

A Preliminary Study on Identifying Sensors from Iris Images

Nathan Kalka
Noblis

Nick Bartlow
Noblis

Bojan Cukic
University of North Carolina
Charlotte

Arun Ross*
Michigan State University

Abstract

In this paper we explore the possibility of examining an iris image and identifying the sensor that was used to acquire it. This is accomplished based on a classical pixel non-uniformity (PNU) noise analysis of the iris sensor. For each iris sensor, a noise reference pattern is generated and subsequently correlated with noise residuals extracted from iris images. We conduct experiments using data from seven iris databases, viz., West Virginia University (WVU) non-ideal, WVU off-angle, Iris Challenge Evaluation (ICE) 1.0, CASIAv2-Device1, CASIAv2-Device2, CASIAv3 interval, and CASIAv3 lamp. Results indicate that iris sensor identification using PNU noise is very encouraging, with rank-1 identification rates ranging from 86%-99% for unit level testing (distinguishing sensors from the same vendor) and 81%-96% for the combination of brand (distinguishing sensors from different vendors) and unit level testing. Our analysis also suggests that in many cases, sensor identification can be performed even with a limited number of training images. We also observe that JPEG compression degrades identification performance, specifically at the sensor unit level.

1. Introduction

Biometric systems, which utilize the physical or behavioral characteristics of an individual for person recognition, have proliferated over the past decade. Advances in data acquisition technology and the need to authenticate or detect individuals in diverse applications, ranging from laptop access control to national security, have spurred the growth of the technology [11]. Notwithstanding the many benefits of this technology, maintaining the security and integrity of biometric data poses significant challenges. For instance, the efficacy of a biometric system can be compromised through malicious attacks such as fabrication or alteration of biometric data, which could occur at many points within the system [17]. Similarly, raw biometric images, such as faces, fingerprints or irises, could be created or altered and then re-introduced into the system. This problem

is exacerbated by the fact that there is often no obvious cue, visual or otherwise, that an image has been fabricated or altered. In order to detect and mitigate vulnerabilities related to data fabrication or alteration, *digital hardware fingerprinting* schemes may be applicable.

Digital hardware fingerprinting is the process of identifying the source hardware used to capture an image¹ regardless of the image content. In the field of digital forensics, digital hardware fingerprinting provides the ability to identify or validate the source hardware that captured an image - a process often referred to as *source identification*. This can be done, for example, by gleaning distinguishing characteristics in images due to sensor imperfections [14, 6]. All sensors are subject to small manufacturing imperfections, resulting from inconsistencies during the production process. Such sensor imperfections manifest as noise in the ensuing image - albeit often undetectable to human observers - and can then be detected and characterized by computer vision or image schemes for the purpose of source identification.

Prior work has focused mainly on measuring imperfections from optical technology such as digital cameras to capture images. Geradts et al. [10] proposed a technique based on sensor imperfections which exploited the presence of defective or dead pixels. Mehdi et al. [15] proposes a supervised learning approach that utilizes 34 image features derived from the spatial and wavelet domains including average pixel values, RGB pair correlations, and neighbor distribution center of mass. Bayram et al. [4] proposed a technique that measures interpolation artifacts introduced by the color filter array (CFA). Lukas et al. [14] measured pixel non-uniformity (PNU) noise, which is defined as “noise” resulting from pixels that have different sensitivity to light, often caused by imperfections in the manufacturing process. Bartlow et al. [3, 2] conducted experiments to show that such PNU noise can be extracted from *fingerprint* sensors and can be used to differentiate between different sensor technologies, brands, models, and/or units.

In this paper, our focus is on “fingerprinting” iris biometric sensors. Unlike traditional digital cameras which

¹Here, we restrict our discussion primarily to visual data.

capture imagery in the visible spectrum of light, the iris sensors studied in this paper capture imagery in the near infrared spectrum of light. Given this difference it is uncertain whether the aforementioned hardware fingerprinting techniques are viable in the infrared spectrum. Another difference is related to the number of iris sensors that are currently available, which pales in comparison to sensors utilized in traditional image forensics such as digital cameras. Nevertheless, we study the feasibility of performing source identification utilizing PNU noise across different iris biometric sensors. The contribution of the work is three-fold: (1) To the best of our knowledge, it is one of the first work to demonstrate the capability to automatically identify the hardware source used to collect biometric iris images. To do so, we adopt the technique presented in [14]; (2) We study the impact of varying the number of images used to derive reference templates for sensors; (3) We establish the impact of JPEG compression on the PNU identification technique.

The remainder of the paper is organized as follows: Section 2 outlines the design of the experiment including a description of the datasets and testing methodology. Section 3 defines the algorithm applied for digital hardware fingerprinting. Section 4 presents the results of the identification experiment and analyzes the robustness of the approach to JPEG compression. Section 5 provides a discussion including considerations of interest, while Section 6 summarizes the contribution of the work.

2. Experimental Design

In this work we assembled a dataset composed of image subsets from seven publicly available iris databases. Specifically, we utilized subsets of imagery from ICE1.0, WVU Non-Ideal, WVU Off-Angle, CASIAv2-Device1, CASIAv2-Device2, CASIAv3-Lamp, and CASIAv3-Interval. Sample images and noise residuals (explained in the next section) can be found in Table 2. The following provides a description of each database.

1. **ICE1.0** - This dataset was collected and utilized for the Iris Challenge Evaluation conducted by NIST [16]. Each image was collected with an LG IrisAccess EOU 2200 having a native resolution of 480x640. A total of 2953 images were collected from 244 classes. It is also important to note that this data was intentionally collected with a broad range of quality in mind.
2. **WVU Non-Ideal** - Collected with a hand held OKI Irispass-h device at a native resolution of 480x640 [7]. Like ICE1.0, this dataset was intentionally collected with a range of quality variations (defocus, motion, non-uniform illumination, and occlusion). We utilized a subset of this database comprising 2424 images from 354 classes.
3. **WVU Off-Angle** - An EverFocus CCD camera with a native resolution of 480x640 was utilized for collection [7]. This data is composed of 146 iris classes with each class represented by 4 images captured at a yaw angle of 0° (2 samples), 15°, and 30° degrees.
4. **CASIAv2-Device1** - This subset is composed of 1200 images, collected from 60 classes. Images were captured with a hand held OKI Irispass-h device which has a native resolution of 480x640 [5].
5. **CASIAv2-Device2** - This subset is composed of 1200 images, collected from 60 classes. Images were captured with a proprietary device developed by CASIA having a native resolution of 480x640 [5].
6. **CASIAv3-Lamp** - Collected at the with a hand held OKI Irispass-h device [5] at a native resolution of 480x640. This data is composed of 16212 images, collected from 819 classes.
7. **CASIAv3-Interval** - Collected with a proprietary device developed by CASIA which has a native resolution of 280x320. Composed of 2639 images, collected from 395 classes [5].

From each iris database, we selected 200 images for our experiments. Selection was done such that a single image was chosen from each class, up to a maximum of 200 images. When the total number of classes was less than 200 for the dataset, multiple images were sequentially selected from each class. For instance, three images per class were selected from the WVU Off-angle dataset until the maximum of 200 was reached. The success of the source identification technique was tested while varying the number of images used to generate reference patterns. A breakdown of the train and test scenarios is provided in Table 1. It is

Train/Test Scenarios					
Train	4	8	16	32	64
Test	196	192	184	168	136

Table 1. Train and test scenarios for all sensors.

also important to note that we employed a 10-fold cross-validation framework for all seven sensors/datasets when testing the technique. Therefore, the total number of tests for each sensor ranges from 1960 (196 * 10) to 1360 (136 * 10).

3. Technical Approach

Recall that *brand level* sensor identification attempts to differentiate sensors manufactured by different vendors while *unit level* identification attempts to differentiate between sensors of the same model manufactured by the same















Dataset	Abbreviation	Sensor Model	Manufacturer	Image	Noise	Width	Height	Format
ICE1.0	ICE-LG	IrisAccess EOU 2200	LG			640	480	TIFF
WVU Non-Ideal	WVU-OKI	Irispass-h	OKI			640	480	BMP
WVU Off-Angle	WVU-EverFocus	Monochrome CCD	EverFocus			640	480	BMP
CASIAv2-Device1	CASIAv2-OKI	Irispass-h	OKI			640	480	BMP
CASIAv2-Device2	CASIAv2 _p	n/a	CASIA			640	480	BMP
CASIAv3-Lamp	CASIAv3-OKI	Irispass-h	OKI			640	480	JPEG
CASIAv3-Interval	CASIAv3 _p	n/a	CASIA			320	280	JPEG

Table 2. Illustration of sample images and corresponding noise residuals from each database.

vendor. As a means to identify iris sensors at the brand level, we first adopt the approach proposed by Lukas et al. in [14]. This approach is based on estimating pixel non-uniformity (PNU), a portion of the photo-response non-uniformity (PRNU) inherent to every image captured by the sensors. The remainder of this section is divided into two parts: a description of the general framework for identifying hardware sources through PNU noise and a description of the wavelet-based denoising algorithm [12].

3.1. Sensor Identification

The process of sensor identification consists of two integral steps: (1) Generating a noise reference pattern for each sensor; (2) Correlating noise residuals generated from test images to the aforementioned reference pattern(s).

1. **Generate reference pattern.** For each iris sensor, a reference pattern is calculated by taking an average of the noise residual estimates across multiple training images. This is mathematically described as follows:

$$\aleph_k = p^{(k)} - F(p^{(k)}), \quad (1)$$

$$\aleph_i = \frac{\sum_{k=1}^N \aleph_k}{N}. \quad (2)$$

Here, N represents the number of training images used to generate the reference pattern, \aleph_i (see Figure 1 for examples). Noise residual, \aleph_k , is generated from training image $p^{(k)}$ while F represents a denoising filter. It should be noted that while F can represent any denoising filter, Lukas et al. found that a wavelet-based approach yielded the best results which is described later in this section.

2. **Correlate test residuals to reference pattern(s).** For each input test image, the noise residual, \aleph_k , is ex-

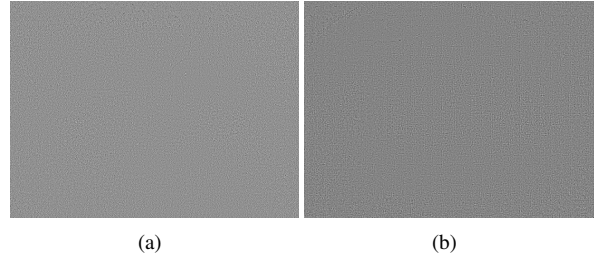


Figure 1. Illustration of reference templates generated from 5 training samples: (a) ICE-LG and (b) WVU-OKI. Visually, structural differences can be perceived between both templates; template differences are important when classifying noise residuals. In general, ideal reference templates for classification will yield low intra-class and high inter-class variation.

tracted and subsequently correlated with each reference pattern \aleph_i . Pearson's product-moment correlation coefficient is adopted for this purpose:

$$\rho_i(\aleph_{(k)}, \aleph_i) = \frac{(\aleph_{(k)} - \bar{\aleph}_{(k)})(\aleph_i - \bar{\aleph}_i)}{\|\aleph_{(k)} - \bar{\aleph}_{(k)}\| \|\aleph_i - \bar{\aleph}_i\|}. \quad (3)$$

Note that $\bar{\aleph}_{(k)}$ and $\bar{\aleph}_i$ represent the sample means of the noise residual and reference pattern, respectively.

3.2. Wavelet Based Denoising Algorithm

The wavelet based denoising algorithm, F , can be summarized in four steps:

1. **Wavelet Decomposition.** The original noisy image is decomposed into four levels utilizing wavelets, specifically, 8-tap Daubechies Quadratic Mirror Filters (QMF). The vertical, horizontal, and diagonal coefficients are denoted as $v(i, j)$, $h(i, j)$, and $d(i, j)$ respectively. Here (i, j) represents the coefficients for each pixel in each of the three sub-bands.

2. **Local MAP variance estimation.** In each sub-band, estimate the local variance of the noise-free image for each wavelet coefficient using MAP estimation for four sizes of a $W \times W$ neighborhood N , where $W \in \{3, 5, 7, 9\}$:

$$\hat{\sigma}_W^2(i, j) = \max \left[0, \frac{1}{W^2} \sum_{(i, j) \in N} h^2(i, j) - \sigma_0^2 \right]. \quad (4)$$

Calculate the minimum of the four local variances:

$$\hat{\sigma}^2(i, j) = \min [\sigma_3^2(i, j), \sigma_5^2(i, j), \sigma_7^2(i, j), \sigma_9^2(i, j)]. \quad (5)$$

3. **Wiener Filtering.** The denoised wavelet coefficients are subsequently obtained after Wiener filtering.

$$h_{den}(i, j) = h(i, j) \frac{\hat{\sigma}^2(i, j)}{\hat{\sigma}^2(i, j) + \sigma_0^2} \quad (6)$$

4. **Repeat.** Steps 1-3 are repeated for each decomposition level and color channel. In [14], the authors used $\sigma_0^2 = 5$ in their experiments as do we in this work. Due to the imaging characteristics of near infrared iris sensors (typically monochrome CCD), the resulting imagery is single channel or grayscale; therefore, it is *not* necessary to perform Step 4 across multiple color channels.

4. Experimental Results

In this section, we present results of sensor identification at the unit level and the brand level. In either case, we wish to determine the sensor that was used to capture the iris image. Therefore, a test noise residual is compared against the reference patterns of each sensor in the dataset under consideration. Results are illustrated in the form of match/non-match histograms, confusion matrices for specific train and test scenarios, and Cumulative Match Characteristic (CMC) curves.

4.1. Unit Level Sensor Identification

The first set of experiments was performed with WVU-OKI, CASIAv2-OKI, and CASIAv3-OKI in the context of unit level identification, since all images were captured with an OKI sensor of the same brand and model. Figure 2 illustrates the difference in correlation between match and non-match comparisons of test noise residuals for the OKI sensor from CASIAv3. Clearly, there is some overlap between the distributions. Specifically, 15 test residuals were misclassified as the CASIAv2-OKI device. Perfect separation was achieved when classifying residuals from the other two OKI devices, which is documented in Table 3. In Figure 3 we plot a CMC curve, which indicates overall accuracy, for all three OKI devices as a function of train/test sizes. When

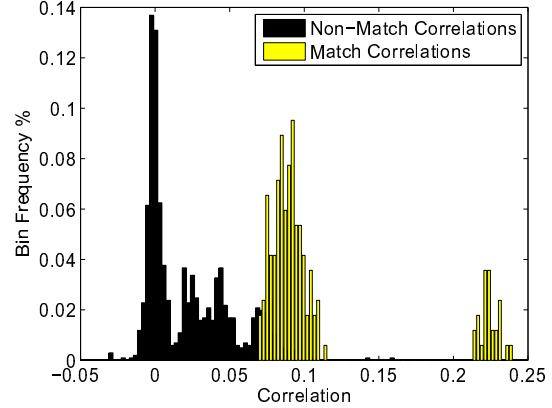


Figure 2. Example match and non-match distributions pertaining to 32 training images for CASIAv3-OKI.

training on only 4 images per sensor, the rank one identification rate is around 86%. This is very encouraging considering the small amount of data utilized for training. This rate increases up to 99.75% when the train size increases to 64 images per sensor. Here the number of test residuals mis-

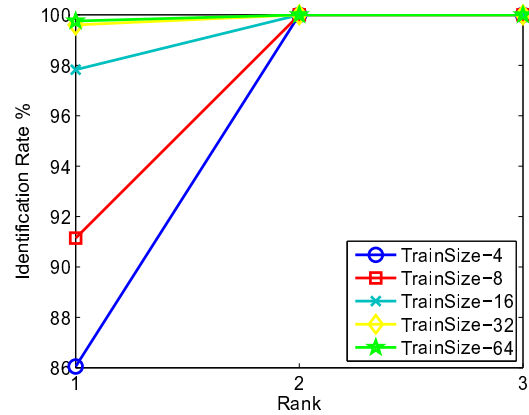


Figure 3. Unit Level Sensor identification as a function of training set size.

classified for the CASIAv3-OKI device reduces to seven.

4.2. Unit and Brand Level Sensor Identification

The following set of experiments explore performance when attempting sensor identification as a combination of the unit and brand levels for all sensors in this study. The ICE-LG, WVU-OKI, CASIAv2-OKI, and CASIAv2_p sensors provided the best performance as perfect separation was achieved (see Table 3) when utilizing 32 training images per reference pattern. An example is provided in Figure 4 which illustrates the separation of match and non-match distributions for the ICE-LG sensor. On

Actual \ Classified	ICE-LG	WVU-OKI	WVU-EverFocus	CASIAv3-OKI	CASIAv3 _p	CASIAv2-OKI	CASIAv2 _p
ICE-LG	1680	0	0	0	0	0	0
WVU-OKI	0	1680	0	0	0	0	0
WVU-EverFocus	9	0	1661	0	0	10	0
CASIAv3-OKI	0	0	0	1665	0	15	0
CASIAv3 _p	103	155	47	210	1009	82	74
CASIAv2-OKI	0	0	0	0	0	1680	0
CASIAv2 _p	0	0	0	0	0	0	1680

Table 3. Confusion matrix when training on 32 images per sensor.

Actual \ Classified	ICE-LG	WVU-OKI	WVU-EverFocus	CASIAv3-OKI	CASIAv3 _p	CASIAv2-OKI	CASIAv2 _p
ICE-LG	1680	0	0	0	0	0	0
WVU-OKI	0	1387	0	293	0	0	0
WVU-EverFocus	15	9	1605	23	19	0	9
CASIAv3-OKI	0	0	0	1664	0	16	0
CASIAv3 _p	102	147	44	234	985	87	81
CASIAv2-OKI	0	0	0	1090	0	590	0
CASIAv2 _p	0	0	0	0	0	0	1680

Table 4. Confusion matrix when training on 32 images per sensor with JPEG compression (Quality=75).

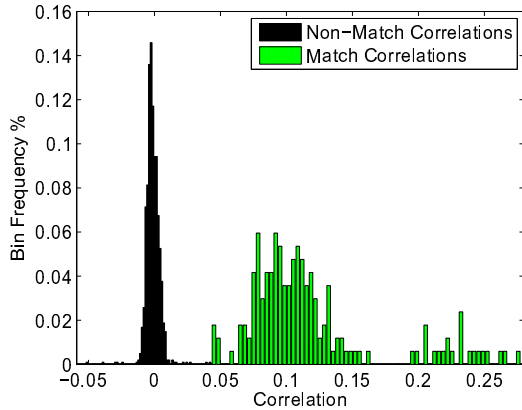


Figure 4. ICE1.0 example match and non-match distributions with 32 training images per sensor.

the other hand, perfect separation was not attainable for the WVU-EverFocus, CASIAv3-OKI, and CASIAv3_p test residuals. Specifically, few errors in classification were observed for the WVU-EverFocus and CASIAv3-OKI test residuals while CASIAv3_p proved the most challenging (see Table 3). Here, misclassification of test residuals is spread almost uniformly across the remaining six sensors. This is also observed when utilizing the largest number of training images to generate a reference pattern. There are a number of plausible explanations for this, not the least of which is JPEG compression (the images were released in JPEG format; further discussion on this issue may be found in

section 5). Overall accuracy in terms of CMC performance across all sensors and train/test scenarios is shown in Figure 5. Here, the rank one identification accuracy approaches

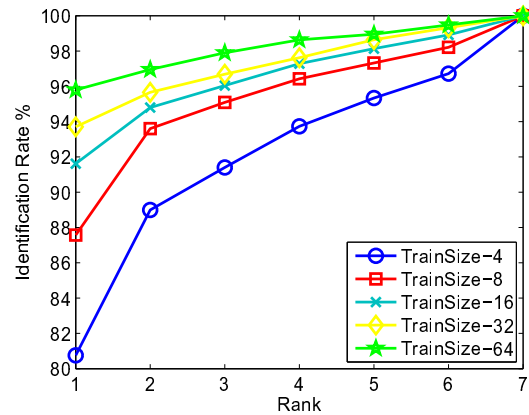


Figure 5. Cumulative match characteristic curve as a function of train/test scenarios when considering all seven sensors.

81% when generating a reference pattern using four training image samples. This number increases to 96% when considering 64 training image samples.

4.3. Impact of JPEG Compression

The last set of experiments analyzes the impact JPEG compression has on the proposed sensor identification technique with respect to CMC performance. More specifically,

this experiment analyzes performance when reference patterns are generated from iris images prior to compression while test residuals are generated from images compressed at JPEG quality levels of 75, 50, and 35. The only exception is for CASIAv3-OKI and CASIAv3_p as the original image data was released in JPEG format. It is also important to note that the ISO/IEC 19794-6 Iris Image Data Standard [1] recommends a JPEG compression ratio of no more than 6 : 1. This ratio is approximately equal to JPEG quality levels between 90 – 95, but is outside the ranges tested in

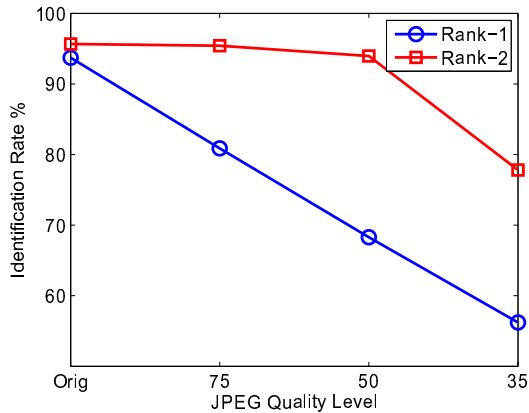


Figure 6. Rank-1 and Rank-2 Identification performance for all seven sensors as a function of JPEG quality level. In this illustration reference patterns were generated from 32 training images.

this paper. However, Daugman [8] has observed that iris recognition accuracy actually improves at JPEG quality 70. Therefore, the aforementioned JPEG quality levels (specifically 75), are reasonable in the current context.

Figure 6 illustrates the impact JPEG compression has on rank-1 and rank-2 identification performance when considering 32 train image samples to generate reference patterns. Most notably, accuracy degrades as the JPEG quality level decreases. This is further noted in the confusion matrix (for JPEG quality level 75) provided in Table 4. Interestingly, ICE-LG and CASIAv2_p test residuals remain unaffected by JPEG compression. On the other hand, heavy degradation was noted during classification of the OKI test residuals, specifically, the WVU-OKI and CASIAv2-OKI residuals.

5. Discussion

When evaluating the experimental results, readers need to be aware of the underpinnings of this study. One consideration of interest relates to the number of iris sensors tested in this study. Unit level testing was limited to three sensors while brand and unit combination testing consisted of seven sensors. If it were possible, increasing the number of iris sensors in either case, in the order of 100’s or even 1000’s, may result in degraded identification performance.

However, currently, iris sensors are not as prolific as other types of digital cameras and are much more expensive. Furthermore, publicly available iris image data sets, pertaining to a large number of iris sensors, do not exist.

Another point of concern is related to correlating residuals of different sizes, as not all iris sensors share the same native resolution. The larger residual is cropped to the smaller for compatibility with the selected method of correlation. Although not tested in this paper, normalized cross correlation may provide a better solution to this problem as in [9, 13]. Alternatively, resizing the image prior to noise extraction may be possible. However, it is unclear if such image operations may introduce new artifacts or even destroy the inherent noise pattern.

As discussed earlier, the PNU noise is a consequence of each pixel in the imaging array having a different sensitivity to visible light. Our results seem to indicate that PNU noise will also accumulate when the light spectrum shifts to near infrared (e.g., iris, particularly at 750nm-850nm).

The last point of interest is related to the classification performance of CASIAv3_p residuals. While JPEG compression may be one of the contributing factors to the observed performance, we believe photometric properties such as image saturation may also be one of the underlying causes. In [14], the authors argue that PNU noise cannot accumulate when pixels are completely saturated (pixel intensity = 255) or under saturated (pixel intensity = 0). After visual evaluation of images captured by the CASIAv3_p device, over-saturation is apparent, varying spatially from image to image (possibly a result of the near infrared ring of LEDs). Naturally, training on more image samples for this specific device may help accommodate spatial variability for over/under saturated pixels.

6. Summary

This paper investigated the feasibility of sensor identification from iris images. We established the prospect of performing sensor identification based on estimating PNU noise inherent in images through a wavelet based denoising algorithm proposed in [12]. We observed that sensor identification for iris sensors can be successfully performed with as little as 4 images per sensor. Our experiments also indicate that JPEG compression had minor impact on brand level testing while strong degradations in performance were observed at the unit level. Future work would involve utilizing iris images pertaining to a larger number of sensors.

References

- [1] Information technology, biometric data interchange format: Iris image data. Standard 19794-6, ISO/IEC, 2005. 6
- [2] N. Bartlow. *Establishing the Digital Chain of Evidence in Biometric Systems*. PhD thesis, West Virginia University, May 2009. 1

- [3] N. Bartlow, N. Kalka, B. Cukic, and A. Ross. Identifying sensors from fingerprint images. In *Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 78–84, June 2009. 1
- [4] S. Bayram, H. Sencar, N. Memon, and I. Avcibas. Source camera identification based on cfa interpolation. In *IEEE International Conference on Image Processing*, volume 3, pages 69–72, September 2005. 1
- [5] CASIA. Casia iris image database. Online, <http://biometrics.idealtest.org/>. 2
- [6] M. Chen, J. Fridrich, M. Goljan, and J. Lukas. Determining image origin and integrity using sensor noise. *IEEE Transactions on Information Forensics and Security*, 3(1):74–90, march 2008. 1
- [7] S. Crihalmeanu, A. Ross, S. Schuckers, and L. Hornak. A protocol for multibiometric data acquisition, storage and dissemination technical report. *West Virginia University, Lane Department of Computer Science and Electrical Engineering*, 2007. 2
- [8] J. Daugman and C. Downing. Effect of severe image compression on iris recognition performance. *IEEE Transactions on Information Forensics and Security*, 3(1):52–61, March 2008. 6
- [9] T. Filler, J. Fridrich, and M. Goljan. Using sensor pattern noise for camera model identification. In *IEEE International Conference on Image Processing*, pages 1296–1299, October 2008. 6
- [10] Z. Geradts, J. Bijhold, M. Kieft, K. Kurosawa, K. Kuroki, and N. Saitoh. Methods for identification of images acquired with digital cameras. *Enabling Technologies for Law Enforcement and Security SPIE*, 4232, February 2001. 1
- [11] A. Jain, L. Hong, and S. Pakanti. Biometric identification. *Communications of ACM.*, 43(2):91–98, February 2000. 1
- [12] M. Kivanc Mihcak, I. Kozintsev, and K. Ramchandran. Spatially adaptive statistical modeling of wavelet image coefficients and its application to denoising. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 6, pages 3253–3256, March 1999. 3, 6
- [13] C.-T. Li. Source camera identification using enhanced sensor pattern noise. *IEEE Transactions on Information Forensics and Security*, 5(2):280–287, June 2010. 6
- [14] J. Lukas, J. Fridrich, and M. Goljan. Digital camera identification from sensor pattern noise. *IEEE Transactions on Information Forensics and Security*, 1(2):205–214, June 2006. 1, 2, 3, 4, 6
- [15] K. Mehdi, H. Sencar, and N. Memon. Blind source camera identification. In *International Conference on Image Processing*, volume 1, pages 709–712, October 2004. 1
- [16] NIST. Iris challenge evaluation. Online, <http://iris.nist.gov/ice/2005/>. 2
- [17] N. Ratha, J. Connell, and R. Bolle. Enhancing security and privacy in biometrics-based authentication systems. *IBM Systems Journal*, 40(3):614–634, 2001. 1