in detail.

7.1 Set-up for Forensic Techniques

This section describes the set-up and parameters specific for each forensic technique proposed in Chapter 5, while the set-up common for all techniques is described in the previous section.

Because the scenario of Bloys [21] and this work have a strong similarity and the results have been very promising, the Blind Camera Fingerprinting and Image Clustering (BCFAIC) has been reimplemented for the investigation

of the *CASIA-Iris V4 Database*. Two different threshold functions t_{BLOY} and t_{STSM} have been used for the experiments, where t_{BLOY} corresponds to the threshold function proposed in [21]. The threshold function t_{STSM} has been calculated using images from the *OKI IRISPASS-h (1)* and *Irisguard H100 IRT* sensors by following the same methodology described by Bloy and by performing the PNRU extraction according to Chapter 2. The parameters for both threshold functions are given in the equations in 7.1.

$$\begin{aligned} t_{Bloy} &= -0.00002438n^2 + 0.0002889n + 0.009 \\ t_{STSM} &= 0.0000007826n^2 - 0.000006373n + 0.0210 \end{aligned} \tag{7.1}$$

Since the number of sensors is unknown for each *CASIA-Iris V4* sub set and the K-Means Clustering (KM) needs a parameter k representing the number of clusters, the experiments have been conducted for $k = 1 \ldots 5$ with the assumption that not more than 5 sensors have been used for a single sub set. This limitation is not mandatory and can be extended if necessary, but increases the computational effort. The normalized cross-correlation (NCC) was chosen as distance metric and the K-Means clustering was repeated five times for each k to avoid local minima.

As in the case of the K-Means Clustering (KM), the experiments for the PCA K-Means Clustering (PCAKM) have been conducted for $k = 1 \ldots 5$ and the first $n = 5$ principal components have been chosen to form the feature vector for each PRNU noise residual. Furthermore the Squared Euclidian distance has been chosen as the distance metric. In contrast to all other forensic techniques, the PRNU for this technique has been extracted from a 256×256 patch in the center of each image under investigation, because it delivered slightly better results than the extraction used for the other techniques during testing.

The Sliding Window Fingerprinting (SWFP) has been performed with a window size of $n = 50$, after testing several window sizes for $n \in \{5, 10, 20, 50, 100, 150\}$. This is a good trade-off between the quality of the PRNU fingerprint and the ability to detect shorter sequences of images, since using too few images to generate the PRNU fingerprint lead to a weak estimate according to Fridich [19]. If one is able to generate a high

quality PRNU fingerprint using fewer images, for example by more sophisticated post processing or denoising approaches, a window size n as small as possible would be preferable.

For the Device Identification on Dataset Partitions (DIODP) all experiments have been conducted with the partition sizes $n \in \{10, 25, 50, 100, 200, 500\}$, where the results for the first two partition sizes 10 and 25 have to be treated with care because of the few amount of images for each partition. The same constraints regarding a weak PRNU estimate hold as previously explained for the SWFP. The interval of confidence for the EER scores has been estimated, as described in section 2.4, at $\alpha = 95\%$, hence the 2.5% lowest and 2.5% highest NCC scores have been excluded.

8 Results

This chapter containing the results of the forensic investigation experiments is divided into two sections, covering the results obtained by applying the forensic techniques to the *Test* data set described in Section 8.1 and the *CASIA-Iris V4* Database described in Section 8.2.

For each data set, the techniques have been applied using no PRNU enhancement (*NoEnh*) and the two PRNU enhancement techniques described in Chapter 7 (*EnhLi3* and *EnhCald*). The forensic techniques have been applied in the following order:

- Blind Camera Fingerprinting and Image Clustering (BCFAIC)
- K-Means Clustering (KM)
- PCA K-Means Clustering (PCAKM)
- Sliding Window Fingerprinting (SWFP)
- Device Identification On Dataset Partitions (DIODP)

The outcome of the experiments, where each listed forensic technique has been applied to the specific data sets, is presented in the following sections and subsections.

8.1 Test Set

This section covers the results for the *Test* data set, with its subsets *test-sequential* (TS) and *test-mixed* (TM). Additionally, the results without the use of PRNU enhancement techniques (*NoEnh*) are compared to the results with the *EnhLi3* and *EnhCald* enhancements applied. These experiments have also been taken as an evaluation basis on how well each forensic technique is able to detect images from multiple sensors in the data sets and to check if

the implemented algorithms are working as intended. The expected results should indicate that the data sets, *test-sequential* (TS) and *test-mixed*, both contain images from two respectively multiple sensors.

8.1.1 BCFAIC

	NoEnh		*EnhLi3*		*EnhCald*	
BCFAIC t_{Bloy}	TS	TM	TS	TM	TS	TM
images	1000	1000	1000	1000	1000	1000
total partitions	2	2	2	3	2	3
partitions > 100	2	2	2	2	2	2
partitions < 10	0	0	0	1	0	1
unassociated images	0	0	0	0	0	0
mismatched images	22	19	26	30	2	5

	NoEnh		*EnhLi3*		*EnhCald*	
BCFAIC t_{STSM}	TS	TM	TS	TM	TS	TM
images	1000	1000	1000	1000	1000	1000
total partitions	3	4	5	5	14	12
partitions > 100	2	2	2	2	2	2
partitions < 10	1	2	3	3	11	9
unassociated images	0	0	0	0	0	0
mismatched images	0	0	0	0	0	0

Table 8.1: Clustering Results of the BCFAIC technique applied on the *test-sequential* (TS) and *test-mixed* (TM) data sets using the threshold functions t_{Bloy} (top) and t_{STSM} (bottom). In addition, the PRNU enhancement techniques *EnhLi3* and *EnhCald* have been applied beside the case without PRNU enhancement applied (*NoEnh*).

First of all the BCFAIC technique known from literature has been applied to evaluate its performance in this scenario with a priori knowledge of correspondence of images and sensors. Looking at the results in Table 8.1 using the t_{Bloy} threshold function on top and t_{STSM} at the bottom, we realize that the algorithm produces generally good with some undesirable side effects.

The results of both threshold functions do not contain unassociated images, which is a positive aspect. Having a closer look at the results using the t_{Bloy}

threshold function, 2 partitions are generated in almost all cases, except for the TM data set when *EnhLi3* and *EnhCald* have been applied 3 partitions are generated by the clustering. However in these cases only 1 additional partition is generated, which contains only a few images from the same sensor. The main issue, which emerges from the usage of the t_{Bloy} threshold function, is that all combinations of PRNU enhancement and data set yield mismatched images. The number of mismatched images accounts for approximately 0.2% for the case where no PRNU enhancement is applied (*NoEnh*), 0.26 - 0.30% for *EnhLi3* and 0.02 - 0.05% for *EnhCald* in respect to the total amount of images, hence *EnhCald* was able to significantly diminish the value of mismatched images without compromising the clustering results.

On the other hand, the t_{STSM} threshold function, shows a slightly higher amount of total partitions (between 3 and 5) for *NoEnh* and *EnhLi3* when compared to t_{Bloy} and a considerably higher amount (10 to 12) for the *EnhCald* enhancement. Even though the number of total clusters increased, 2 large clusters have been generated throughout all cases. Hence mostly small clusterst with less than 10 images are responsible for the increase, especially for *EnhCald*. The main advantage over the t_{Bloy} threshold function is that there are no mismatched images, which was bought at the price of a higher amount of clusters, but since the additional clusters are very small in size, they could be categorized as some kind of "noise".

8.1.2 KM

Table 8.2 shows the results of the KM technique for the number of clusters $k = 1 \ldots 5$. As it can be clearly seen the technique performs as expected and shows the highest MSV for the correct number of clusters in all cases. Furthermore the MSV for the correct number of sensors, which is 2, is also significantly higher than for the incorrect numbers of clusters $1, 3, 4$ and 5. The largest MSV difference between the correct k and the others is obtained without applying a PRNU enhancement (*NoEnh*). Applying a PRNU enhancement, either *EnhLi3* or *EnhCald*, results in a lower MSV difference between correct and incorrect values for k although the value for the correct k remains always significantly higher than for the incorrect ones.

KM	*NoEnh*		*EnhLi3*		*EnhCald*	
k	TS	TM	TS	TM	TS	TM
1	0.0230	0.0235	0.0214	0.0219	0.0139	0.0143
2	**0.0384**	**0.0388**	**0.0351**	**0.0353**	**0.0234**	**0.0237**
3	0.0161	0.0257	0.0152	0.0230	0.0129	0.0126
4	0.0030	0.0037	0.0042	0.0041	0.0127	0.0022
5	0.0043	0.0035	0.0049	0.0040	0.0051	0.0028

Table 8.2: Mean silhouette value (MSV) for K-Means clustering (KM) performed on the *test-sequential* (TS) and *test-mixed* (TM) data sets using different PRNU enhancement techniques (*NoEnh*, *EnhLi3* and *EnhCald*). The highest MSV for each data set and PRNU enhancement is highlighted in bold.

8.1.3 PCAKM

PCAKM	*NoEnh*		*EnhLi3*		*EnhCald*	
k	TS	TM	TS	TM	TS	TM
1	0.0112	0.0111	0.0221	0.0257	0.0173	0.0176
2	**0.0381**	**0.0370**	**0.0662**	**0.0656**	0.0392	0.0416
3	0.0271	0.0269	0.0344	0.0312	**0.0416**	**0.0417**
4	0.0150	0.0245	0.0224	0.0191	0.0172	0.0168
5	0.0140	0.0134	0.0195	0.0175	0.0171	0.0163

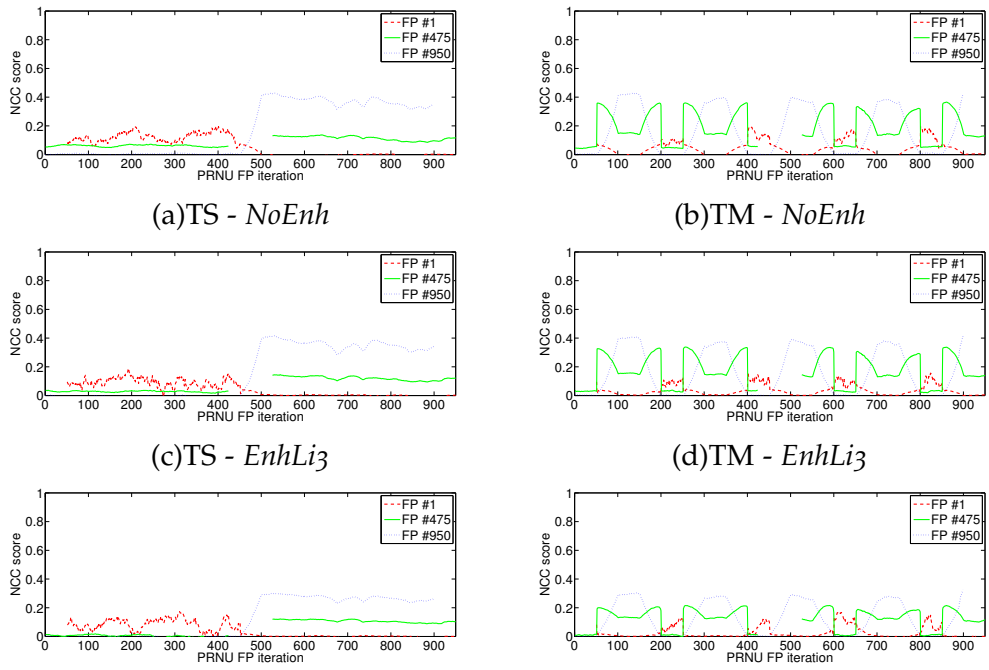
Table 8.3: Mean silhouette value (MSV) for PCA K-Means clustering (PCAKM) performed on the *test-sequential* (TS) and *test-mixed* (TM) data sets using different PRNU enhancement techniques (*NoEnh*, *EnhLi3* and *EnhCald*). The highest MSV for each data set and PRNU enhancement is highlighted in bold.

The results of the PCAKM technique, shown in Table 8.3, behave similar to the KM results for *NoEnh* and *EnhLi3* where $k = 2$ yields the highest MSV value as expected. However the MSV values for incorrect k values tend to be higher in comparison with the KM results and *EnhLi3* also shows much higher MSV values for the correct k. The *EnhCald* enhancement induces the highest MSV value for $k = 3$, which is incorrect for the two data sets TS and TM, but the MSVs for $k = 2$ are very close to the highest scores. Therefore no reliable assumption on the exact number of sensors can be established in this case for *EnhCald* and it can only be observed that multiple sensors have

been used, but not how many.

8.1.4 SWFP

Figure 8.1 shows the results of the SWFP technique with a window size of 50. We can see three different curves that are representing the NCC correlation scores of three different fingerprint iterations with all other fingerprints, which have been generated according to Section 5.3. The dashed (red) line shows the NCC scores for the fingerprint with iteration 1, the solid (green) line for iteration 475 and the dotted (blue) line for iteration 950.



(a)TS - *NoEnh* (b)TM - *NoEnh*

(c)TS - *EnhLi3* (d)TM - *EnhLi3*

(e)TS - *EnhCald* (f)TM - *EnhCald*

Figure 8.1: Results of the SWFP experiment on the *test-sequential* (TS) and *test-mixed* (TM) data sets using different PRNU enhancement techniques (*NoEnh*, *EnhLi3* and *EnhCald*). The different curves represent the NCC scores of a specific PRNU fingerprint iteration with all other iterations. The transitions indicated a change of the image source sensor in the TS (a, c, e) and the TM data set (b, d, f).

The NCC scores for all three emphasized iterations (1, 475 and 950) show a collective transition at the fingerprint iterations 451 until 500 in the *test-sequential* (TS) data set. For this data set, according to the data set specification in Section 6.2, these are the iterations where the alteration of the image source sensors happens. Until iteration 451 all fingerprints contain solely images from the first sensor (images 1 to 500). From iteration 452 until 500, the fingerprints are composed by images from both sensors (images 452 to 500 from the first sensor, images 501 to 549 from the second one) and from iteration 501 on the fingerprints are composed of solely images from the second sensor (images 501 to 1000). Hence the transitions in all three emphasized iterations can be observed like expected from the specification of the data set.

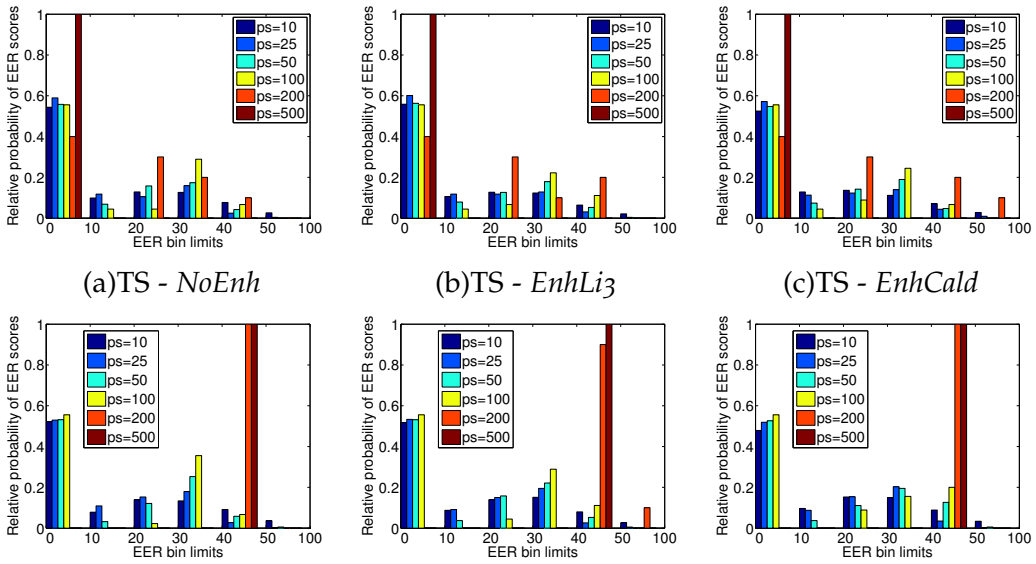
Applying the SWFP technique to the *test-mixed* (TM) data set, one would expect from the data set specification in Section 6.2 that the first transition in the data should occur at iteration 51 and that the transitions should repeat every 100 iterations, since the images in the data set are always alternating in blocks of 100 images from each sensor. This can exactly be observed in the subfigures (a), (c) and (e) of Figure 8.1, where the dashed (red) line shows the NCC scores for the fingerprint with iteration 1, the solid (green) line for iteration 475 and the dotted (blue) line for iteration 950. The previously described transitions can be observed for all three emphasized iterations (1, 475 and 950).

For both data sets, TS and TM, the two PRNU enhancement techniques *EnhLi3* and *EnhCald* lead to lower NCC scores in general. Both techniques tend to show transitions in the data where no transitions would be expected, e.g. around iteration 200 and 375 for *EnhCald* in Subfigure (e) of Figure 8.1, or diminish the effect of the transitions compared to *NoEnh*, as it can be seen in Subfigure (f) of Figure 8.1. These effects are visible for the *EnhLi3* technique and even more pronounced for the *EnhCald* technique.

8.1.5 DIODP

The last technique applied to the *Test* data set is the Device Identification on Dataset Partitions (DIODP), which has been performed using different partition sizes $n \in \{10, 25, 50, 100, 200, 500\}$. As described in Section 5.4 for

this technique the distribution of EER scores is analyzed to determine the presence of images from various sensors in the data set under investigation. Figure 8.2 shows the EER scores distribution, both for the *test-sequential* TS and the *test-mixed* TM data sets with (*EnhLi3*, *EnhCald*) and without (*NoEnh*) PRNU enhancement techniques applied.



(a)TS - *NoEnh* (b)TS - *EnhLi3* (c)TS - *EnhCald*

(d)TM - *NoEnh* (e)TM - *EnhLi3* (f)TM - *EnhCald*

Figure 8.2: Results of the DIODP experiment with different partition sizes (ps) on the *test-sequential* (TS) and *test-mixed* (TM) data sets using different PRNU enhancement techniques (*NoEnh*, *EnhLi3* and *EnhCald*) showing the distribution of EER scores. A significant amount of very low EER scores can be observed for both data sets, which indicates the presence of images from multiple sensors.

All cases show a high amount of very low EER scores. This is the expected result for this technique, since both data sets, TS and TM, contain images from multiple sensors. Having a look at the different partition sizes, it can be seen that the larger partition sizes lead to unreliable results when images from more than one sensor are put together into a single partition. This becomes clearly visible for the TM data set, where images from both sensors are alternated in blocks of 100 images. Therefore the results from all partition sizes need to be taken into consideration. The PRNU enhancing

techniques have only had a very small impact on the scores distributions, as it can be seen in Figure 8.2.

8.2 CASIA Iris-V4 Database

In the following section the proposed forensic techniques are applied to the *CASIA-Iris V4* data set in order to detect if multiple sensors have been used to acquire the data sets under investigation, which are:

- CASIA-Iris-Interval: *intv*
- CASIA-Iris-Lamp: *lamp*
- CASIA-Iris-Twins: *twin*
- CASIA-Iris-Distance: *dist*
- CASIA-Iris-Thousand: *thou*

All the experiments have been conducted both with the use of PRNU enhancement techniques *EnhLi3* and *EnhCald* and without (*NoEnh*). The results should clarify the open question whether each data set has been acquired using a single or multiple sensors for each data set.

8.2.1 BCFAIC

The Blind Camera Fingerprinting and Image Clustering (BCFAIC) has been the first technique applied to the CASIA-Iris V4 database to see how it performs in a real world scenario. Mismatched images could not be identified in this scenario, because no documented ground truth on the source sensor of every image exists. This technique creates clusters of associated images (images with a high NCC score) and partitions the data set. The resulting partitions should reflect the number of distinct sensors used in the data set. Unassociated images have a very low NCC score among each other, so that they are classified as being all from different sensors because they could not be clustered properly.

The results in Table 8.4 shows a generally high number of clusters for the *intv* and *lamp* data sets and a relatively low number for the *twin*, *dist*

BCFAIC *NoEnh*			t_{Bloy}					t_{STSM}		
	intv	*lamp*	*twin*	*dist*	*thou*	*intv*	*lamp*	*twin*	*dist*	*thou*
# IMG	1307	6855	1095	1566	2000	1307	6855	1095	1566	2000
# P	10	16	3	1	2	143	216	20	1	6
P $>$ 100	4	5	1	1	1	2	11	1	1	1
P $<$ 10	3	4	1	0	1	18	159	18	0	4
UI	0	0	0	0	0	0	0	0	0	0

BCFAIC *EnhLi3*			t_{Bloy}					t_{STSM}		
	intv	*lamp*	*twin*	*dist*	*thou*	*intv*	*lamp*	*twin*	*dist*	*thou*
# IMG	1307	6855	1095	1566	2000	1307	6855	1095	1566	2000
# P	11	15	3	1	2	186	266	24	1	14
P $>$ 100	4	5	1	1	1	1	12	1	1	2
P $<$ 10	3	5	1	0	0	168	129	19	0	12
UI	0	0	0	0	0	0	0	0	0	0

BCFAIC *EnhCald*			t_{Bloy}					t_{STSM}		
	intv	*lamp*	*twin*	*dist*	*thou*	*intv*	*lamp*	*twin*	*dist*	*thou*
# IMG	1307	6855	1095	1566	2000	1307	6855	1095	1566	2000
# P	17	20	6	1	4	6	2867	307	1	193
P $>$ 100	4	7	1	1	2	1	0	3	1	3
P $<$ 10	5	4	3	0	1	4	260	254	0	188
UI	0	0	0	0	0	928	0	0	0	0

Table 8.4: Clustering Results of the BCFAIC technique applied on the *CASIA-Iris V4* data sets using the threshold functions t_{Bloy} (left) and t_{STSM} (right) and the *EnhLi3* (middle) and *EnhCald* (bottom) enhancements applied beside the case without PRNU enhancement *NoEnh* (top). The tables show the number of images (# IMG), total number of partitions (# P), large partitions (P $>$ 100), small partitions (P $<$ 10) and images that could not be associated to any partition (UI).

and *thou* data sets. The *EnhCald* produces the highest amounts of clusters among all enhancement techniques, especially when combined with the t_{STSM} threshold function, which even leads to 928 out of 1307 unassociated images for the *intv* data set. Since these results do not look realistic at all, the results of *EnhCald* are discarded for the further analysis of the BCFAIC results.

The t_{STSM} function produces a much higher number of clusters for all data sets and all PRNU enhancement techniques compared to the t_{Bloy} function,

except for the *dist* data set, which results in 1 cluster for all cases. Therefore it is safe to say that *dist* has been acquired with a single sensor according to the BCFAIC technique.

NoEnh			KM		
k	*intv*	*lamp*	*twin*	*dist*	*thou*
1	0.0024	**0.0391**	**0.0398**	**0.3821**	**0.0121**
2	0.0034	0.0100	0.0052	0.0632	0.0074
3	0.0031	0.0004	0.0063	−0.0076	0.0064
4	0.0032	0.0003	0.0057	−0.0069	0.0045
5	0.0030	0.0001	0.0053	−0.0071	0.0042
EnhLi3			KM		
k	*intv*	*lamp*	*twin*	*dist*	*thou*
1	0.0024	**0.0391**	**0.0398**	**0.3821**	**0.0121**
2	0.0034	0.0100	0.0052	0.0632	0.0074
3	0.0031	0.0004	0.0063	−0.0076	0.0064
4	0.0032	0.0003	0.0057	−0.0069	0.0045
5	0.0030	0.0001	0.0053	−0.0071	0.0042
EnhCald			KM		
k	*intv*	*lamp*	*twin*	*dist*	*thou*
1	0.0025	**0.0167**	**0.0167**	**0.1619**	**0.0051**
2	0.0045	0.0009	0.0014	0.0264	0.0021
3	0.0051	0.0005	0.0015	−0.0049	0.0018
4	0.0054	0.0001	0.0013	−0.0049	0.0014
5	0.0053	0.0001	0.0011	−0.0061	0.0014

Table 8.5: Clustering Results of the KM technique applied on the *CASIA-Iris V4* data sets using the *EnhLi3* (middle) and *EnhCald* (bottom) enhancements beside the case without PRNU enhancement *NoEnh* (top). The highest MSV value is highlighted in bold for each data set if it is much higher than the other values.

As mentioned before, the *twin* and *thou* data sets result in a relatively low number of clusters, with mostly small clusters contributing to the amount of clusters and yielding only 1 large cluster in almost all cases. The *intv* and *lamp* data sets show mixed results, with a high number of clusters for t_{Bloy} and a very high number for t_{STSM}. Having a closer look again at the large clusters and small clusters for those two data sets, the number of large clusters is clearly higher than 1 and there are also a high amount of clusters

that contain between 10 and 100 images, which leads to the assumption that multiple sensors might have been used for those two data sets. In conclusion the results of the BCFAIC are hard to interpret in general, except for the *dist* data set.

8.2.2 KM

The results of the K-Means (KM) experiments are presented in Table 8.5. The KM clustering generates groups images from the same source and the results show that one sensor was used to acquire all subsets except the *intv* data set, for which no clear result can be observed. This holds for all enhancement techniques, whereat the *EnhCald* yields the largest differences between the highest MSV and the second highest. The *EnhLi3* seems to have no impact when applied for the KM technique because the resulting MSVs do not differ from the *NoEnh* results.

8.2.3 PCAKM

The PCA K-Means results, shown in Table 8.6, actually do not permit any statements to be made on the number of sensors because of the mostly very close MSVs results for different numbers of clusters k. Though there is an observable trend in the data, where the MSVs for 2 clusters have the largest value and the 1 cluster results have the lowest values among all values for k. This indicates that multiple sensors have been used to acquire all the data sets, which is different from the results of the other forensic techniques applied until this point.

8.2.4 SWFP

The Sliding Window Fingerprinting (SWFP) moves a window with a defined size over the data image after image and a PRNU fingerprint from the data within this window is calculated in each step. The presence of images from multiple sensors in the data set should express in a sudden increase or decrease of the correlation score. If only images from one sensor are present

NoEnh			PCAKM		
k	*intv*	*lamp*	*twin*	*dist*	*thou*
1	0.0003	0.0010	0.0024	0.0019	0.0006
2	0.0098	**0.0185**	0.0197	**0.0234**	**0.0474**
3	0.0090	0.0123	0.0141	0.0186	0.0427
4	0.0088	0.0097	0.0119	0.0180	0.0391
5	0.0089	0.0086	0.0105	0.0177	0.0405

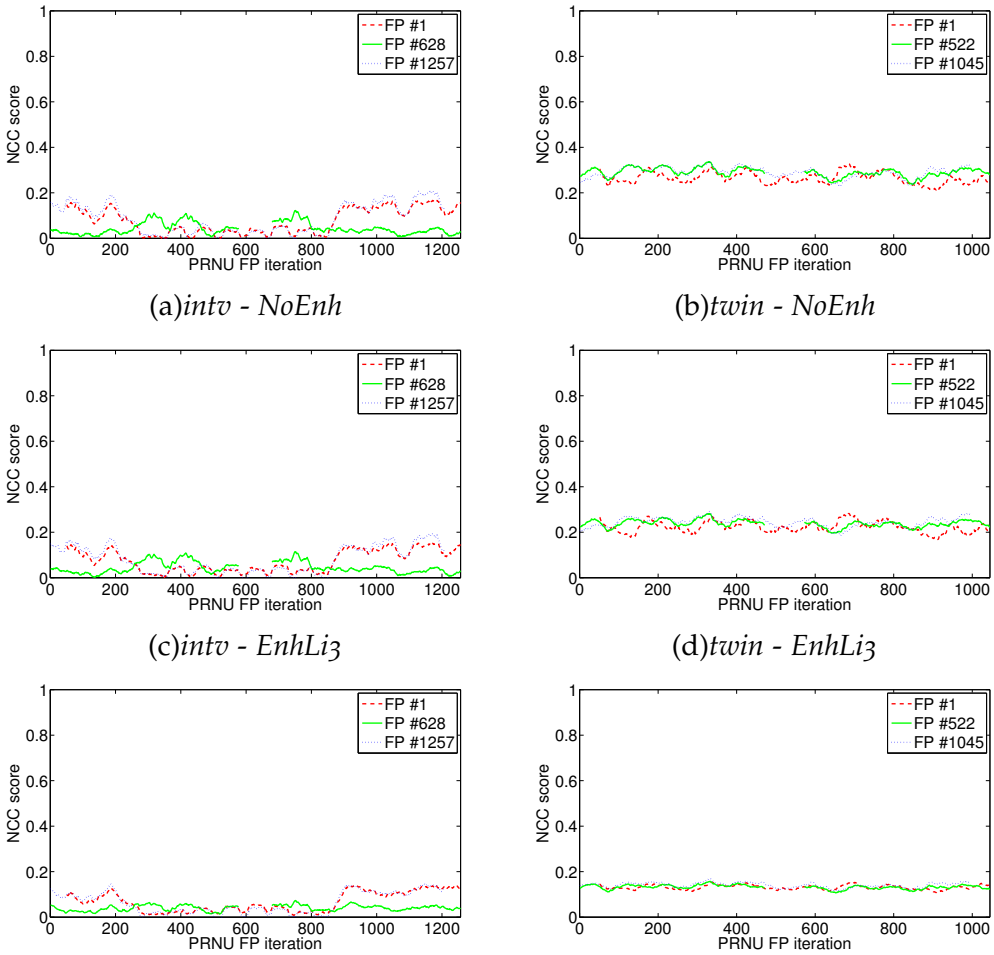
EnhLi3			PCAKM		
k	*intv*	*lamp*	*twin*	*dist*	*thou*
1	0.0004	0.0006	0.0016	0.0013	0.0005
2	0.0093	**0.0139**	**0.0194**	0.0134	0.0184
3	0.0083	0.0104	0.0148	0.0118	0.0164
4	0.0082	0.0080	0.0115	0.0111	0.0152
5	0.0084	0.0074	0.0104	0.0110	0.0145

EnhCald			PCAKM		
k	*intv*	*lamp*	*twin*	*dist*	*thou*
1	0.0063	0.0001	0.0003	0.0014	0.0002
2	0.0144	0.0066	0.0088	0.0124	0.0089
3	0.0193	0.0057	0.0080	0.0123	0.0082
4	0.0239	0.0055	0.0079	0.0124	0.0077
5	0.0275	0.0053	0.0081	0.0140	0.0078

Table 8.6: Clustering Results of the PCAKM technique applied on the *CASIA-Iris V4* data sets using the *EnhLi3* (middle) and *EnhCald* (bottom) enhancements beside the case without PRNU enhancement *NoEnh* (top). The tables MSV results for the given values of *k* which represent the number of clusters (sensors). The highest MSV value is highlighted in bold for each data set if it is much higher than the other values.

in the data set, the correlation scores among all images should be quite stable around a certain level.

The results for the *intv* and *twin* data set in Figure 8.3 show two transitions with a large offset at iteration 200 - 275 and the other one at around 850 - 900 for the *intv* data set. Additionally the results also show multiple small transitions between the iterations 300 and 850. These observations are consistent throughout all PRNU enhancements. The outcome of the *twin* data set does not show any noticeable transitions at all and all scores are



(a)*intv - NoEnh* (b)*twin - NoEnh*

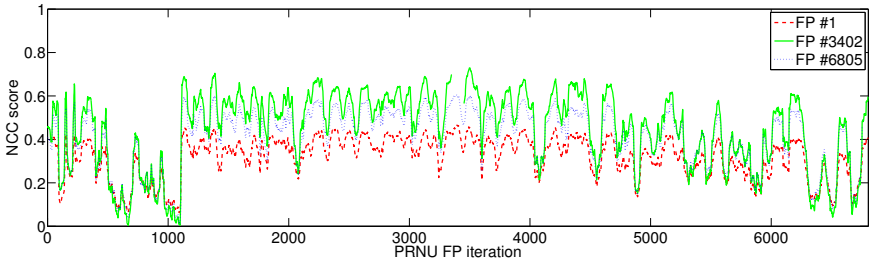
(c)*intv - EnhLi3* (d)*twin - EnhLi3*

(e)*intv - EnhCald* (f)*twin - EnhCald*

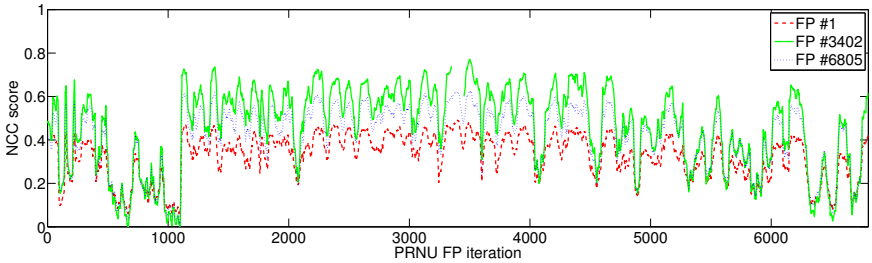
Figure 8.3: Results of the SWFP experiment on the *intv* (a, c, e) and *twin* (b, d, f) data sets using different PRNU enhancement techniques (*NoEnh*, *EnhLi3* and *EnhCald*). The different curves represent the NCC scores of a specific PRNU fingerprint iteration with all other iterations.

quite stable. Regarding the PRNU enhancements, the *EnhLi3* enhancement produces comparable results as if no enhancement is applied both data sets. Only a small offset is noticeable in the correlation scores between the enhanced experiments *EnhLi3* and *EnhCald* and the un-enhanced experiment

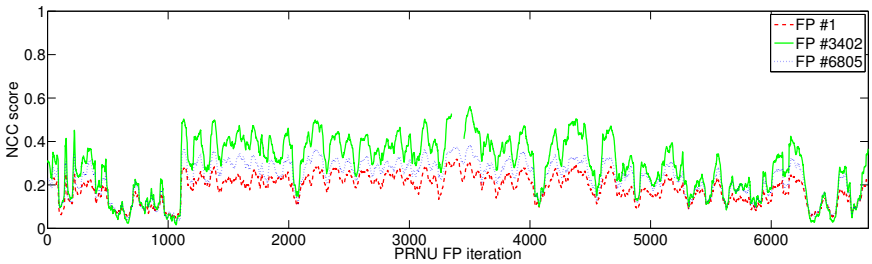
NoEnh, where the *EnhLi3* are almost equal as the *NoEnh* and the *EnhCald* scores are slightly lower, but the transitions are similar for all experiments.



(a)*lamp - NoEnh*



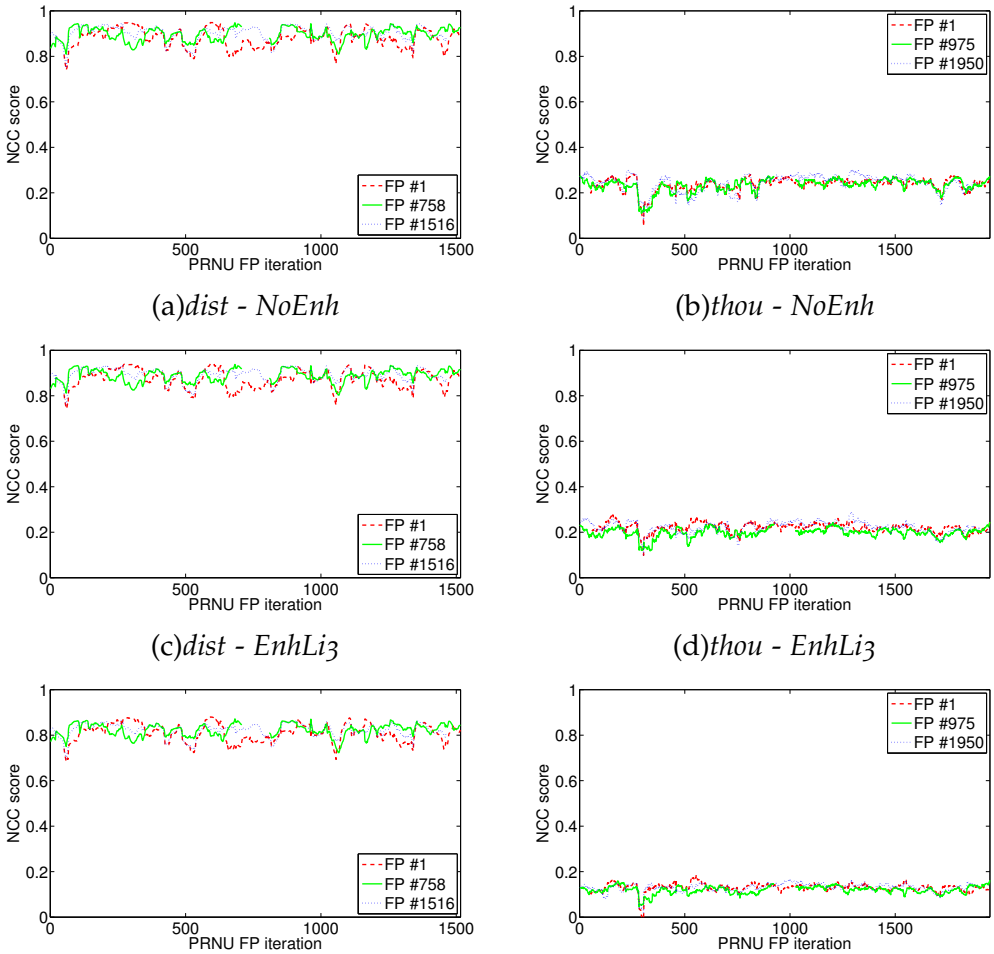
(b)*lamp - EnhLi3*



(c)*lamp - EnhCald*

Figure 8.4: Results of the SWFP experiment on the *lamp* data set using different PRNU enhancement techniques (*NoEnh*, *EnhLi3* and *EnhCald*). The different curves represent the NCC scores of a specific PRNU fingerprint iteration with all other iterations.

Figure 8.4 shows the outcome of experiments on the *lamp* data set. The first thing to be noticed is that the NCC score curve is has a high variance, hence in this case only large *jumps* are accounted as a transition. Such a



(a)*dist - NoEnh* (b)*thou - NoEnh*

(c)*dist - EnhLi3* (d)*thou - EnhLi3*

(e)*dist - EnhCald* (f)*thou - EnhCald*

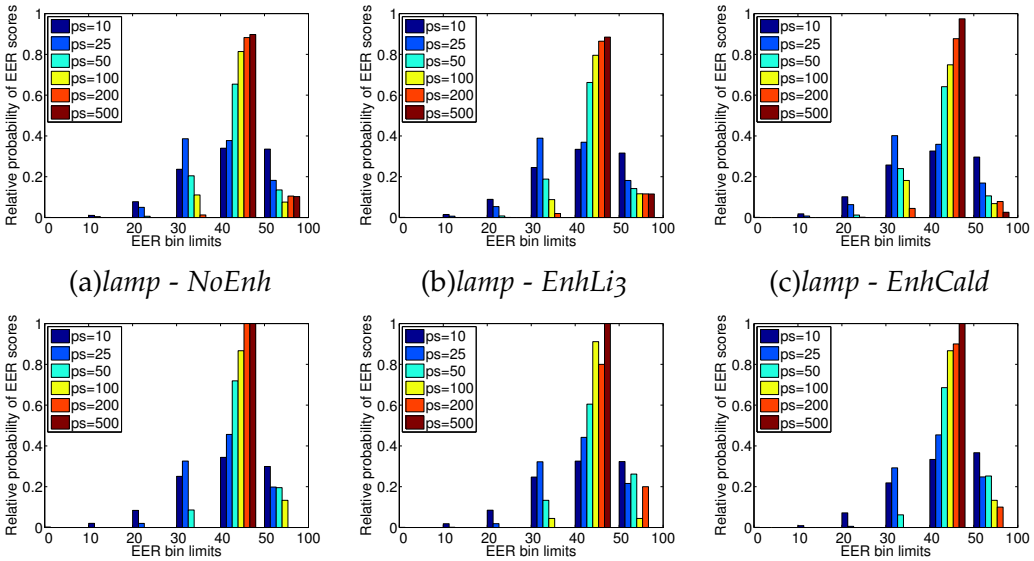
Figure 8.5: Results of the SWFP experiment on the *dist* (a, c, e) and *thou* (b, d, f) data sets using different PRNU enhancement techniques (*NoEnh*, *EnhLi3* and *EnhCald*). The different curves represent the NCC scores of a specific PRNU fingerprint iteration with all other iterations.

transition can be observed at approximately iteration 700 and 1050. Again, the different PRNU enhancement techniques show similar results.

For the remaining data sets, *dist* and *thou*, no transitions in the correla-

tion scores can be identified. They are comparable for *EnhLi3*, *EnhLi3* and *EnhCald*, therefore these data sets have probably been acquired with a single sensor according to this experiment. The only difference is an offset in the correlation scores for the individual enhancement configurations.

Summing up, this technique suggests that all data sets, with the exception of *lamp* and *intv*, have been acquired with a single sensor. Regarding the PRNU enhancements it can be observed that the two PRNU enhancements *EnhLi3* and *EnhCald* exhibit decreased mean correlation scores.



(a)*lamp - NoEnh* (b)*lamp - EnhLi3* (c)*lamp - EnhCald*

(d)*twin - NoEnh* (e)*twin - EnhLi3* (f)*twin - EnhCald*

Figure 8.6: Results of the DIODP experiment with different partition sizes (ps) on the *lamp* (a, b, c) and *twin* (d, e, f) data sets using different PRNU enhancement techniques (*NoEnh*, *EnhLi3* and *EnhCald*) showing the distribution of EER scores.

8.2.5 DIODP

The Device Identification on Dataset Partitions (DIODP) experiment divides the data sets into n partitions with the same size and treat the disjoint partitions as n different sensors. If the resulting EER score is low (e.g. 0%),

the extracted PRNU and respectively the PRNU fingerprint is different for both partitions, hence they must have been acquired with different sensors.

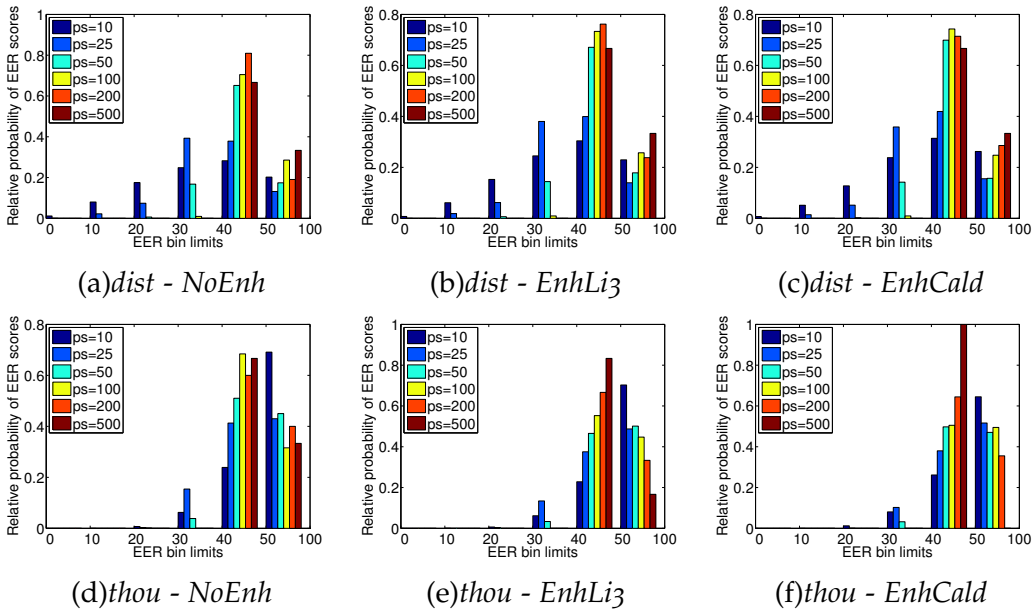


Figure 8.7: Results of the DIODP experiment with different partition sizes (ps) on the *dist* (a, b, c) and *thou* (d, e, f) data sets using different PRNU enhancement techniques (*NoEnh*, *EnhLi3* and *EnhCald*) showing the distribution of EER scores. No significant amount of very low EER scores can be observed for both data sets, which indicates that the data sets have been acquired each with a single sensor.

On the other hand, if the resulting EER score is high (e.g. 50%), the extracted PRNU very similar for both partitions and their images have all likely been acquired with the same sensor. After calculating the pairwise EER scores for all partition combinations P_i and Pj, where $i \neq j$, the EER score distribution is evaluated. If the distribution contains mostly high EER scores, the data set probably contains images from a single sensor. On the other hand, if the distribution contains very low EER scores, the data set is suspicious of containing images from multiple sensors.

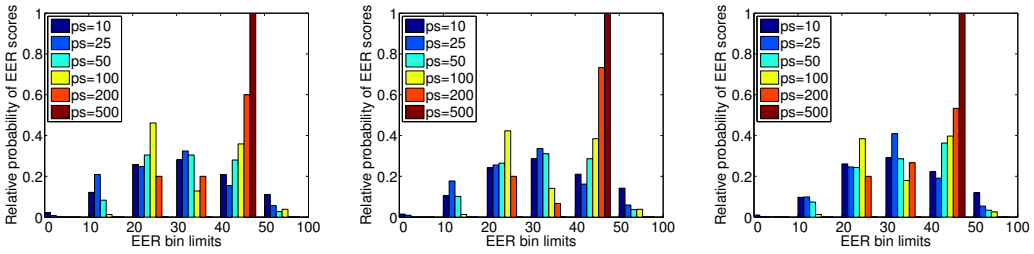
To be able to clearly represent the resulting EER scores we performed a

binning of the scores into six bins with the following limits: scores below 10%, between 10% and 20%, between 20% and 30%, between 30% and 40%, between 40% and 50%, and scores above 50%, where the lower bounds are inclusive and the upper bounds are exclusive.

From the Device Identification on Dataset Partitions (DIODP) experiments in the figures 8.6, 8.7 and 8.8, it can be observed that for almost all data sets other than *intv* the EER scores are higher than 30% (figures 8.6 and 8.7), which indicates that the *lamp*, *twin*, *dist* and *thou* data sets might have been acquired with one sensor.

Having a closer look at the *intv* data set in Figure 8.8 with different partition sizes, the results indicate that this data set might have been acquired with more than one sensor, because the distribution of the EER scores is similar to the one from the two sensors in the *Test* data sets, only with higher EER scores.

Similar to the previous forensic techniques, the results of the two PRNU enhancement approaches are quite similar to the un-enhanced ones for all data sets under investigation.



(a)*intv - NoEnh* (b)*intv - EnhLi3* (c)*intv - EnhCald*

Figure 8.8: Results of the DIODP experiment with different partition sizes (ps) on the *intv* data sets using different PRNU enhancement techniques (*NoEnh*, *EnhLi3* and *EnhCald*). A significant amount of low scores can be observed in the EER score distribution for this data set, which indicates the presence of images from multiple sensors.

9 Conclusion and Future Work

The goal of this thesis was to determine whether the various *CASIA-Iris V4* data sets have been acquired using multiple sensors (of the same model) or if a single sensor has been used, since no documentation on this exists and the images do not contain any kind of metadata information. In addition, in the case that images from multiple sensors are detected, establishing a ground truth for the *CASIA-Iris V4* database was intended. This investigation has been conducted by developing and applying different forensic methods to detect the number of sensors used to acquire the images within the analyzed data sets. To address the issue with the highly correlated image content in the data sets, the application of two PRNU enhancement techniques have been included in this work and the impact has been evaluated.

The novel forensic techniques proposed in this work detect the number of sensors in an image data set based on various aspects. The *K-Means Clustering* (KM) and *PCA-Kmeans Clustering* (PCAKM) techniques aim at grouping the images into different numbers of clusters representing the different sensors and then evaluate the optimal number of partitions based on intra- and inter-partition similarities. The *Sliding Window Fingerprinting* (SWFP) technique aims at detecting the presence of multiple sensors based on the similarity between calculated PRNU fingerprints and shows if a different sensor has been used during the acquisition of the images by exploiting the decrease of the similarity scores of consecutive PRNU fingerprints. The *Device Identification On Dataset Partitions* (DIODP technique divides the data set into equally sized partitions and performs a device identification with the assumption that each partition has been acquired with a distinct sensor. The distribution of the device identification scores is finally evaluated to determine if only a single sensor or if multiple sensors have been used to acquire the images.

All the proposed forensic techniques have consistently been able to detect the presence of multiple sensors in the investigated *Test* data sets, for which the number of sensor was known, and outperformed the reimplemented related technique (BFAIC), which had some issues with some images being associated with the wrong sensor. Some of the developed techniques were also able to exactly point out the number of different sensors (KM, PCAKM). The DIODP technique was able to show the presence of multiple sensors, while the SWFP was able to show the alteration of sequences of images from different sensors in the data set. The impact of the PRNU enhancement techniques on the results of the experiments on the *Test* data sets has been almost negligible, which confirms that the highly correlated content in the data sets has virtually no influence on the discriminative power of the sensors and corresponds to the observations in [39], where the use of uncorrelated data did not improve the discriminability.

The investigation results on the *CASIA-Iris V4* database are not as clear as the *Test* set results, hence only an assumption based on the trends in the results can be given because this is a completely blind investigation without any a priori knowledge about the sensors. The output of the BFAIC shows that the *dist* data set has been acquired using a single sensor, while for the other data sets the number of clusters always exceeds one. Having a closer look at the large and small clusters, containing more than 100 and less than 10 images respectively, the *twin* and *thou* data sets show a single large cluster and a few small ones, which could indicate that some images have simply been misclassified. On the other hand the results of the *intv* and *lamp* data sets show multiple large clusters, which could indicate the presence of images from multiple sensors. The KM technique shows the highest scores for the single cluster configuration for the *lamp*, *twin*, *dist* and *thou* data sets, indicating that a single sensor has been used for the image acquisition. The results for the *intv* data set are not clear, since the scores are similar for all the investigated numbers of clusters. The PCAKM technique shows the highest scores for 2 clusters across all data sets, but the highest scores are mostly very close to the scores obtained with other cluster configurations. Therefore no clear statement on the number of sensors can be made with the PCAKM technique. The SWFP technique does not show any transitions in the correlation scores for the *twin*, *dist* and *thou*, while it shows observable transitions for the *intv* and *lamp* data sets. This suggests

that the latter two have been acquired using multiple sensors. Finally, the DIODP technique shows a suspicious EER score distribution for the *intv* data set indicating the presence of images from multiple sensors, while all other data sets (*lamp*, *twin*, *dist* and *thou*) do not show such distributions.

Summarizing the outcome of all the forensic techniques applied to the *CASIA-Iris V4* database, the results indicate that the *intv* data set might have been acquired with more than one sensor, while all other *CASIA-Iris V4* sub sets have been acquired with one sensor. The *lamp* and *twin* data sets, suspicious of containing images from multiple sensors because of the same model denoted in the specification, seem both to be acquired with just a single distinct sensor. These data sets were also responsible for the high EER scores in the experiment of Höller *et al.* [64].

The detection of multiple sensors in the *CASIA-Iris V4* data sets remains a challenging task, since it is a completely blind approach without any a priori knowledge of the sensors. Unknown factors could have had an impact on the quality of the PRNU noise residuals and hence tampered the results, therefore further studies have to be conducted to be able to use sensor fingerprints as an authentication measure for biometric systems. These factors could include some internal processing of the sensors or an unknown image processing which affects the PRNU in all images.

The performance of the proposed forensic techniques has been evaluated on two data sets containing images from two different sensors in fixed sequences, where the results have been quite satisfactory. However this data sets are not very representative for a real world scenario and more sophisticated data sets with random occurrences of images from more than two sensors should be used as benchmark. For this purpose the *Dresden Image Database* [23] would be an appropriate candidate, since it incorporates several sensors from different models as well as different manufacturers and it is very well documented.

Artifacts degrading the quality of the PRNU, such as the NUAs observed for the sensors in the *Dresden Image Database* [23], have to be investigated for the sensors of the *CASIA-Iris V4* database and also other similarity measures in the literature show improved results, such as the peak correlation energy (PCE) [19] or the circular correlation norm (CCN) [36], which are able to lower the false positive ratio. PCE and CCN have also the ability to align

images or crops from images, which was not needed in this scenario since all images have the full sensor resolution, but this might be interesting for further studies.

An alternative way to enhance the quality of the PRNU noise extraction is to apply various denoising and filtering techniques. First of all a Wiener filtering should be applied to deal with periodic artifacts, as described in literature, and additionally varying the denoising filter may improve the quality of the extracted PRNU, e.g. the filter proposed by Cooper [9] which besides is able to deal with JPEG compression artifacts that are peresent in the *CASIA-Iris V4* database. Another very promising denoising filter is the BM3D technique by Dabov *et al.* [12], which has been reported to improve the results in [10].

For this work the PRNU has been extracted from the images corners, which might also not be optimal because of the quality of the PRNU is related to the location, where the vignetting effect is said to have a negative impact, as determined by Li *et al.* [46] and it was proposed to extract the PRNU from the image center to obtain a high quality estimate of the PRNU.

If it is possible to enhance the quality of the PRNU by applying the mentioned enhancements and to obtain a more accurate estimate of the PRNU using less images, especially the proposed SWFP and DIODOP techniques would profit substantially from this because a lesser number of images implies a more accurate detection of images from multiple sensors. Obviously, also the other proposed forensic techniques would benefit from a more accurate and hence higher quality PRNU estimate.

Bibliography

[1] Farid Ahmed and Ira S. Moskowitz. "Composite Signature Based Watermarking for Fingerprint Authentication." In: *Proceedings of the 7th Workshop on Multimedia and Security*. MM&Sec '05. New York, NY, USA: ACM, 2005, pp. 137–142. ISBN: 1-59593-032-9. DOI: 10.1145/1073170.1073195. URL: http://doi.acm.org/10.1145/1073170.1073195 (cit. on p. 9).

[2] E.J. Alles, Z.J. Geradts, and C.J. Veenman. "Source Camera Identification for Heavily JPEG Compressed Low Resolution Still Images." In: *Journal of Forensic Sciences* 54.3 (2009), pp. 628–638 (cit. on p. 9).

[3] I. Amerini et al. "Blind image clustering based on the Normalized Cuts criterion for camera identification." In: *Signal Processing: Image Communication* 29 (2014), pp. 831–843 (cit. on pp. 39, 42).

[4] N. Bartlow et al. "Identifying sensors from fingerprint images." In: *Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops 2009. IEEE Computer Society Conference on*. June 2009, pp. 78–84. DOI: 10.1109/CVPRW.2009.5204312 (cit. on p. 2).

[5] Ed Bott. *Microsoft to add 'enterprise grade' biometric security to Windows 10*. 2015. URL: http://www.zdnet.com/article/microsoft-to-add-enterprise-grade-biometric-security-to-windows-10/ (visited on 05/07/2015) (cit. on p. 1).

[6] R. Caldelli et al. "Fast image clustering of unknown source images." In: *IEEE International Workshop on Information Forensics and Security (WIFS) 2010*. 2010, pp. 1–5. DOI: 10.1109/WIFS.2010.5711454 (cit. on pp. 20, 36, 42, 43, 62).

[7] Mo Chen et al. "Determining Image Origin and Integrity Using Sensor Noise." In: *IEEE Transactions on Information Security and Forensics* 3.1 (Mar. 2008), pp. 74–90 (cit. on pp. 9, 19).

[8] V. Christlein et al. "An Evaluation of Popular Copy-Move Forgery Detection Approaches." In: *Information Forensics and Security, IEEE Transactions on* 7.6 (Dec. 2012), pp. 1841–1854. ISSN: 1556-6013. DOI: 10.1109/TIFS.2012.2218597 (cit. on p. 4).

[9] Alan J. Cooper. "Improved photo response non-uniformity (PRNU) based source camera identification." In: *Forensic Science International* (2013) (cit. on pp. 21, 86).

[10] Andrea Cortiana et al. "Performance Comparison of Denoising Filters for Source Camera Identification." In: *SPIE Conference on Media Watermarking, Security, and Forensics*. San Francisco, CA, 2011 (cit. on p. 86).

[11] Luca Cuccovillo et al. "Blind Microphone Analysis and Stable Tone Phase Analysis for Audio Tampering Detection." In: *Proceedings of the 135th Audio Engineering Society (AES) Convention*. New York, USA, Oct. 2013 (cit. on p. 3).

[12] K. Dabov et al. "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering." In: *Image Processing, IEEE Transactions on* 16.8 (Aug. 2007), pp. 2080–2095. ISSN: 1057-7149. DOI: 10.1109/TIP.2007.901238 (cit. on p. 86).

[13] A. De Rosa et al. "Investigating multimedia contents." In: *Security Technology (ICCST), 2014 International Carnahan Conference on*. 2014, pp. 1–6 (cit. on pp. 3, 4).

[14] MohammadOmar Derawi, Bian Yang, and Christoph Busch. "Fingerprint Recognition with Embedded Cameras on Mobile Phones." English. In: *Security and Privacy in Mobile Information and Communication Systems*. Ed. by Ramjee Prasad et al. Vol. 94. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Springer Berlin Heidelberg, 2012, pp. 136–147. ISBN: 978-3-642-30243-5. DOI: 10.1007/978-3-642-30244-2_12. URL: http://dx.doi.org/10.1007/978-3-642-30244-2_12 (cit. on p. 1).

[15] Shumin Ding, Chunlei Li, and Zhoufeng Liu. "Protecting Hidden Transmission of Biometrics Using Authentication Watermarking." In: *Information Engineering (ICIE), 2010 WASE International Conference on*. Vol. 2. Aug. 2010, pp. 105–108. DOI: 10.1109/ICIE.2010.120 (cit. on p. 9).

[16] Jozsef Dudas et al. "Identification of in-field defect development in digital image sensors." In: *SPIE-IS&T/ Vol. 6502 65020Y-1*. 2007 (cit. on p. 12).

[17] J. Dudas et al. "On-Line Mapping of In-Field Defects in Image Sensor Arrays." In: *Defect and Fault Tolerance in VLSI Systems, 2006. DFT '06. 21st IEEE International Symposium on*. Oct. 2006, pp. 439–447. DOI: 10.1109/DFT.2006.48 (cit. on p. 12).

[18] George H. Dunteman. *Principal Components Analysis*. A Sage Publications Nr. 69. SAGE Publications, 1989. ISBN: 9780803931046. URL: http://books.google.de/books?id=Pzwt-CMMt4UC (cit. on p. 49).

[19] J. Fridrich. "Digital image forensics." In: *Signal Processing Magazine, IEEE* 26.2 (Mar. 2009), pp. 26–37. ISSN: 1053-5888. DOI: 10.1109/MSP.2008.931078 (cit. on pp. 4, 5, 11–13, 25, 34, 36, 43, 52, 63, 85).

[20] Jessica Fridrich. "Sensor Defects in Digital Image Forensics." In: *Digital Image Forensics: There is more to a picture than meets the eye*. Ed. by H.T. Sencar and N. Memon. Springer Verlag, 2012. Chap. 6, pp. 179–218 (cit. on pp. 11, 13–15).

[21] G. Bloy. "Blind Camera Fingerprinting and Image Clustering." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30.3 (Mar. 2008), pp. 532–534 (cit. on pp. 43, 62, 63).

[22] T. Gloe, S. Pfennig, and M. Kirchner. "Unexpected artefacts in PRNU-based camera identification: a 'Dresden Image Database' case-study." In: *MM&Sec'12: Proceedings of the 14th ACM Multimedia and Security Workshop*. ACM, Sept. 2012, pp. 109–114 (cit. on p. 19).

[23] Thomas Gloe and Rainer Böhme. "The Dresden Image Database for benchmarking digital image forensics." In: *SAC 2010: Proceedings of the 2010 ACM Symposium on Applied Computing*. ACM, 2010, pp. 1584–1590 (cit. on pp. 4, 14, 19, 85).

[24] Miroslav Goljan and Jessica Fridrich. "Sensor-Fingerprint Based Identification of Images Corrected for Lens Distortion." In: *Proceedings of SPIE* (2012) (cit. on p. 19).

[25] Miroslav Goljan, Jessica Fridrich, and Mo Chen. "Defending Against Fingerprint-Copy Attack in Sensor-Based Camera Identification." In: *IEEE Transactions on Information Security and Forensics* 6.1 (June 2011), pp. 227–236 (cit. on p. 9).

[26] Miroslav Goljan, Jessica Fridrich, and Tomas Filler. "Large Scale Test of Sensor Fingerprint Camera Identification." In: *Proceedings of SPIE, Electronic Imaging, Security and Forensics of Multimedia Contents XI*. San Jose, CA, USA: SPIE, Jan. 18, 2009 (cit. on p. 9).

[27] Dan Goodin. *Fingerprint lock in Samsung Galaxy 5 easily defeated by whitehat hackers*. 2015. URL: http://arstechnica.com/security/2014/04/fingerprint-lock-in-samsung-galaxy-5-easily-defeated-by-whitehat-hackers/ (visited on 05/07/2015) (cit. on p. 8).

[28] J. Hämmerle-Uhl, K. Raab, and A. Uhl. "Experimental Study on the Impact of Robust Watermarking on Iris Recognition Accuracy (Best Paper Award, Applications Track)." In: *Proceedings of the 25th ACM Symposium on Applied Computing*. 2010, pp. 1479–1484 (cit. on p. 9).

[29] J. Hämmerle-Uhl, K. Raab, and A. Uhl. "Watermarking as a Means to Enhance Biometric Systems: A Critical Survey." In: *Proceedings of the 2011 Information Hiding Conference (IH'11)*. Ed. by T. Filler et al. Vol. 6958. Springer LNCS. Prague, Czech Republic, May 2011, pp. 238–254 (cit. on p. 8).

[30] G.E. Healey and R. Kondepudy. "Radiometric CCD camera calibration and noise estimation." In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 16.3 (Mar. 1994), pp. 267–276. ISSN: 0162-8828. DOI: 10.1109/34.276126 (cit. on p. 1).

[31] M. R. Islam, M.S. Sayeed, and A. Samraj. "Biometric Template Protection Using Watermarking with Hidden Password Encryption." In: *Proceedings of International Symposium on Information Technology*. Malaysia, 2008, pp. 296–303 (cit. on p. 9).

[32] J. Fridrich. "Digital Image Forensic Using Sensor Noise." In: *IEEE Signal Processing Magazine* 26.2 (Mar. 2009) (cit. on pp. 11, 15, 61, 62).

[33] Anil K. Jain, Karthik Nandakumar, and Abhishek Nagar. "Biometric Template Security." In: *EURASIP J. Adv. Signal Process* 2008 (Jan. 2008), 113:1–113:17. ISSN: 1110-8657. DOI: 10.1155/2008/579416. URL: http://dx.doi.org/10.1155/2008/579416 (cit. on p. 6).

[34] James R. Janesick et al. "Scientific Charge-Coupled Devices." In: *Optical Engineering* 26.8 (1987). DOI: 10.1117/12.7974139. URL: http://dx.doi.org/10.1117/12.7974139 (cit. on p. 1).

[35] Nalini K. Ratha, Jonathan H. Connell, and Ruud M. Bolle. "Secure data hiding in wavelet compressed fingerprint images." In: *ACM Multimedia 2000*. Los Angeles, CA, USA, Nov. 2000. URL: a%20href=%22http://woodworm.cs.uml.edu/~rprice/ep%22%3Einfo%3C/a (cit. on p. 8).

[36] Xiangui Kang, Guangdong Wu, and K.J.R. Liu. "A context adaptive predictor of sensor pattern noise for camera source identification." In: *Image Processing (ICIP), 2012 19th IEEE International Conference on*. Sept. 2012, pp. 237–240. DOI: 10.1109/ICIP.2012.6466839 (cit. on pp. 20, 85).

[37] Nikos Komninos and Tassos Dimitriou. "Protecting Biometric Templates with Image Watermarking Techniques." In: *ICB*. Ed. by Seong-Whan Lee and Stan Z. Li. Vol. 4642. Lecture Notes in Computer Science. Springer, Aug. 31, 2007, pp. 114–123. ISBN: 978-3-540-74548-8. URL: http://dblp.uni-trier.de/db/conf/icb/icb2007.html#KomninosD07 (cit. on p. 9).

[38] L. Debiasi and A. Uhl. "Techniques for a Forensic Analysis of the CASIA-Iris V4 Database." In: *Proceedings of the 3rd International Workshop on Biometrics and Forensics (IWBF'15)*. 2015 (cit. on pp. v, 47).

[39] L. Debiasi, Z. Sun, and A. Uhl. "Generation of iris sensor PRNU fingerprints from uncorrelated data." In: *Proceedings of the 2nd International Workshop on Biometrics and Forensics (IWBF'14)*. 2014 (cit. on pp. v, 25–27, 31, 32, 84).

[40] Andreas Lang and Jana Dittmann. *Digital watermarking of biometric speech references: impact to the EER system performance*. 2007. DOI: 10.1117/12.703890. URL: http://dx.doi.org/10.1117/12.703890 (cit. on p. 9).

[41] Justin Lee. *Spoofing iris recognition technology with pictures*. 2015. URL: http://www.biometricupdate.com/201503/spoofing-iris-recognition-technology-with-pictures (visited on 05/07/2015) (cit. on p. 8).

[42] J. Leung et al. "Automatic Detection of In-field Defect Growth in Image Sensors." In: *Defect and Fault Tolerance of VLSI Systems, 2008. DFTVS '08. IEEE International Symposium on*. Oct. 2008, pp. 305–313. DOI: 10.1109/DFT.2008.58 (cit. on p. 12).

[43] J. Leung et al. "Quantitative analysis of in-field defects in image sensor arrays." In: *Defect and Fault-Tolerance in VLSI Systems, 2007. DFT '07. 22nd IEEE International Symposium on*. Sept. 2007, pp. 526–534. DOI: 10.1109/DFT.2007.59 (cit. on p. 12).

[44] Chang-Tsun Li. "Source camera identification using enhanced sensor pattern noise." In: *Proceedings of the IEEE Internationel Conference on Image Processing, ICIP '09*. Cairo, Egypt: IEEE, Nov. 7, 2009, pp. 1509–1512 (cit. on pp. 19, 20, 34).

[45] Chang-Tsun Li. "Unsupervised classification of digital images using enhanced sensor pattern noise." In: *ISCAS*. IEEE, 2010, pp. 3429–3432 (cit. on pp. 33, 36, 38).

[46] Chang-Tsun Li and R. Satta. "On the location-dependent quality of the sensor pattern noise and its implication in multimedia forensics." In: *Imaging for Crime Detection and Prevention 2011 (ICDP 2011), 4th International Conference on*. Nov. 2011, pp. 1–6. DOI: 10.1049/ic.2011.0134 (cit. on p. 86).

[47] Ch.-T. Li. "Source camera identification using enhanced sensor pattern noise." In: *IEEE Transactions on Information Forensics and Security* 5.2 (2010), pp. 280–287 (cit. on p. 62).

[48] ChunLei Li et al. "Protecting Biometric Templates Using Authentication Watermarking." In: *PCM (1)*. Ed. by Guoping Qiu et al. Vol. 6297. Lecture Notes in Computer Science. Springer, 2010, pp. 709–718. ISBN: 978-3-642-15701-1. URL: http://dblp.uni-trier.de/db/conf/pcm/pcm2010-1.html#LiMWZ10 (cit. on p. 9).

[49] Ching-yung Lin and Shih-fu Chang. "Generating Robust Digital Signature for Image/Video Authentication." In: *Multimedia and Security Workshop at ACM Multimedia*. 1998 (cit. on p. 8).

[50] Chun-Shien Lu and Hong-Yuan Mark Liao. "Structural Digital Signature for Image Authentication: An Incidental Distortion Resistant Scheme." In: *Proceedings of the 2000 ACM Workshops on Multimedia*. MULTIMEDIA '00. Los Angeles, California, USA: ACM, 2000, pp. 115–118. ISBN: 1-58113-311-1. DOI: 10.1145/357744.357893. URL: http://doi.acm.org/10.1145/357744.357893 (cit. on p. 8).

[51] Shuhan Luan et al. "Silhouette coefficient based approach on cellphone classification for unknown source images." In: *ICC*. IEEE, 2012, pp. 6744–6747. ISBN: 978-1-4577-2052-9 (cit. on p. 42).

[52] Jan Lukas, Jessica J. Fridrich, and Miroslav Goljan. "Digital camera identification from sensor pattern noise." In: *IEEE Transactions on Information Forensics and Security* 1.2 (Oct. 29, 2008), pp. 205–214 (cit. on pp. 13, 14, 20, 61).

[53] Jan Lukas, Jessica Fridrich, and Miroslav Goljan. "Determining digital image origin using sensor imperfections." In: *Proceedings of the SPIE*. Vol. 5685. 1. SPIE, 2005, pp. 249–260.

[54] M. Goljan, J. Fridrich, and T. Filler. "Managing a Large Database of Camera Fingerprints." In: *Proceedings of SPIE, Media Forensics and Security XII*. San Jose, CA, USA: SPIE, Jan. 17, 2010 (cit. on p. 9).

[55] Minerva M. Yeung and Sharath Pankanti. "Verification Watermarks on Fingerprint Recognition and Retrieval." In: *Proceedings of the 11th SPIE Annual Symposium, Electronic Imaging '99, Security and Watermarking of Multimedia Contents*. Ed. by Ping Wah Wong and Edward J. Delp. Vol. 3657. San Jose, CA, USA, Jan. 1999, pp. 66–78 (cit. on p. 8).

[56] Babak Mahdian and Stanislav Saic. "A Bibliography on Blind Methods for Identifying Image Forgery." In: *Image Commun.* 25.6 (July 2010), pp. 389–399. ISSN: 0923-5965. DOI: 10.1016/j.image.2010.05.003. URL: http://dx.doi.org/10.1016/j.image.2010.05.003 (cit. on p. 9).

[57] N.K. Ratha, J.H. Connell, and R.M. Bolle. "Enhancing security and privacy in biometrics-based authentication systems." In: *IBM Systems Journal* 40.3 (Apr. 2001), pp. 614–634. ISSN: 0018-8670 (cit. on p. 6).

[58] Nalini K. Ratha et al. "A Secure Protocol for Data Hiding in Compressed Fingerprint Images." In: *ECCV Workshop BioAW*. Ed. by Davide Maltoni and Anil K. Jain. Vol. 3087. Lecture Notes in Computer Science. Springer, 2004, pp. 205–216. ISBN: 3-540-22499-8. URL: http://dblp.uni-trier.de/db/conf/eccv/eccv2004bioaw.html#RathaFCB04 (cit. on p. 8).

[59] K. Rosenfeld and H.T. Sencar. "A study of the robustness of PRNU-based camera identification." In: *Proceedings of SPIE, Media Forensics and Security XI*. Vol. 7254. San Jose, CA, USA: SPIE, Jan. 18, 2009, pp. 72540M–725408M (cit. on p. 9).

[60] Peter Rousseeuw. "Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis." In: *J. Comput. Appl. Math.* 20.1 (Nov. 1987), pp. 53–65. ISSN: 0377-0427. DOI: 10.1016/0377-0427(87)90125-7. URL: http://dx.doi.org/10.1016/0377-0427(87)90125-7 (cit. on p. 49).

[61] M. Schneider and Shih-Fu Chang. "A robust content based digital signature for image authentication." In: *Image Processing, 1996. Proceedings., International Conference on*. Vol. 3. Sept. 1996, 227–230 vol.3. DOI: 10.1109/ICIP.1996.560425 (cit. on p. 8).

[62] Husrev T. Sencar, Mahalingam Ramkumar, and Ali N. Akansu. *Data hiding fundamentals and applications. Content security in digital multimedia.* English. Amsterdam: Elsevier/Academic Press. xv, 252 p., 2004 (cit. on p. 8).

[63] Chih-Hsuan Tzeng and Wen-Hsiang Tsai. "A New Technique for Authentication of Image/Video for Multimedia Applications." In: *Proceedings of the 2001 Workshop on Multimedia and Security: New Challenges*. MM&Sec '01. Ottawa, Ontario, Canada: ACM, 2001, pp. 23–26. ISBN: 1-58113-393-6. DOI: 10.1145/1232454.1232464. URL: http://doi.acm.org/10.1145/1232454.1232464 (cit. on p. 8).

[64] Andreas Uhl and Yvonne Höller. "Iris-Sensor Authentication using Camera PRNU Fingerprints." In: *Proceedings of the 5th IAPR/IEEE International Conference on Biometrics (ICB'12)*. New Delhi, India, Mar. 2012, pp. 1–8 (cit. on pp. 2, 8, 9, 14, 23, 24, 26, 27, 30, 32, 53, 61, 62, 85).

[65] Sergei Vassilvitskii and David Arthur. "K-means++: The Advantages of Careful Seeding." In: *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. SODA '07. New Orleans, Louisiana: Society for Industrial and Applied Mathematics, 2007, pp. 1027–1035. ISBN: 978-0-898716-24-5. URL: http://dl.acm.org/citation.cfm?id=1283383.1283494 (cit. on p. 48).

[66] Min Wu and Bede Liu. *Multimedia Data Hiding*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2002. ISBN: 0387954260 (cit. on p. 8).