Chapter 21

Identifying Iris Sensors from Iris Images

Luca Debiasi, Christof Kauba and Andreas Uhl

The base component of iris sensors deployed in practical applications is a digital image sensor, mostly supported by a near infra-red (NIR) light source to improve the iris recognition results [1]. These sensors acquire digital images, which are then further processed and inserted into a biometric system's processing chain.

The authenticity and integrity of the acquired iris images plays an important role for the overall security of a biometric system. Ratha *et al.* [2] identified eight stages in a generic biometric system where attacks may occur. Figure 21.1 shows an insertion and presentation attack on an exemplary biometric system. An insertion attack bypasses the biometric sensor by inserting data (biometric sample) into the transmission from the sensor to the feature extractor. This transmission is the most relevant point for an attack on the integrity and authenticity of the acquired iris images, where the iris image inserted during the attack could be acquired with another sensor offsite, even without the knowledge of a genuine user, or a manipulated image to spoof the biometric recognition system. In contrast to the insertion attack, in case of the presentation attack a forged or fake biometric trait, i.e. an artificially manufactured fake fingerprint or a print of an iris image, is presented to the genuine sensor installed in the biometric system. The presentation of a forged biometric trait can usually be detected by deploying different liveness detection systems.

Encryption and other classical authentication techniques like digital signatures or data-hiding have been suggested to secure the previously mentioned transmis-

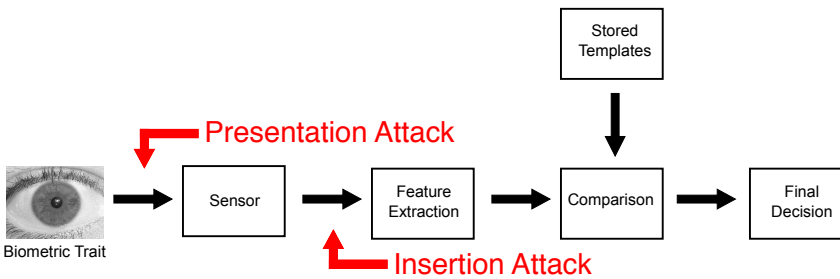

Figure 21.1: Exemplary biometric system and point of insertion and presentation attacks.

sion channel 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 [3, 4] and the digital signature approaches [5, 6, 7, 8]. Höller *et al.* [9] 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 [10].
- **Fragile watermarks** (as proposed for these tasks in e.g. [11, 12, 13]) 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 [14, 15, 16, 17].

Höller *et al.* [9] 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 [18], speech recognition [19], and fingerprint recognition [20].

Passive-blind approaches, in contrast to active methods, do not need any prior information about the image. As stated in [21], 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.

The field of digital image forensics deals with still images and analyzing traces in still image data. Two major tasks in this field are establishing an image's origin and its integrity. In contrast to digital watermarking as authenticity technique, as mentioned in [22], digital image forensics do 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).

To determine an image's origin several approaches have been proposed exploiting hardware and software related artifacts. Investigated hardware related artifacts cover optical defects, like chromatic aberrations [23] or lens distortions [24], or sensor artifacts, like sensor defects [25] and noise. Software artifacts are introduced during the processing of the images in the cameras and can be unveiled using statistical features [26] or by analysing the common image processing pipeline of the images [27].

The photo-response non-uniformity (PRNU) of imaging sensors, as described in [28, 29], 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.

Even though the PRNU is stochastic in nature, it is a relatively stable component of the sensor over its life span, providing a unique sensor fingerprint with the following important properties [29]:

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 (except for ageing related sensor defects) 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.

Slight variations of individual pixels during the conversion of photons to electrons in digital image sensors are the source of the PRNU, thus it is considered an intrinsic property which is present in all digital imaging sensors. Every digital image sensor adds this weak, noise-like pattern into every image acquired with it. The sensor identification can be performed at different levels, as described by Bartlow *et al.* [30]: Technology, brand, model, unit. Due to the datasets evaluated in this work we focus on the model level, which corresponds to a differentiation according to model and brand.

The PRNU can also be used for the verification of an image's integrity. The integrity is compromised if an 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, altered). These manipulations lead to changes in the PRNU which can be detected as shown in [31, 32].

In the context of biometric systems security the PRNU fingerprint of a sensor can be used to ensure the integrity and authenticity of images acquired with a biometric sensor. Höller *et al.* [9] 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 sensor's PRNU [31]. Besides image integrity, this technique can also provide authenticity by identifying the source sensor uniquely and impor-

tant properties as required in a biometric scenario have been demonstrated: suitability to manage large datasets [33, 34], robustness against common signal processing operations like compression and malicious signal processing [35, 36], and finally methodology to reveal forged PRNU fingerprints has been established [37].

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 [9] and [30] using the PRNU. The results from Höller *et al.* [9], where the discriminative power of five iris sensors from the *CASIA-Iris V4* database has been evaluated show high variations. Other work by Kalka *et al.* [38] regarding the differentiability of iris sensor showed varying results, while studies conducted on fingerprint sensors by Bartlow *et al.* [30] showed more satisfactory results. In order for PRNU fingerprints beeing useful as an authentication measure for biometric systems, the sources of the poor differentiation results have to be determined. Some possible explanations are given in [38] and [9] and consist of the highly correlated data of biometric datasets, saturated pixels and the use of multiple sensors of the same model. An additional caveat for the PRNU extraction is the image content. Since the PRNU covers the high frequency components of an image, it is contaminated with other high frequency components within the images, such as edges. Li [39] proposed an approach for attenuating the influence of details from scenes on the PRNU so as to improve the device identification rate of the identifier. Moreover the PRNU fingerprint can be extracted from images of a biometric sensor and injected into forged images, as described by Goljan *et al.* [40]. Using several images captured by the sensor deployed in the biometric system a suitable PRNU fingerprint can be generated by the attacker. This attack can only be detected with a triangle test [9], which requires additional genuine images acquired under controlled conditions.

To overcome the reported problems with the PRNU extraction for some sensors and the injection attack, there exist other approaches like the one by El-Naggar and Ross [41], who proposed a passive approach tailored to iris recognition. At first the ocular image is segmented to get the iris region, then the iris texture is unwrapped, followed by a normalisation step to get a normalised iris image. Only the inner half of this normalised iris image is used and further split into a set of overlapping blocks. For each block 50 Gabor and 68 statistical features are extracted to form a 118 dimensional feature vector representing the iris image. These feature vectors are then classified using a 3-layer artificial neural network. They were able to achieve accuracies of $80-85\%$. We propose a similar approach which follows their evaluation methodology but uses different features and a SVM classifier. In this chapter we evaluate this approach and a PRNU based one, trying to identify the iris data set which an iris image belongs to. If the correct data set can be determined for a given iris image, these approaches could be used to secure an iris recognition system against insertion attacks.

The chapter is organised as follows: In Section 21.1 we describe the two different approaches and the examined iris data sets are listed in Section 21.2. The experimental setup and the results are illustrated in Section 21.3 and 21.4, respectively. In Section 21.5 we discuss how the previously examined techniques can be used in a practical application. Finally Section 21.6 concludes the chapter.

21.1 Techniques for Sensor Identification/Dataset Classification

In this section we present two different techniques that allow to infer which dataset an iris image originates from. The first technique, PRNU based Sensor Identification (PSI), does this by identifying the sensor used to acquire the image. The second technique, Iris Texture Classification (ITC), makes use of the iris texture and its inherent features to classify the iris images according to the source sensor. Both techniques are presented in detail in the following section.

21.1.1 PRNU based Sensor Identification (PSI)

A digital image sensor consists of lots of small photosensitive, usually rectangular detectors that capture the incident light and generate an electric signal. These detectors are commonly known as pixels. The image acquired with the sensor is constructed by the aggregate of all pixels. Due to imperfections in the manufacturing and the inhomogeneity of the manufacturing material, silicon, the efficiency of each pixel of converting photons to electrons varies slightly. According to Fridrich [29], the raw output of a sensor with $w \times h$ pixels can be modeled as:

$$\begin{aligned} Y &= I + I \circ K + \tau D + C + \Theta \\ with \quad & Y, I, K, D, C, \Theta \in \mathbb{R}^{w \times h}; \tau \in \mathbb{R} \end{aligned} \tag{21.1}$$

where Y is the sensor output (image). I represents the incoming light, $I \circ K$ 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 (e.g. readout noise, shot noise or photonic noise, quantization noise, etc.). Since all pixels are independent 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 $\theta \in \Theta$.

The extraction of the PRNU noise residuals is performed as indicated by Fridrich in [42]. For each image I the noise residual W_I is estimated:

$$W_I = I - F(I) \tag{21.2}$$

where F is a denoising function filtering out the sensor pattern noise. We used two different denoising techniques to extract the PRNU from the images: The wavelet-based denoising filter as described in Appendix A of [43] and the BM3D filter proposed in [44], which is reported to produce better and more consistent results in filtering out the PRNU in [45]. The extracted PRNU noise residual is then normalised in respect to the L_2-norm because its embedding strength is varying between different sensors as explained by [9]. As additional post processing steps a zero mean operation is applied to each extracted PRNU noise residual to suppress artifacts with regular grid structure.

To reduce the PRNU contamination effect from scene details, we apply an image content attenuating PRNU enhancement technique (Model 3 in [39]), subsequently denoted as *ELi*.

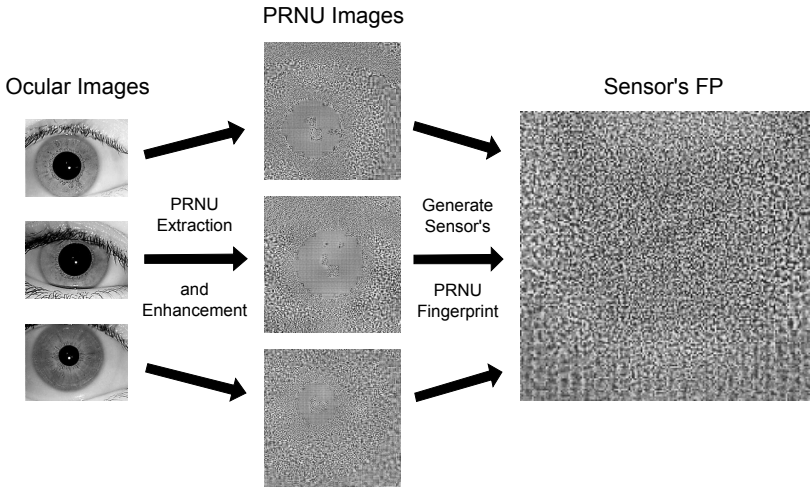

Figure 21.2: PRNU noise residual extraction and fingerprint generation with multiple iris images of the same sensor.

Estimating a sensor's PRNU from a single image is usually not sufficient, because that specific image may contain various kinds of disturbances as modeled by Θ in Equation (21.1). Thus multiple images from the same sensor are averaged to isolate the systematic components of all images and suppress these random noise components, as shown in Figure 21.2. This averaged noise is denoted as PRNU fingerprint or reference pattern noise (RPN) in literature. The PRNU fingerprint $\hat{K}$ of a sensor is then estimated using a maximum likelihood estimator for images I_i with $i = 1...N$.

$$\hat{K} = \frac{\sum_{i=1}^{N} W_I^i I^i}{\sum_{i=1}^{N} (I^i)^2} \tag{21.3}$$

The PRNU fingerprint is enhanced using a Wiener filter applied in the DFT domain to suppress periodic artifacts as described in [46].

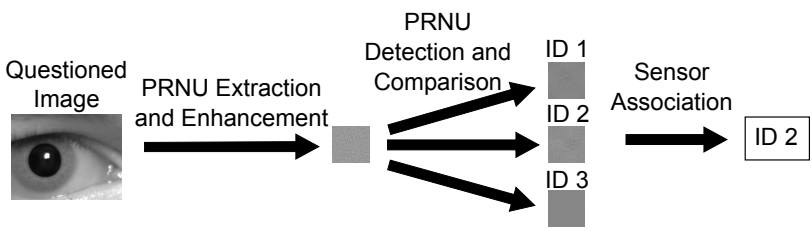

Figure 21.3: PRNU noise residual extraction and identification of corresponding sensor.

To determine if an image has been acquired with a specific sensor, the presence of a sensor's PRNU fingerprint in the questioned image has to be detected. Since

images acquired with iris sensors are usually not geometrically transformed, this can be done by means of calculating the normalised Cross Correlation (NCC):

$$\rho_{[J,\hat{K}]} = NCC(W_J, J\hat{K}) \tag{21.4}$$

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

An alternative correlation measure to detect the presence of a PRNU fingerprint $\hat{K}$ in an Image I is the Peak Correlation Energy (PCE), proposed by Fridrich in [46]. Fridrich notes that with the PCE the detection threshold will not vary as much as for NCC detector with varying signal length, different cameras and their on-board image processing. It is applied like the NCC detector:

$$\rho_{[J,\hat{K}]} = PCE(W_J, J\hat{K}) \tag{21.5}$$

A schematic illustration for the detection of the correct PRNU fingerprint in a questioned image is given in Figure 21.3.

21.1.2 Iris Texture Classification (ITC)

The input for the Iris Texture Classification (ITC) approach are the preprocessed, segmented, unrolled and normalised iris images originating from various different iris datasets and the output of the classifier is a prediction of the iris sensor used to capture the image or the dataset where the input iris image belongs to, respectively. As the ITC is SVM based, a training phase is needed prior to the use of the classifier, similar to generating a PRNU fingerprint for the PSI approach. In the following the three chosen feature extraction methods, namely DenseSIFT, DMD and LBP are briefly explained. Then the classification approach using a GMM, Fisher Vector encoding and a SVM classifier is described.

21.1.2.1 Feature Extraction

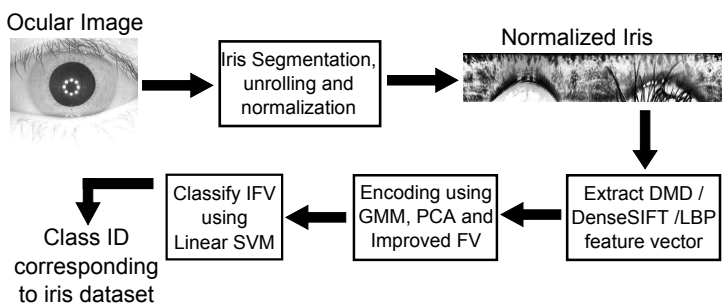

Figure 21.4: Flowchart of the Iris Texture Classification (ITC) approach.

DenseSIFT: Is a variant of SIFT. SIFT, the scale invariant feature transform, is a general purpose feature extraction technique used in object recognition proposed

by Lowe [47]. It is invariant to image scale and rotation and robust against various affine distortions, addition of noise, illumination changes and changes of the viewpoint. SIFT locates extrema in the scale-space, localises keypoints, determines their dominant orientation and finally constructs a local descriptor for the keypoint based on a region around it. Fei-Fei et al. [48] proposed to use the local SIFT descriptors on a predefined grid defined across the whole image instead of localising their positions according to scale space extrema. This approach is known as dense SIFT. A 128-dimensional SIFT feature vector is extracted each 3 pixels in 5 different scales (2^0, $2^{-1/2}$, 2^{-1}, $2^{-3/2}$, 2^{-2}). The spatial bins of the SIFT feature descriptor histogram consist of 4 bins in x, 4 bins in y and 8 orientation bins. vl_feat's (`http://www.vlfeat.org`) implementation of DenseSIFT is utilised.

DMD: Dense Micro-block difference is a local feature extraction and texture classification technique proposed by Mehta and Egiazarian [49]. It captures the local structure from image patches (9×9 to 15×15 pixels) at high scales. Instead of the pixels, small blocks of the image which capture the micro-structure are processed. Therefore the pairwise intensity differences of smaller blocks (e.g. 2×2 or 3×3 pixel blocks) calculated in several different directions (not only radial like in LBP) in combination with the average intensity of the whole patch are used to encode the local structure of the patch. Difference values of block pairs located near the centre of the patch are given higher weights than blocks towards the patch boundaries. This should be able to capture the repetitively characteristic local structure providing discriminative information.

LBP: The local binary patterns proposed by Ojala [50] observe the variations of pixels in a local neighborhood. These variations are thresholded by the central pixel value to obtain a binary decision, which is then encoded as a scalar value. The occurrences of each scalar value for all pixels in the image are represented in a histogram, which forms the extracted feature vector.

21.1.2.2 Feature Encoding

We utilize the Improved Fisher Vector Encoding (IFV) scheme [51] in the same way as it is done in [52, 49]. IFV is usually used in object recognition. Fisher vector encoding starts by extracting local SIFT descriptors densly (DenseSIFT) and at multiple scales to get a feature vector f. We not only use DenseSIFT features but also DMD and LBP ones as input for the next steps. The feature vector f is then soft-quantised using a Gaussian Mixture Model (GMM) with K modes where the Gaussian covariance matrices are assumed to be diagonal. The local descriptors present in f are first decorrelated and then dimensionality reduced (optional) by PCA. So far this describes the standard Fisher Vector encoding [53]. The IFV now adds signed square rooting and l^2 normalization as described in [51].

21.1.2.3 Classification

A linear SVM is then used to classify the IFV encoded features. We experimented with different types of kernels $K(x',x'')$ (linear, Hellinger, exponential) and the linear kernel lead to the most promising results. The input data to the SVM (IFV encoded feature vectors) is normalised such that $K(x',x'') = 1$ which usually improves the

performance. The SVM is trained using a standard non-linear SVM solver on a subset of the unrolled, normalised iris images which is subsequently not used for the testing (evaluation) step.

21.2 Datasets

To enable a meaningful comparision with the previous work of El-Naggar and Ross [41] we attempted to use the same iris datasets they originally used and extend the number of datasets. Unfortunately we were not able to acquire the MGBC and the WVU iris dataset. Thus we use the remaining 6 datasets they used plus 3 additional iris dataset which are described in the following. All of them are publicly available, common datasets which have been utilised in many different iris recognition related works. Figure 21.5 shows some example images for each of the datasets. Table 21.1 summarizes the most important attributes of the datasets. In the following a short description of each single dataset is given:

CASIA V2: We use the first subset of the CASIA V2 iris database [54] (device 1). This subset consists of 1200 images and was captured using an OKI Irispass-h sensor by the Chinese Academy of Sciences Institue of Automation (CASIA).
CASIA V3: was again captured by the Chinese Academy of Sciences Institue of Automation (CASIA) [54] and consists of several different subsets. We used the CASIA V3 Interval subset in accordance with the work of El-Naggar and Ross. This subset consists of 2639 images captured with a self-developed close-up iris camera.
CASIA V4: This is the V4 version of the iris dataset provided by CASIA [54]. Again it consists of several subsets, where we used the Thousands subset. It consists of 20000 images which were collected using an IrisKing IKEMB-100 camera.
ICE2005: NIST, the National Institute of Standards and Technology in the US conducted a series of biometric recognition contests, one of them was the Iris Challenge Evaluation (ICE) in 2005. The ICE2005 [55] images where captured at the University of Notre Dame with a LG EOU 2200 iris camera and consists of 2953 images.
IITD: The IIT Delhi Iris Database [56] consists of 1120 images and was acquired by the Biometrics Research Laboratory in the Indian Institute of Technology Delhi (IITD) in 2007. The images were captures with an JIRIS JPC1000 digital CMOS iris camera.
MMU2: The MMU V2 iris database [57] consists of 995 iris images. These images are collected using a Panasonic BM-ET100US Authenticam.
UBIRIS: The Noisy Visible Wavelength Iris Image Database UBIRIS V1 [58] consists of 1877 images collected in 2004. The images were captured with a Nikon E5700 digital camera in two sessions.
UPOL: The Univerzita Palackho v Olomouci iris dataset [59] consists of 384 images. The irises were scanned by TOPCON TRC50IA optical device connected with SONY DXC-950P 3CCD camera.
UTIRIS: University of Tehran IRIS (UTIRIS) image dataset [60] consists of two different sessions, one captured using visible wavelength illumination and the other one using near-infrared illumination. We only used the near infrared subset, which consists of 793 images. The infrared images were captured with an ISG Lightwise LW iris camera.

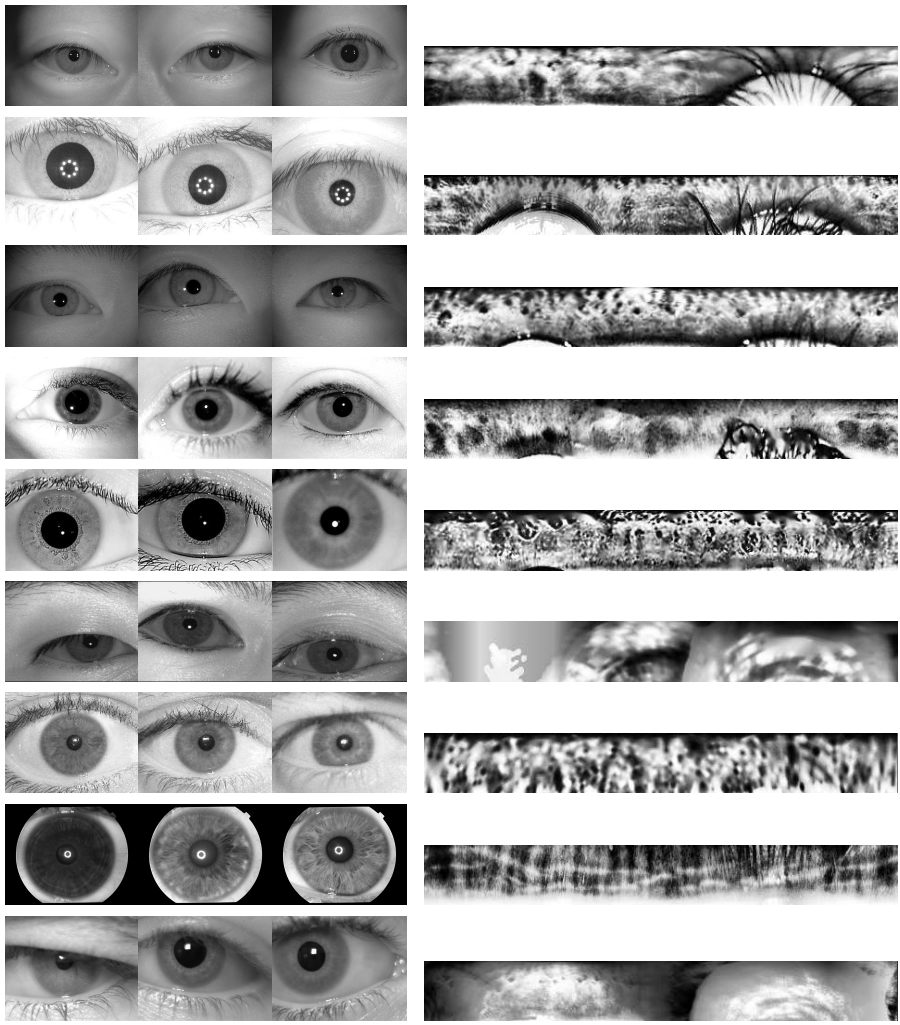

Figure 21.5: Ocular image and normalised iris image samples from different datasets, from top to bottom: CASIA V2, CASIA V3, CASIA V4, ICE2005, IITD, MMU2, UBIRIS, UPOL, UTIRIS

Dataset	# IMG	Sensor	Illumination	Resolution	Class ID
CASIA V2	1200	OKI IRISPASS-h	near infrared	480×640	1
CASIA V3	2639	CASIA Iris camera	near infrared	320×280	2
CASIA V4	20000	IrisKing IKEMB-100	near infrared	640×480	3
ICE2005	2953	LG EOU 2200 iris camera	near infrared	480×640	4
IITD	1120	JIRIS, JPC1000 camera	near infrared	240×320	5
MMU2	995	Panasonic BM-ET100US Authenticam	near infrared	320×238	6
UBIRIS	1877	Nikon E5700	natural lighting	200×150	7
UPOL	384	SONY DXC-950P 3CCD camera	camera flash	768×576	8
UTIRIS	793	ISG Lightwise LW	near infrared	1000×776	9

Table 21.1: Attributes of iris datasets.

21.3 Experimental Setup

We follow the same test methodology as El-Naggar and Ross [41]. For the Iris Texture Classification (ITC) approach as described in Section 21.2 each dataset is randomly split into two distinct subsets, a training and a testing one. UPOL is the iris dataset containing the least images, 384 images only, thus it is split 50:50 into 192 training and 192 testing images. Consequently, for all other images also 192 training and 192 testing images are chosen for the corresponding subsets. The first step in the processing chain is the preprocessing of the ocular images, including iris segmentation and iris unrolling. The unrolled iris patches are then normalised and all having a size 512×64 pixels. This is done utilizing the USIT (University of Salzburg Iris Toolkit, Version 2.0 available at `http://www.wavelab.at/sources/USIT/`) software toolkit in version 1.0.3. For the segmentation step the WAHET (Weighted Adaptive Hough and Ellipsopolar Transform) method is used. Figure 21.5 shows one example of an unrolled and normalised iris image for each dataset. For further details on the exact implementation of WAHET and the iris unrolling the interested reader is referred to [61]. The next step is the feature extraction using DenseSIFT, DMD and LBP. Afterwards, the features are dimensionality reduced using a GMM and then Fisher Vector encoding is applied before they are put into a linear SVM for classification. A 5-fold cross validation is performed and the mean results of all 5 runs are used as final results shown below.

For the PRNU based Sensor Identification (PSI) approach we decided to extract the PRNU from a central patch with varying sizes ranging from 64×64 up to 576×576 pixels because of the varying image size of the data sets. We furthermore evaluate the effect of applying the content attenuation PRNU enhancement *ELi*, described in section 21.1.1, in contrast to not applying it. The configurations for the extraction and post-processing of the PRNU are: Extracted PRNU sizes (from 64×64 up to 576×576 pixels), denoising filters (Wavelet and BM3D), PRNU enhancements (ELi and NoEnh) and PRNU detectors (NCC and PCE). Due to the different image sizes of the datasets the number of sensors to discriminate decreases for an increasing PRNU size as shown by the value in parentheses next to the PNRU size in Table 21.8. For each run we selected 192 random images for each data set for the

Method	DenseSIFT	DMD	LBP
mAcc	0.9838	0.9688	0.8715
mAP	0.9968	0.9878	0.9172

Table 21.2: Mean accuracies (mACC) and mean average precisions (mAP) for DenseSIFT, DMD and LBP

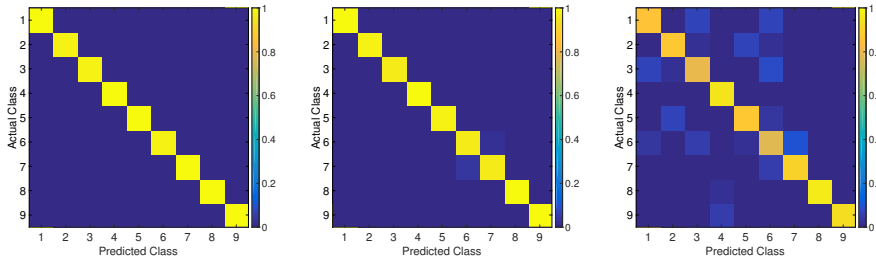

Figure 21.6: Confusion matrix for DenseSIFT, DMD and LBP

generation of the PRNU fingerprint ("training" set) and another 192 random images as the test set, without overlapping images between both sets. We compute the NCC and PCE correlation scores for all test images with all generated PRNU fingerprints, where the predicted sensor (or class) is determined by means of the highest (rank one) correlation score. The larger the size of the extracted PRNU, the less sensors could be used for evaluating the identification performance for the sensors. The experiment was repeated 5 times (5-fold cross validation), where the final result is the average of all 5 runs. The described parameters have been chosen to make the results of both identification/classification approaches as comparable as possible.

21.4 Experimental Results

Table 21.2 summarises the results of the ICT approach. It lists the mean accuracy (mAcc) as well as the mean average precision (mAP) over all 5 runs. The accuracy describes the number of correctly classified items (true positives + true negatives) over the number of total items per class calculated per class. The mAcc is just the mean over all single accuracies. The average precision (AP) describes the area under the precision/recall curve calculated per query/class. The mAP is the mean over all AP values. It can be clearly seen that the ICT approach works best using DenseSIFT features. Using DMD and LBP the recognition performance both in terms of the mACC and the mAP is still clearly over 90%. The results show that the ICT approach is able to determine the source of an unrolled iris texture image with a very high accuracy considering the nine iris datasets.

Figure 21.6 shows the confusion matrices for ICT. The numbers on the axes are corresponding to the class IDs in table 21.1. Considering DenseSIFT it can be seen

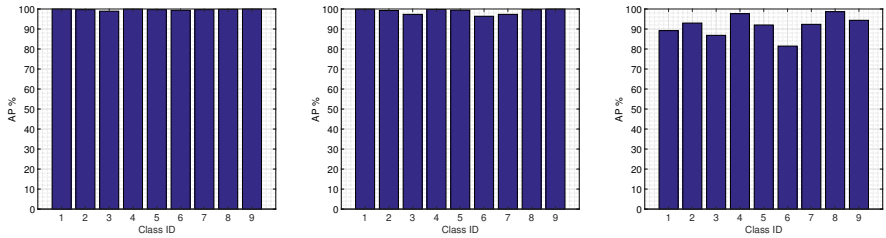

Figure 21.7: Average precision for DenseSIFT, DMD and LBP

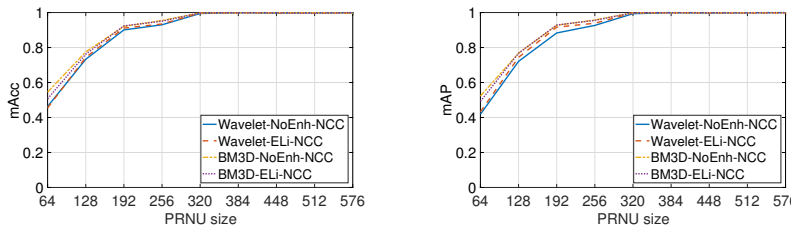

Figure 21.8: Mean accuracy (mAcc) and mean average precision (mAP) for selected PRNU patch sizes and PRNU extraction configurations.

that for CASIA V2, ICE2005, MMU2, UBIRIS, UPOL and UTIRIS all images are correctly classified as belonging to the actual dataset. Only some of the CASIA V3, CASIA V4 and IITD images are misclassified.

Figure 21.7 shows the average precision plots for ICT. Again it can be seen that classification works perfectly for CASIA V2, MMU2, UBIRIS, UPOL, ICE2005 and UTIRIS considering DenseSIFT. Considering DMD it still works perfectly for CASIA V2, ICE2005, UPOL and UTIRIS but no longer for MMU2 and UBIRIS though still quite acceptably. LBP's performance is a bit worse.

Table 21.3 lists the PSI results which show that the PRNU size affects the identification performance most across all configurations. Reasonable mAcc and mAP rates can already be achieved with 192×192 pixel patches, while a patch size larger than 320×320 yields very good results for the identification of the different iris sensors through their PRNU fingerprint. Neither the choice of PRNU detector nor PRNU enhancement makes a big difference in this case, but better results can be achieved by choosing the BM3D denoising filter over the Wavelet filter for smaller PRNU sizes, as shown in Figure 21.8.

Next we are having a closer look at the results for the single classes or sensors. Figure 21.9 shows the confusion matrix for both denoising filters using a small patch size of 64×64 pixels, no content attenuation PRNU enhancement (NoEnh) and the NCC detector. It can be seen that the identification performance varies highly among the different classes, where class 9 shows very good results and the classes 1, 2 and 7 show very low identification performance independent of the denoising filter. All other classes show higher accuracies when the BM3D filter is used.

		Wavelet				BM3D			
	PRNU	NoEnh		ELi		NoEnh		ELi	
	Size	NCC	PCE	NCC	PCE	NCC	PCE	NCC	PCE
mAcc	64 (9)	0.4633	0.4587	0.4545	0.4456	0.5451	0.5397	0.5068	0.4845
	128 (9)	0.7326	0.7300	0.7437	0.7380	0.7752	0.7696	0.7620	0.7554
	192 (8)	0.9007	0.8932	0.9112	0.9008	0.9210	0.9279	0.9237	0.9197
	256 (6)	0.9300	0.9358	0.9347	0.9326	0.9545	0.9507	0.9505	0.9446
	320 (5)	0.9946	0.9963	0.9973	0.9988	0.9981	0.9994	0.9981	0.9983
	384 (5)	0.9988	0.9987	0.9990	0.9990	0.9983	0.9981	0.9992	0.9992
	448 (5)	0.9973	0.9988	0.9983	0.9992	0.9990	0.9998	0.9983	0.9971
	512 (2)	0.9990	0.9974	0.9984	0.9932	0.9984	0.9990	0.9958	0.9943
	576 (2)	0.9984	0.9979	0.9974	0.9995	0.9964	0.9974	0.9984	0.9964
mAP	64 (9)	0.4184	0.4033	0.4325	0.4135	0.5227	0.5146	0.4925	0.4691
	128 (9)	0.7209	0.7114	0.7438	0.7352	0.7674	0.7619	0.7675	0.7572
	192 (8)	0.8832	0.8844	0.9166	0.9092	0.9271	0.9300	0.9287	0.9236
	256 (6)	0.9264	0.9268	0.9389	0.9384	0.9576	0.9530	0.9549	0.9506
	320 (5)	0.9934	0.9949	0.9977	0.9986	0.9985	0.9992	0.9989	0.9986
	384 (5)	0.9989	0.9989	0.9992	0.9990	0.9988	0.9987	0.9994	0.9994
	448 (5)	0.9976	0.9988	0.9985	0.9988	0.9991	0.9994	0.9986	0.9977
	512 (2)	0.9991	0.9984	0.9987	0.9966	0.9990	0.9997	0.9977	0.9973
	576 (2)	0.9995	0.9989	0.9990	0.9998	0.9982	0.9991	0.9994	0.9982

Table 21.3: Mean accuracy (mAcc) and mean average precision (mAP) for all tested PRNU patch sizes and PRNU extraction configurations. The number in parentheses next to the PRNU size indicates the number of different sensors.

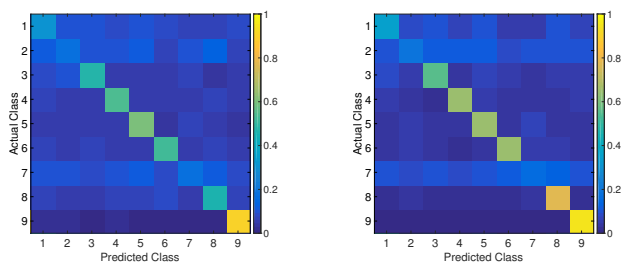

Figure 21.9: Confusion matrices using the Wavelet (left) and BM3D (right) denoising filters for 64×64 *pixels PRNU patch size, no content attenuation PRNU enhancement (NoEnh) and NCC detector.*

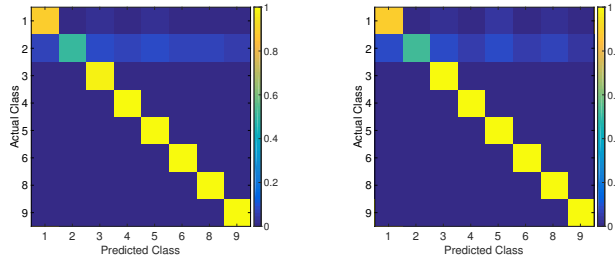

Figure 21.10: Confusion matrices using the NCC (left) and PCE (right) detectors for 192×192 *pixels PRNU patch size, ELi content attenuation PRNU enhancement and BM3D denoising filter.*

TS size	DenseSIFT	DMD	LBP	PRNU128	PRNU256	PRNU512
192	0.9937	0.9810	0.9116	0.8227	0.9540	0.9999
96	0.9919	0.9547	0.8143	0.7870	0.9348	0.9979
48	0.9833	0.9564	0.5038	0.7426	0.9069	0.9978
24	0.9668	0.9277	-	0.7073	0.8594	0.9937
12	0.9367	0.8805	-	0.6363	0.8003	0.9796
6	0.8766	0.8062	-	0.5244	0.7476	0.9565
3	0.7897	0.6921	-	0.4817	0.6905	0.9348
1	0.6320	0.5749	-	0.3297	0.5734	0.8623

Table 21.4: Mean average precisions (mAP) for DenseSIFT, DMD, LBP, and PRNU with sizes 128×128*,* 256×256 *and* 512×512 *for different training set sizes (TS size).*

Having a look at a larger PNRU size of 192×192 pixels, BM3D denoising filter and ELi PRNU enhancement, as shown in Figure 21.10, reveals that the identification performance for both detectors, NCC and PCE, is practically identical with only slight differences for all classes. This indicates that the choice of detector is not critical for the overall performance.

The ITC approach in general and the PSI approach with PRNU sizes at least 512×512 pixels outperform the approach by El-Naggar and Ross [41].

21.5 Practical Discussion

In order to secure a biometric system against insertion attacks (described in the introduction) the authenticity of the biometric samples, i.e. images, has to be verified. We examined two approaches tailored to identify the sensor an iris image was captured with. Both approaches can be used for existing biometric systems and while setting-up new ones because they rely solely on intrinsic image properties. The first one is based on the PRNU of the iris sensors and the second one on texture features. We examined different training set sizes. The results shown in Table 21.4, indicate that both achieve good classification results if some conditions are met. The ITC ap-

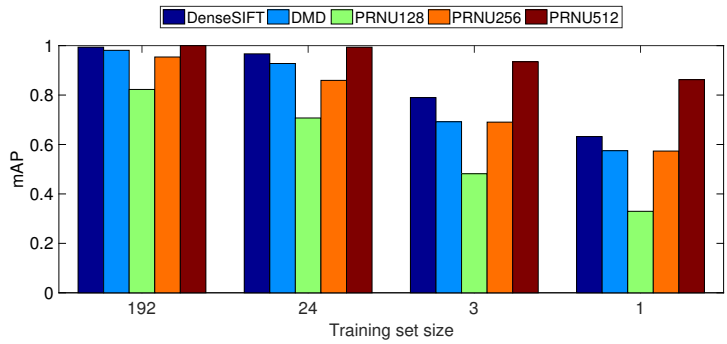

Figure 21.11: Exemplary mean average precision (mAP) scores for selected training set sizes.

proach works well for a broad range of image resolutions as long as there are enough training images available, i.e. at least 10 training images should be available. Using the LBP feature extractor for the ITC approach causes the training to fail for smaller training set sizes, as shown by missing values in Table 21.4. In these cases the LBP feature vectors are not distinctive enough and cannot be soft-quantized by the GMM. The different image resolutions and iris sizes in the images among the datasets require varying unrolling parameters. Therefore, unrolling and normalisation cause a separate level of interpolation for each dataset. Due to the texture classification nature of the ITC approach the results could be biased by the interpolation artifacts that may improve the discriminative power of the datasets but not the sensors themselves. This effect could eventually be mitigated by using the ocular images as input in combination with different features like BSIF [62], which is designed to capture image characteristics similar to the PRNU.

The PSI approach works well for bigger PRNU sizes and also works for very few training images, c.f. it even works with one single training image for the PRNU512 case. The PRNU is extracted directly from the ocular images as opposed to the unrolled iris texture in the ITC approach. Thus no additional bias is introduced and the discrimination relies solely on the sensors' characteristics.

In this work we only investigated different types of sensors, but not multiple sensors of the same model and manufacturer. As mentioned in the introduction the PRNU is able to distinguish between specific sensor instances of the same model, which is an advantage over the ITC approach in practical deployments since the attacker may have access to the same sensor model. It needs to be clarified whether the ITC approach is able to handle this kind of set-up as well.

Both approaches are suitable if it comes to securing a biometric system depending on the system's configuration. For biometric systems dealing with smaller images but with many training images available the ITC approach is favourable, for systems dealing with larger images but only very few training images available the PSI approach should be used. This is further illustrated in Figure 21.11. If only few training images are available and the images are small then a fusion of the two ap-

proaches could improve the results. The image size cannot be changed easily but it is easy to provide some more training images (just capture additional data with the sensor) and thus the ITC approach can be used again. For securing a biometric system at first the respective approach has to be trained (PRNU fingerprint generation for the PSI approach) using images of the biometric sensor(s). Every time a new biometric sample is captured the image is analysed using the pretrained classifier which then tells if the image was captured by one of the biometric sensors it was trained with or not. In the latter case it is very likely that an insertion attack happened and the authentication process is aborted.

21.6 Conclusion

In this chapter we examined two passive approaches to secure an iris recognition system against insertion attacks by verifying the authenticity of the iris images. The first one, named PSI, is based on the photo response non-uniformity (PRNU) of image sensors and the second one, named ITC, exploits the texture information of unrolled iris images. The examination was performed using images from 9 distinct iris databases or sensors, respectively.

The results show that both approaches perform well in identifying the correct sensor an iris image was captured with, though the performance of the PSI approach is dependent on the size of the extracted PRNU. The ITC approach worked well for all datasets. We furthermore examined the impact of the number of images available for the training of both approaches.

Each approach has its advantages and drawbacks depending on the configuration of the biometric system: The PSI approach gives better results if only a small number of high-resolution images is available, while the ITC approach needs a higher number of images to achieve an acceptable performance, but the advantage is they do not need to be of high resolution. In addition the PSI approach is also suited to distinguish different sensors of the same model. This helps in detecting an injection of images from the same sensor model as deployed in the biometric system. If only a small number of low resolution images is available, a fusion of both approaches is likely to improve the overall performance.

Bibliography

[1] J. Daugman, "How iris recognition works," *International Conference on Image Processing*, vol. 1, pp. I–33–I–36, 2002.

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

[3] H. T. Sencar, M. Ramkumar, and A. N. Akansu, *Data hiding fundamentals and applications. Content security in digital multimedia.*, 2004.

[4] C. Wu and C. Kuo, "Comparison of two speech content authentication approaches," in *Proceedings of SPIE, Security and Watermarking of Multimedia Contents IV*, vol. 4675, San Jose, CA, USA, 2002.

[5] M. Schneider and S.-F. Chang, "A robust content based digital signature for image authentication," in *Proceedings of the IEEE International Conference on Image Processing (ICIP'96)*, Lausanne, Switzerland, 1996.

[6] J. Tzeng, W.-L. Hwang, and I. Chern, "Enhancing image watermarking methods by second order statistics," in *Proceedings of the IEEE International Conference on Image Processing (ICIP'01)*, Thessaloniki, Greece, 2001.

[7] C.-S. Lu and H.-Y. M. Liao, "Oblivious watermarking using generalized gaussian," in *Proceedings of the 7th International Conference on Fuzzy Theory and Technology*, Atlantic City, NJ, USA, 2000, pp. 260–263.

[8] W.-K. Lin and N. Burgess, "Listless zerotree coding for color images," in *32nd Asilomar Conference on Signals, System and Computers*, CA, USA, 1998.

[9] A. Uhl and Y. Höller, "Iris-sensor authentication using camera PRNU fingerprints," in *Proceedings of the 5th IAPR/IEEE International Conference on Biometrics (ICB'12)*, New Delhi, India, 2012, pp. 1–8.

[10] 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)*, ser. Springer LNCS, vol. 6958, Prague, Czech Republic, 2011, pp. 238–254.

[11] M. M. Yeung and S. 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*, vol. 3657, San Jose, CA, USA, 1999, pp. 66–78.

[12] N. K. Ratha, J. H. Connell, and R. M. Bolle, "Secure data hiding in wavelet compressed fingerprint images," in *ACM Multimedia 2000*, Los Angeles, CA, USA, 2000.

[13] N. K. Ratha, M. A. Figueroa-Villanueva, J. H. Connell, and R. M. Bolle, "A secure protocol for data hiding in compressed fingerprint images." in *ECCV Workshop BioAW*, ser. Lecture Notes in Computer Science, vol. 3087, 2004, pp. 205–216.

[14] N. Komninos and T. Dimitriou, "Protecting biometric templates with image watermarking techniques." in *ICB*, ser. Lecture Notes in Computer Science, vol. 4642, 2007, pp. 114–123.

[15] L. Li, C. S. Tong, and S. K. Choy, "Texture classification using refined histogram," *IEEE Transactions on Image Processing*, vol. 19, no. 5, pp. 1371–1378, 2010.

[16] S. Ding, C. Li, and Z. Liu, "Protecting hidden transmission of biometrics using authentication watermarking," in *Information Engineering (ICIE), 2010 WASE International Conference on*, vol. 2, 2010, pp. 105–108.

[17] F. Ahmed and I. 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, 2005, pp. 137–142.

[18] 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.

[19] A. Lang and J. Dittmann, "Digital watermarking of biometric speech references: impact to the eer system performance," in *Electronic Imaging 2007*. International Society for Optics and Photonics, 2007, pp. 650 513–650 513.

[20] M. R. Islam, M. Sayeed, and A. Samraj, "Biometric template protection using watermarking with hidden password encryption," in *Proceedings of International Symposium on Information Technology*, 2008, pp. 296–303.

[21] B. Mahdian and S. Saic, "A bibliography on blind methods for identifying image forgery," *Image Commun.*, vol. 25, no. 6, pp. 389–399, 2010.

[22] T. Gloe and R. Böhme, "The dresden image database for benchmarking digital image forensics," in *SAC 2010: Proceedings of the 2010 ACM Symposium on Applied Computing*, 2010, pp. 1584–1590.

[23] M. K. Johnson and H. Farid, "Exposing digital forgeries in complex lighting environments," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 3, pp. 450–461, 2007.

[24] K. Choi, E. Lam, and K. Wong, "Automatic source camera identification using the intrinsic lens radial distortion," *OPTICS EXPRESS*, vol. 14, no. 24, pp. 11 551–65, 2006.

[25] Z. Geradts, J. Bijhold, M. Kieft, K. Kurosawa, K. Kuroki, and N. Saitoh, "Methods for identification of images acquired with digital cameras," in *Proceedings of SPIE, Enabling Technologies for Law Enforcement and Security*, vol. 4232, 2001, pp. 505–512.

[26] M. Kharrazi, H. T. Sencar, and N. Memon, "Blind source camera identification," in *Image Processing, 2004. ICIP '04. 2004 International Conference on*, vol. 1, 2004, pp. 709–712 Vol. 1.

[27] S. Bayram, H. Sencar, N. Memon, and I. Avcibas, "Source camera identification based on CFA interpolation," in *Proceedings of the IEEE International Conference on Image Processing, ICIP '05*, vol. 2, Genoa, Italy, 2005, pp. 69–72.

[28] A. De Rosa, A. Piva, M. Fontani, and M. Iuliani, "Investigating multimedia contents," in *Security Technology (ICCST), 2014 International Carnahan Conference on*, 2014, pp. 1–6.

[29] J. Fridrich, *Steganography in Digital Media: Principles, Algorithms, and Applications*, 2009.

[30] N. Bartlow, N. Kalka, B. Cukic, and A. Ross, "Identifying sensors from fingerprint images," in *Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops 2009. IEEE Computer Society Conference on*, 2009, pp. 78–84.

[31] M. Chen, J. Fridrich, M. Goljan, and J. Lukas, "Determining image origin and integrity using sensor noise," *IEEE Transactions on Information Security and Forensics*, vol. 3, no. 1, pp. 74–90, 2008.

[32] G. Chierchia, G. Poggi, C. Sansone, and L. Verdoliva, "A bayesian-mrf approach for prnu-based image forgery detection," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 4, pp. 554–567, 2014.

[33] M. Goljan, J. Fridrich, and T. 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, 2009.

[34] 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, 2010.

[35] K. Rosenfeld and H. 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, 2009, pp. 72 540M – 725 408M.

[36] E. Alles, Z. Geradts, and C. Veenman, "Source camera identification for heavily jpeg compressed low resolution still images," *Journal of Forensic Sciences*, vol. 54, no. 3, pp. 628–638, 2009.

[37] M. Goljan, J. Fridrich, and M. Chen, "Defending against fingerprint-copy attack in sensor-based camera identification," *IEEE Transactions on Information Security and Forensics*, vol. 6, no. 1, pp. 227–236, 2011.

[38] N. Kalka, N. Bartlow, B. Cukic, and A. Ross, "A preliminary study on identifying sensors from iris images," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2015.

[39] C.-T. Li, "Source camera identification using enhanced sensor pattern noise," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 2, pp. 280–287, 2010.

[40] M. Goljan, J. Fridrich, and M. Chen, "Sensor noise camera identification: Countering counter-forensics," in *Proceedings of SPIE, Media Forensics and Security XII*, San Jose, CA, USA, 2010.

[41] S. El-Naggar and A. Ross, "Which dataset is this iris image from?" in *Information Forensics and Security (WIFS), 2015 IEEE International Workshop on*. IEEE, 2015, pp. 1–6.

[42] J. Fridrich, "Digital image forensic using sensor noise," *IEEE Signal Processing Magazine*, vol. 26, no. 2, 2009.

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

[44] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising with block-matching and 3d filtering," in *Electronic Imaging 2006*. International Society for Optics and Photonics, 2006, pp. 606 414–606 414.

[45] A. Cortiana, V. Conotter, G. Boato, and F. D. Natale, "Performance comparison of denoising filters for source camera identification," in *Media Watermarking, Security, and Forensics XIII*, ser. Proceedings of SPIE, vol. 7880, 2011, p. 788007.

[46] J. Fridrich, "Sensor defects in digital image forensics," in *Digital Image Forensics: There is more to a picture than meets the eye*, 2012, ch. 6, pp. 179–218.

[47] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[48] L. Fei-Fei and P. Perona, "A bayesian hierarchical model for learning natural scene categories," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 2. IEEE, 2005, pp. 524–531 vol. 2.

[49] R. Mehta and K. Egiazarian, *Texture Classification Using Dense Micro-block Difference (DMD)*, ser. Lecture Notes in Computer Science. Springer-Verlag, Berlin, 2015, pp. 643–658.

[50] T. Ojala, M. Pietikainen, and D. Harwood, "Performance evaluation of texture measures with classification based on kullback discrimination of distributions," in *Proceedings of the 12th IAPR International Conference on Pattern Recognition*, vol. 1, 1994, pp. 582 –585 vol.1.

[51] F. Perronnin, J. Sánchez, and T. Mensink, "Improving the fisher kernel for large-scale image classification," in *European conference on computer vision (ECCV10)*. Springer, 2010, pp. 143–156.

[52] M. Cimpoi, S. Maji, I. Kokkinos, S. Mohamed, and A. Vedaldi, "Describing textures in the wild," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'14)*, 2014, pp. 3606–3613.

[53] F. Perronnin and C. Dance, "Fisher kernels on visual vocabularies for image categorization," in *2007 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2007, pp. 1–8.

[54] N. L. of Pattern Recognition. Casia iris v4 database. http://biometrics.idealtest.org/.

[55] P. J. Phillips, K. W. Bowyer, P. J. Flynn, X. Liu, and W. T. Scruggs, "The iris challenge evaluation 2005," in *Biometrics: Theory, Applications and Systems, 2008. BTAS 2008. 2nd IEEE International Conference on*. IEEE, 2008, pp. 1–8.

[56] A. Kumar and A. Passi, "Comparison and combination of iris matchers for reliable personal authentication," *Pattern recognition*, vol. 43, no. 3, pp. 1016–1026, 2010.

[57] Cteo, "Mmu2 iris image database," *Available at: http://pesona.mmu.edu.my/ ccteo/*, 2008.

[58] H. Proenca and L. Alexandre, "UBIRIS: a noisy iris image database," in *Image Analysis and Processing - ICIAP 2005*, ser. Lecture Notes on Computer Science, vol. 3617. Cagliari, Italy: Springer-Verlag, 2005, pp. 970–977.

[59] M. Dobes and L. Machala, "Upol iris image database, 2004," *Available at: http;//www.phoenix.inf.upol.cz/iris*, 2013.

[60] M. Hosseini, B. Araabi, and H. Soltanian-Zadeh, "Pigment melanin: Pattern for iris recognition," *Instrumentation and Measurement, IEEE Transactions on*, vol. 59, no. 4, pp. 792 –804, 2010.

[61] C. Rathgeb, A. Uhl, and P. Wild, *Iris Recognition: From Segmentation to Template Security*, ser. Advances in Information Security, 2013, vol. 59.

[62] J. Kannala and E. Rahtu, "BSIF: binarized statistical image features," in *Proceedings of the 21st International Conference on Pattern Recognition, ICPR 2012, Tsukuba, Japan, November 11-15, 2012*, 2012, pp. 1363–1366.