# Digital Imaging Sensor Identification (Further Study)

Mo Chen, Jessica Fridrich[*], and Miroslav Goljan
Department of Electrical and Computer Engineering
SUNY Binghamton, Binghamton, NY 13902-6000

## ABSTRACT

In this paper, we revisit the problem of digital camera sensor identification using photo-response non-uniformity noise (PRNU). Considering the identification task as a joint estimation and detection problem, we use a simplified model for the sensor output and then derive a Maximum Likelihood estimator of the PRNU. The model is also used to design optimal test statistics for detection of PRNU in a specific image. To estimate unknown shaping factors and determine the distribution of the test statistics for the image-camera match, we construct a predictor of the test statistics on small image blocks. This enables us to obtain conservative estimates of false rejection rates for each image under Neyman-Pearson testing. We also point out a few pitfalls in camera identification using PRNU and ways to overcome them by preprocessing the estimated PRNU before identification.

## 1. INTRODUCTION

The problem of establishing the origin of digital media obtained through an imaging sensor is important whenever digital content is presented as silent witness in the court. For example, in a child pornography case, establishing solid evidence that a given image was obtained using a suspect's camera is obviously very important. The identification technology could also be used to link a camcorder confiscated inside a movie theater to other, previously pirated content. In these applications, it is necessary to link an image or a video-clip to a *specific piece of hardware* and not just a camera brand or model [1–3]. To solve this problem, an equivalent of "biometrics for cameras" or fingerprint is needed. One of the first fingerprints for digital imaging sensors were defective pixels (hot and dead pixels) [4]. Lukas et al. [5] proposed the PRNU and its detection inside a given image using correlation in a manner similar to digital watermarking methods. In this paper, we also use the PRNU but construct the estimation and detection methods from a model of the sensor output using the apparatus of statistical signal estimation and detection. This enables us to estimate the detection errors more accurately and make better use of available data. In particular, the number of images needed to estimate the PRNU can be significantly smaller than what was reported in [5]. Moreover, by estimating the error probabilities *for each image separately*, rather than for all images from the camera, more reliable results can be obtained.

In Section 2, we describe a simplified sensor output model that will be used in this paper. Estimation of PRNU and its detection is detailed in Section 3. The predictor of the test statistics is described in Section 4. The Neyman-Pearson testing methodology employed in this paper appears in Section 5. Experimental results for 5 cameras are discussed in Section 6. In Section 7, we point out the importance of preprocessing the PRNU to decrease the false alarm rate (false camera identification). The paper is summarized in Section 8, where we also outline future research directions.

Everywhere in this paper, boldface font will denote vectors of length specified in the text, e.g., $\mathbf{X}$ and $\mathbf{Y}$ are vectors of length $n$ and $\mathbf{X}[i]$ denotes the $i$-th component of $\mathbf{X}$. Unless mentioned otherwise, all operations among vectors, such as product, ratio, raising to a power, etc., are *element-wise*. The norm of $\mathbf{X}$ is denoted as $\|\mathbf{X}\| = \sqrt{\mathbf{X} \odot \mathbf{X}}$ with $\mathbf{X} \odot \mathbf{Y} = \sum_{i=1}^{n} \mathbf{X}[i]\mathbf{Y}[i]$ being the dot product of two vectors. Denoting the mean values with a bar, the normalized correlation is

$$corr(\mathbf{X}, \mathbf{Y}) = \frac{(\mathbf{X} - \bar{\mathbf{X}}) \odot (\mathbf{Y} - \bar{\mathbf{Y}})}{\|\mathbf{X} - \bar{\mathbf{X}}\| \cdot \|\mathbf{Y} - \bar{\mathbf{Y}}\|}.$$

---

[*] fridrich@binghamton.edu; phone +001-607-777-6177; fax +001-607-777-4464.

## 2. IMAGING SENSOR OUTPUT MODEL

The chain of processing that occurs in digital cameras is quite complex and varies greatly for different camera types and manufacturers. It includes signal quantization, white balance, demosaicking (color interpolation), color correction, gamma correction, filtering, and, optionally, JPEG compression. Because the processing details are not always easily available (e.g., hard-wired or proprietary), we decided to use a simplified model [6] that captures the most essential elements of typical in-camera processing. This enables us to develop a low-complexity camera ID procedure applicable to a wide spectrum of cameras. We acknowledge that a more accurate model tailored to a specific camera would likely produce more reliable camera ID results at the cost of increased complexity.

Let $\mathbf{I}[i]$ be the signal in one color channel at pixel $i$, $i = 1, \ldots, n$, generated by the sensor before demosaicking is applied and let $\mathbf{Y}[i]$ be the incident light intensity at pixel $i$. Here, we assume that the pixels are indexed, for example, in a row-wise manner. Dropping the pixel indices for better readability, we use the following model of the sensor output

$$\mathbf{I} = g^{\gamma} \cdot \left[ (\mathbf{1} + \mathbf{K})\mathbf{Y} + \mathbf{\Lambda} + \mathbf{\Theta}_s + \mathbf{\Theta}_r \right]^{\gamma} + \mathbf{\Theta}_q , \tag{1}$$

where $g$ is the color channel gain, $\gamma$ is the gamma correction factor (typically, $\gamma \approx 1/2.2$), $\mathbf{K}$ is a zero-mean multiplicative factor responsible for PRNU (the sensor fingerprint), and $\mathbf{\Lambda}, \mathbf{\Theta}_s, \mathbf{\Theta}_r, \mathbf{\Theta}_q$ stand for the following noise sources: dark current, shot noise, read-out noise, and quantization noise. The gain factor $g$ adjusts the pixel intensity level according to the sensitivity of the pixel in the red, green, and blue spectral bands to obtain the correct white balance. We remind that all operations in (1) are element-wise.

Because the dominant term in the square bracket in (1) is the light intensity $\mathbf{Y}$, we can factor it out. Using the first two terms in Taylor expansion of $(1 + x)^{\gamma} = 1 + \gamma x + O(x^2)$, from (1) we obtain

$$\mathbf{I} = \mathbf{I}^{(0)} + \gamma \mathbf{I}^{(0)} \mathbf{K} + \mathbf{\Theta} , \tag{2}$$

where $\mathbf{I}^{(0)} = (g\mathbf{Y})^{\gamma}$ is the sensor output in the absence of noise; $\mathbf{\Theta}$ is a complex of independent random noise components.

## 3. CAMERA IDENTIFICATION METHODOLOGY

The camera identification methodology is a joint estimation and detection problem. First, we estimate the PRNU $\mathbf{K}$ from a set of $N$ images taken by the camera. Then, using hypothesis testing we detect the presence of the term $\mathbf{I}^{(0)}\mathbf{K}$ in a specific image $\mathbf{I}$ whose origin is in question. In this section, we describe the details of both tasks.

The first step is host signal rejection to improve the SNR between the signal of interest and observed data. We suppress the influence of the noiseless image $\mathbf{I}^{(0)}$ by subtracting from both sides of (2) an estimate $\hat{\mathbf{I}}^{(0)} = F(\mathbf{I})$ of $\mathbf{I}^{(0)}$ obtained using a denoising filter[†] $F$

$$\mathbf{W} = \mathbf{I} - \hat{\mathbf{I}}^{(0)} = \gamma \mathbf{IK} + \mathbf{I}^{(0)} - \hat{\mathbf{I}}^{(0)} + \gamma(\mathbf{I}^{(0)} - \mathbf{I})\mathbf{K} + \mathbf{\Theta} , \text{ or}$$
$$\mathbf{W} = \gamma \mathbf{IK} + \mathbf{\Xi} . \tag{3}$$

The term $\mathbf{\Xi}$ is a combination of $\mathbf{\Theta}$ with the additional distortion introduced by the denoising filter. Working with the noise residual $\mathbf{W}$ significantly improves the SNR for our signal of interest $\mathbf{IK}$ and thus improves the reliability of the camera identification process. However, the denoising filter also shapes the signal we are trying to estimate or detect and it also makes the noise $\mathbf{\Xi}$ highly non-stationary. For example, in textured areas $\mathbf{\Xi}$ is larger and the signal $\mathbf{IK}$ is attenuated by a multiplicative factor that depends on local texture.

---

[†] We use a wavelet based denoising filter $F$ that removes from images Gaussian noise with variance $\sigma_F^2$ (in this paper, we used $\sigma_F^2 = 3$.) More details about this filter can be found in our previous work [5] or in the original publication [7].

### 3.1 ML estimation of PRNU

We now explain the methodology for estimating the PRNU $\mathbf{K}$ from $N$ images $\mathbf{I}_1$, …, $\mathbf{I}_N$ obtained by the same camera. We will assume for simplicity that the images are relatively smooth and non-saturated, such as blue sky images. Because we can make arbitrary test images with the camera, this assumption is reasonable. For such images, the model (3) is approximately accurate. From (3), we have for each $k = 1$, …, $N$

$$\frac{\mathbf{W}_k}{\gamma \mathbf{I}_k} = \mathbf{K} + \frac{\mathbf{\Xi}_k}{\gamma \mathbf{I}_k}, \quad \mathbf{W}_k = \mathbf{I}_k - \hat{\mathbf{I}}_k^{(0)}, \; \hat{\mathbf{I}}_k^{(0)} = F(\mathbf{I}_k) \; . \tag{4}$$

Assuming that for each pixel $i$ the sequence $\mathbf{\Xi}_1[i], \mathbf{\Xi}_2[i], \ldots, \mathbf{\Xi}_N[i]$, is WGN (white Gaussian noise) with variance $\sigma^2$, the log-likelihood of obtaining the measured data $\mathbf{W}_k / \gamma \mathbf{I}_k$ given $\mathbf{K}$ is

$$L(\mathbf{W}_k / \gamma \mathbf{I}_k \mid \mathbf{K}) = -\frac{N}{2} \sum_{k=1}^{N} \log(2\pi\sigma^2 / (\gamma \mathbf{I}_k)^2) - \sum_{k=1}^{N} \frac{(\mathbf{W}_k / (\gamma \mathbf{I}_k) - \mathbf{K})^2}{2\sigma^2 / (\gamma \mathbf{I}_k)^2} \; . \tag{5}$$

Taking partial derivatives of (5) with respect to individual elements of $\mathbf{K}$ and simplifying, we finally obtain the ML estimate for the PRNU

$$\frac{\partial L(\mathbf{W}_k / \gamma \mathbf{I}_k \mid \mathbf{K})}{\partial \mathbf{K}} = \sum_{k=1}^{N} \frac{\mathbf{W}_k / (\gamma \mathbf{I}_k) - \mathbf{K}}{\sigma^2 / (\gamma \mathbf{I}_k)^2} = 0 \; ,$$

$$\gamma \hat{\mathbf{K}} = \frac{\sum_{k=1}^{N} \mathbf{W}_k \mathbf{I}_k}{\sum_{k=1}^{N} (\mathbf{I}_k)^2} \; . \tag{6}$$

We also compute the second partial derivative and obtain the Cramer-Rao Lower Bound (CRLB) on the variance of $\hat{\mathbf{K}}$,

$$\frac{\partial^2 L(\mathbf{W}_k / \gamma \mathbf{I}_k \mid \mathbf{K})}{\partial \mathbf{K}^2} = -\frac{1}{\sigma^2} \sum_{k=1}^{N} (\mathbf{I}_k)^2 \Rightarrow \mathrm{var}(\hat{\mathbf{K}}) \geq \left( -E\left[ \frac{\partial^2 L(\mathbf{K})}{\partial \mathbf{K}^2} \right] \right)^{-1} = \frac{\sigma^2}{\sum_{k=1}^{N} (\mathbf{I}_k)^2} \; . \tag{7}$$

Due to the linearity of the signal model, our ML estimator is actually the minimum variance unbiased (MVU) estimator and the CRLB is its variance. Expression (7) also gives us clues on what images should be used for the best estimation of PRNU. First, we want the luminance $\mathbf{I}_k$ to be as high as possible but not saturated because saturated pixels ($\mathbf{I}_k[i] \approx 255$ for an 8-bit grayscale image) carry no information about PRNU. Second, $\mathrm{var}(\hat{\mathbf{K}})$ is proportional to $\sigma^2$, which is the combination of various noise sources and the term $\mathbf{I}_0 - \hat{\mathbf{I}}_0$ introduced by denoising (3). Thus, the test images should be as smooth as possible. In summary, the best candidates for PRNU estimation are images of bright (but not saturated) uniformly white scenes. In practice, one might take out-of-focus images of cloudy sky after zooming in. Finally, we note that $\mathrm{var}(\hat{\mathbf{K}})$ is approximately inversely proportional to the number of images $N$ used for estimation and thus decreases as $1/N$.

### 3.2 PRNU detection

The model (3) with the assumption that $\mathbf{\Xi}$ is a WGN is an approximately valid representation of reality as long as the image is spatially homogenous with no saturated areas, such as sky images. As explained in the previous section, such images are the best for PRNU estimation. In most real-life images, however, the PRNU noise $\gamma \mathbf{I} \hat{\mathbf{K}}$ is modulated. It is attenuated in areas of the image that were "flattened" by processing, such as JPEG compression. Also, in overexposed regions of the image saturated at the maximum value of the dynamic range, the effect of the PRNU (and any other noise) is not present at all. Additional attenuation is produced by the denoising filter because we may be subtracting a portion of the PRNU noise with the denoised image. As a result, we accept the following model for the camera output

$$\mathbf{W} = \gamma \mathbf{TI\hat{K}} + \boldsymbol{\Xi} \,, \tag{8}$$

where now $\mathbf{T}[i]$ is a multiplicative attenuation factor and $\boldsymbol{\Xi}[i]$ is a sequence of independent Gaussian variables with unequal variances $\boldsymbol{\sigma}_{\boldsymbol{\Xi}}^2[i]$ (colored Gaussian noise). We formulate the problem of detection of PRNU in the noise residual $\mathbf{W} = \mathbf{I} - \hat{\mathbf{I}}_0$ of a given image $\mathbf{I}$ as a binary hypothesis testing

$$\begin{aligned} \mathrm{H}_0\!: \ & \mathbf{W} = \boldsymbol{\Xi} \\ \mathrm{H}_1\!: \ & \mathbf{W} = \mathbf{TX} + \boldsymbol{\Xi}, \end{aligned} \tag{9}$$

where $\mathbf{X} = \gamma \mathbf{I\hat{K}}$ is the non-attenuated PRNU signal. Here, more accurately we should have written $\mathrm{H}_0$: $\mathbf{W} = \gamma \mathbf{TI\hat{K}}' + \boldsymbol{\Xi}$, where $\hat{\mathbf{K}}'$ corresponds to a PRNU from some other camera. However, because the combined noise term $\boldsymbol{\Xi}$ dominates the contribution from the PRNU, we consider the PRNU as a weak signal and include it in the noise term.

To estimate the shaping factor $\mathbf{T}$ and the variance $\boldsymbol{\sigma}_{\boldsymbol{\Xi}}^2$, we can either accept a parametric model for them and estimate the parameters, or we can estimate $\mathbf{T}$ and $\boldsymbol{\sigma}_{\boldsymbol{\Xi}}^2$ locally from the image. Finding a good model, however, is not an easy task because both factors depend on the complex interplay between the denoising filter and the local texture and content of the image. Likewise, due to insufficient data, it is not possible to accurately estimate these two non-stationary factors at every pixel. Instead, we opted for the following approach. We divide the image into $M$ smaller disjoint blocks and assume that within each block $b \in \{1, \ldots, M\}$ $\mathbf{T}$ and $\boldsymbol{\sigma}_{\boldsymbol{\Xi}}^2$ are constant equal to $T_b$ and $\sigma_b$, respectively[‡]. Both $T_b$ and $\sigma_b$ will be estimated from a predictor constructed in the next section. Allowing these estimates to be accurate up to an unknown multiplicative factor common to all blocks (see the discussion at the end of Section 4), we arrive at the following hypothesis testing problem

$$\begin{aligned} \mathrm{H}_0\!: \ & \mathbf{W} = \gamma \boldsymbol{\Xi} \\ \mathrm{H}_1\!: \ & \mathbf{W} = a\mathbf{TX} + \gamma \boldsymbol{\Xi} \,. \end{aligned} \tag{10}$$

where now $\boldsymbol{\Xi}[i]$ is a sequence of zero-mean independent Gaussian variables with known variance equal to $\hat{\sigma}_b^2$ on the $b$-th block, and $\mathbf{T}[i]$ is piece-wise constant equal to $\hat{T}_b$, which is also known. Both $a$ and $\gamma$ are unknown multiplicative factors. The optimal detector for (10) is the normalized Generalized Matched Filter (see Chapter 4.4 in [8])

$$\rho = \frac{\displaystyle\sum_{b=1}^{M} \frac{\hat{T}_b}{\hat{\sigma}_b^2} (\mathbf{X}_b \odot \mathbf{W}_b)}{\sqrt{\displaystyle\sum_{b=1}^{M} \frac{\hat{T}_b^2}{\hat{\sigma}_b^2} \| \mathbf{X}_b \|^2} \sqrt{\displaystyle\sum_{b=1}^{M} \frac{1}{\hat{\sigma}_b^2} \| \mathbf{W}_b \|^2}} \,. \tag{11}$$

We now explain the estimation of the shaping factor $T_b$ and variance $\sigma_b$. Under hypothesis $\mathrm{H}_1$, the noise residual $\mathbf{W}$ comes from the tested camera and the normalized correlation $\rho_b$ calculated for pixels in the $b$-th block is

$$\rho_b = corr(T_b \mathbf{X}_b, \mathbf{W}_b) = corr(\mathbf{X}_b, T_b \mathbf{X}_b + \boldsymbol{\Xi}_b) = \frac{T_b \| \mathbf{X}_b \|^2 + \mathbf{X}_b \odot \boldsymbol{\Xi}_b}{\| \mathbf{X}_b \| \sqrt{T_b^2 \| \mathbf{X}_b \|^2 + 2 T_b \mathbf{X}_b \odot \boldsymbol{\Xi}_b + \| \boldsymbol{\Xi}_b \|^2}} \,. \tag{12}$$

Because $\boldsymbol{\Xi}_b$ is zero mean and uncorrelated with $\mathbf{X}[i]$, the mixed term $\mathbf{X}_b \odot \boldsymbol{\Xi}_b$ is small compared to the other terms in (12). Thus,

$$\rho_b \approx \frac{T_b \| \mathbf{X}_b \|}{\sqrt{T_b^2 \| \mathbf{X}_b \|^2 + \| \boldsymbol{\Xi}_b \|^2}} = \frac{1}{\sqrt{1 + \frac{\| \boldsymbol{\Xi}_b \|^2}{T_b^2 \| \mathbf{X}_b \|^2}}} = \frac{1}{\sqrt{1 + \frac{E_{\boldsymbol{\Xi},b}}{E_{\mathbf{TX},b}}}} \,, \tag{13}$$

---

[‡] The blocks will be described by their index sets $B_b \subset \{1, \ldots, n\}$. Signals constrained to the $b$-th block will be denoted with subscript $b$, e.g., $\mathbf{X}_b$, $\mathbf{W}_b$, etc.

where $E_{\mathbf{TX},b} = \| \mathbf{TX} \|^2 = T_b^2 \| \mathbf{X}_b \|^2$, $E_{\Xi,b} = \| \Xi_b \|^2 = \sigma_b^2 | B_b |$ denote the energy of the signal of interest $\mathbf{TX}$ and the noise in the $b$-th block, respectively ($|B_b|$ is the cardinality of $B_b$). Finally, we obtain from (13) an estimate for the SNR between the signal of interest and noise

$$\frac{E_{\Xi,b}}{E_{\mathbf{TX},b}} = \frac{1}{\rho_b^2} - 1 . \tag{14}$$

Expression (14) assumes knowledge of the correlation $\rho_b$, which we only know under hypothesis H$_1$ but not under H$_0$. We address this problem by constructing a predictor of the normalized correlation $\rho_b$ on small blocks based on our knowledge of $\hat{\mathbf{K}}$, $\hat{\mathbf{I}}_0$, and some features extracted from the block that we expect to most influence $\rho_b$ (see Section 4). In other words, we will calculate for each block an estimate (prediction) $\hat{\rho}_b$ of what the correlation should be if $\mathbf{I}$ was obtained by the camera with PRNU $\hat{\mathbf{K}}$ (hypothesis H$_1$).

From (8), $\| \mathbf{W}_b \|^2 = E_{\mathbf{W},b} = E_{\mathbf{TX},b} + E_{\Xi,b}$ and thus using (14) we obtain estimates for both $T_b$ and $\sigma_b^2$:

$$\hat{\sigma}_b^2 = (1 - \hat{\rho}_b^2) E_{\mathbf{W},b} , \; \hat{T}_b = \frac{\hat{\rho}_b}{\| \mathbf{X}_b \|} \sqrt{E_{\mathbf{W},b}} . \tag{15}$$

In the expression for $\hat{\sigma}_b^2$ in (15), we skipped the multiplicative factor $1/|B_b|$ because it is the same for all blocks if all blocks are of the same size and thus does not influence the value of $\rho$ (11). We note that once the shaping factors $T_b$ and the noise variance $\sigma_b^2$ are estimated, the test statistics for the generalized matched filter (10) is calculated by evaluating (11) after substituting for $\hat{T}_b$ and $\hat{\sigma}_b^2$ their corresponding estimates (15)

$$\rho = \sum_{b=1}^{M} \beta_b \rho_b , \tag{16}$$

where

$$\beta_b = \frac{\dfrac{\hat{T}_b}{\hat{\sigma}_b^2} \| \mathbf{X}_b \| \cdot \| \mathbf{W}_b \|}{\sqrt{\displaystyle\sum_{i=1}^{M} \frac{\hat{T}_i^2}{\hat{\sigma}_i^2} \| \mathbf{X}_i \|^2} \sqrt{\displaystyle\sum_{i=1}^{M} \frac{1}{\hat{\sigma}_i^2} \| \mathbf{W}_i \|^2}} . \tag{17}$$

## 4. CORRELATION PREDICTOR

In this section, we construct a predictor of the correlation $\rho_b = corr(\mathbf{X}_b, \mathbf{W}_b)$ on small blocks for images coming from the same camera as the PRNU (the matched case corresponding to hypothesis H$_1$). From experiments, we determined that the most influential factors are

1)  Image intensity;
2)  Texture;
3)  Signal flattening.

The predictor will be a mapping from some feature vector to a real number in the interval [0,1] – the predicted value of $\rho_b$. The block size can not be too small to avoid the lack of statistically significant data but not too large because then the assumption of stationarity of the shaping factor $\mathbf{T}$ and the variance $\sigma_\Xi^2$ is less likely to hold. For typical sizes of digital camera images with 1 million pixels or more, we recommend square blocks with $|B_b| = 128{\times}128$ pixels.

**Image intensity.** Because the PRNU noise is multiplicative, the correlation is higher in areas of high intensity. However, due to the finite dynamic range, the PRNU noise is not present in saturated regions ($\mathbf{I}[i] = 255$) and is

attenuated for $I_{crit} \leq \mathbf{I}[i] < 255$, where the critical value of intensity $I_{crit}$ depends on the camera. Thus, we define the first feature as the average image intensity attenuated close to the maximum dynamic range

$$f_{\mathbf{I}} = \frac{1}{|B_b|} \sum_{i \in B_b} att(\mathbf{I}[i]) , \tag{18}$$

where $att(x)$ is the attenuation function

$$att(\mathbf{I}[i]) = \begin{cases} e^{-(\mathbf{I}[i]-I_{crit})^2/\tau}, & \mathbf{I}[i] > I_{crit}, \\ \mathbf{I}[i]/I_{crit}, & \mathbf{I}[i] \leq I_{crit}, \end{cases} \tag{19}$$

where $\tau$ is a constant. For example, for our tested Canon G2 camera, we determined from experiments $I_{crit} = 250$, $\tau = 6$.

**Texture.** The correlation tends to be smaller in textured areas for two reasons – the variance $\sigma_{\underline{\underline{\Xi}}}^2$ is larger (and thus $\rho_b$ (13) smaller) and the shaping factor $\mathbf{T} < 1$ (the denoising filter removes part of the signal of interest). Since the denoising filter that we use is constructed in the wavelet domain (8-tap Daubechis), we conveniently use this transformed signal and calculate the texture feature $f_{\mathbf{T}}$ as the average reciprocal power of wavelet coefficients $\mathbf{w}$ in the LH, HL, and HH subbands in the first three levels of the wavelet decomposition of the block $B_b$:

$$f_{\mathbf{T}} = \frac{1}{|S_b|} \sum_{i \in S_b} \frac{1}{1 + \mathrm{var}(\mathbf{w}[i])} , \tag{20}$$

where $S_b$ is the union of indices of all wavelet coefficients from all 9 subbands and $\mathrm{var}(\mathbf{w}[i]) = \min \{\mathrm{var}_3(\mathbf{w}[i]), \mathrm{var}_5(\mathbf{w}[i]), \mathrm{var}_7(\mathbf{w}[i]), \mathrm{var}_9(\mathbf{w}[i])\}$ is the variance of wavelet coefficients in the neighborhood of the coefficient $\mathbf{w}[i]$ estimated from local 3×3, 5×5, 7×7, and 9×9 square neighborhoods (all these quantities are calculated during denoising). The purpose of the reciprocal function is to normalize the feature to the interval [0, 1].

**Signal flattening.** Image processing that is of low-pass filtering nature, such as JPEG compression, further attenuates the PRNU noise and thus decreases the correlation. In a relatively flat and high intensity (but not saturated) area, the predictor would thus incorrectly predict a high correlation. These "flattened" areas will typically have a low value of the local variance. Thus, we added the third feature $f_{\mathbf{S}}$ defined as the ratio of pixels in the block with average local variance below a certain threshold

$$f_{\mathbf{S}} = \frac{1}{|B_b|} \left| \{i \in B_b \mid \sigma_{\mathbf{I}}[i] < c\mathbf{I}[i]\} \right| , \tag{21}$$

where $c$ is an appropriately chosen constant that depends on the variance of the PRNU $\mathbf{K}$ (e.g., $c = 0.03$ for Canon G2) and $\sigma_{\mathbf{I}}^2[i]$ is the local variance of image intensity $\mathbf{I}[i]$ at pixel $i$ estimated from a local 5×5 neighborhood.

From experiments, the correlation coefficients strongly depend on the collective influence of texture and intensity. Sometimes, highly textured regions are also high-intensity regions. Thus, we included the following combined texture-intensity feature calculated in the wavelet domain from the same wavelet coefficients as $f_{\mathbf{T}}$

$$f_{\mathbf{TI}} = \frac{1}{|S_b|} \sum_{i \in S_b} \frac{att(\mathbf{I}[i])}{1 + \mathbf{w}^2[i]} , \tag{22}$$

where in (22) for each $i$, $att(\mathbf{I}[i])$ is the attenuated pixel intensity after subsampling the image to an appropriate size that matches the corresponding subband.

Having specified the features, machine learning can now be used to learn the relationship between the features and the correlation. In this paper, we opted for a simple polynomial multivariate least square fitting. Let $\boldsymbol{\rho}$ be a column vector of $K$ normalized correlations (12) calculated for $K$ image blocks and $\mathbf{f_I}$, $\mathbf{f_T}$, $\mathbf{f_S}$, and $\mathbf{f_{TI}}$ be the corresponding $K$-dimensional feature vectors. We model $\boldsymbol{\rho}$ as a linear combination of the features and their second-order terms

$$\rho[k] = \theta_0 + \theta_1 \mathbf{f_I}[k] + \theta_2 \mathbf{f_T}[k] + \theta_3 \mathbf{f_S}[k] + \theta_4 \mathbf{f_{TI}}[k] + \theta_4 \mathbf{f_I}[k]\mathbf{f_I}[k] + \theta_5 \mathbf{f_I}[k]\mathbf{f_T}[k] + \ldots + \psi[k] , \tag{23}$$

where $\psi[k]$ is the modeling noise and $\theta$ is the vector of coefficients to be determined. In (23), there are total of 1+4+10=15 terms. Rewriting (23) in a matrix form, we have

$$\rho = \mathbf{H}\theta + \psi , \tag{24}$$

where $\mathbf{H}$ is a $K \times 15$ matrix of features and their multiplications and $\theta = (\theta_1, \ldots, \theta_{15})$ is the unknown vector parameter. Applying the least square estimator (LSE) to estimate $\theta$, we obtain

$$\hat{\theta} = \left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\rho \tag{25}$$

and the estimated correlation $\hat{\rho}_b = [f_\mathbf{I}, f_\mathbf{T}, f_\mathbf{S}, f_\mathbf{TI}, f_\mathbf{I}f_\mathbf{I}, f_\mathbf{I}f_\mathbf{T}, \ldots]\hat{\theta}$ .

$$\tag{26}$$



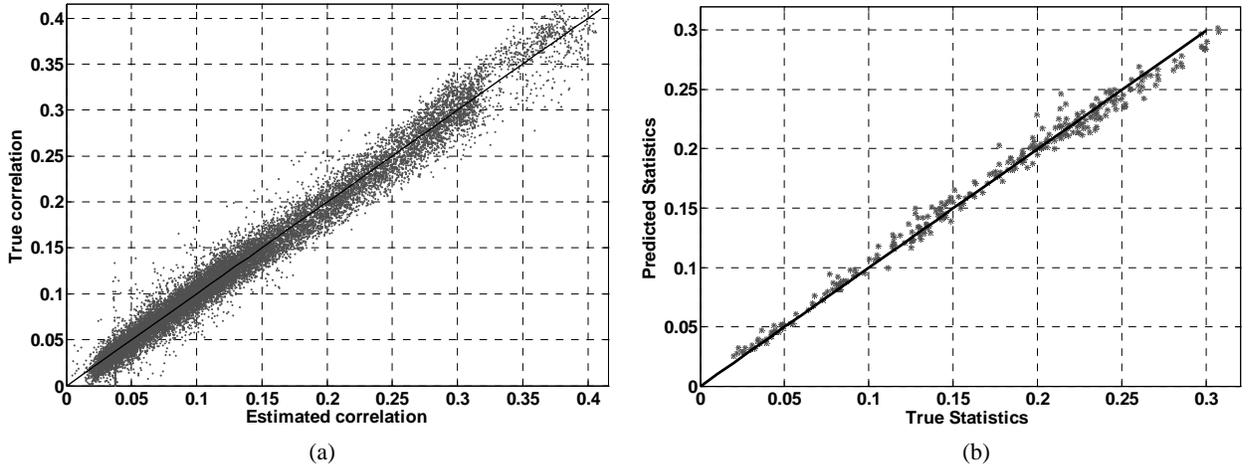(a)                                                        (b)

**Figure 1**. a) Scatter plot of $\rho_b$ vs. $\hat{\rho}_b$ for $K$=30,000 128×128 blocks from 300 images for Canon G2; b) True test statistics $\rho$ (16) vs. $\hat{\rho} = \sum_{b=1}^{M} \beta_b \hat{\rho}_b$ for 285 Canon G2 images not used for calculating the PRNU or the predictor.

The images used to build the predictor should be as diverse as possible so that the features extracted from the blocks cover as large portion of the feature space as possible. This is in contrast to the requirements for calculating the PRNU where we desire to have images mostly without texture. By overlapping the blocks, one can extract several hundreds of blocks from one image, depending on the image size. In practice, we have obtained very good predictors even from as few as 10 images.

We note that if the image under investigation $\mathbf{I}$ is a JPEG image, the predictor should also be trained on JPEG images of approximately the same quality factor because JPEG compression influences the correlation. If the image has undergone an unknown processing that influences the values of correlations, the predictions may be off by a multiplicative factor because the PRNU will be attenuated in such images. This is why we allow in our model (10) unknown multiplicative factors $a$ and $\gamma$.

We remark that the features can be defined in other ways and evaluated in different domains. We tested predictors based on features calculated all from the spatial domain and obtained a very similar performance. Likewise, other machine learning tools that we tested, e.g. neural networks, provided very similar results. It is very likely that a more detailed study of the influence of image properties on the test statistics combined with better machine-learning tools will further improve the predictor and lead to a more accurate camera ID process.

# 5. NEYMAN-PEARSON TESTING

We now describe the Neymann-Pearson hypothesis testing approach to decide whether or not a given image $\mathbf{I}$ was taken with a specific camera whose PRNU $\mathbf{K}$ has been estimated in Section 3. We would like to estimate the probability of making an incorrect decision. Recalling (9), false acceptance occurs when hypothesis $H_0$ is true but we decide $H_1$, while false rejection occurs when we accept $H_0$ when $H_1$ is true.

To estimate the false acceptance ratio (FAR), we tested 2500 images from www.flickr.com coming from over 1000 different cameras against the PRNU $\mathbf{K}$. The test statistics (11) was modeled as a Generalized Gaussian [§] $GG(\mu_0, \alpha_0, \sigma_0)$ where $\mu_0$ is the mean, $\alpha_0$ the shape parameter, and $\sigma_0$ the width parameter. To estimate the false rejection rate (FRR), we need to obtain a pdf for the test statistics for images from the same camera. However, because the statistics heavily depend on the image content, modeling the distribution for very diverse images is nearly impossible as it strongly depends on the available images. Moreover, this would lead to overly conservative estimates for "good" images and too optimistic estimates for highly textured images. We should be evaluating the FRR against images of approximately the same content (close in the feature space). The predictor constructed in the previous section will enable us to achieve this goal.

First, note that the experimental predictor data displayed graphically in Fig. 1 can be used to estimate the pdf of the prediction error $\upsilon_b$ for each estimated block correlation $\hat{\rho}_b$

$$\rho_b = \hat{\rho}_b + \upsilon_b \,. \tag{27}$$

From Figure 1a we see that the variance of $\upsilon_b$ does not depend on the predicted correlation value. Thus, we model the prediction error $\upsilon_b$ as identically distributed (but not necessarily independent) Generalized Gaussian variables with pdf $f_\upsilon(x) = f_{GG}(x; 0, \alpha_\upsilon, \sigma_\upsilon)$. The errors from different blocks, however, are *not independent* because it is likely that neighboring blocks will have similar texture and thus the prediction errors will also be similar. To obtain a conservative estimate of the FRR, we made the assumption that the prediction errors are *completely correlated*, in which case the pdf of the test statistics $\rho = \sum_{b=1}^{M} \beta_b \rho_b = \sum_{b=1}^{M} \hat{\rho}_b + \left(\sum_{b=1}^{M} \beta_b\right)\upsilon$ becomes $f_\upsilon((x - \mu_1)/c)/c$, where $c = \sum_{b=1}^{M} \beta_b$ and $\mu_1 = \sum_{b=1}^{M} \hat{\rho}_b$. This is the pdf of $\rho$ for images of approximately the same content (measured in the feature space).

We set the decision threshold $Th$ to obtain $\text{FAR} = \int_{Th}^{\infty} f_{GG}(x; \mu_0, \alpha_0, \sigma_0)dx = 10^{-5}$, and calculate the FRR for images of approximately the same content as

$$\text{FRR} = \int_{-\infty}^{Th} f_\upsilon((x - \mu_1)/c)/c \, dx \,. \tag{28}$$

We note that we decide that $\mathbf{I}$ was taken by the camera when $\rho > Th$, $\rho$ given by (16).

# 6. EXPERIMENTS

We selected five cameras for our tests: Canon G2 with a 4 megapixel (MP) CCD, Olympus C765-1 and Olympus C765-2 with a 4MP CCD, Sigma SD-9 with a 3MP CMOS Foveon sensor, and Olympus C3030 with a 3MP CCD. We note that the two Olympus C765 cameras are of exactly the same brand. The PRNU was calculated from 30 blue sky images or uniformly lit test images obtained using a light box. The predictors were trained on more than 10,000 blocks from 20 images. We first calculated the test statistics for the $H_0$ hypothesis by testing the PRNU from each camera against 2500 images from 1000 different cameras (including the images from the other four tested cameras). The distribution of the test statistics was used to calculate the threshold $Th$ giving $\text{FAR} = 10^{-5}$. Then, we tested 200–300 images from the correct camera and evaluated the FRR for each image using (28) and (29).

---

[§] $f_{GG}(x; \mu, \alpha, \sigma) = (\alpha / 2\sigma\Gamma(1/\alpha))e^{-\left(\frac{|x-\mu|}{\sigma}\right)^{\alpha}}$

**Table 1.** True and predicted values of test statistics $\rho$ and FRR for eight test images from each of five cameras.

| Canon G2 | RAW | | JPEG 90 | | JPEG 75 | |
|---|---|---|---|---|---|---|
| | True/Pred. | FRR | True/Pred. | FRR | True/Pred. | FRR |
| 1 | 0.026/0.030 | 3.4e–2 | 0.021/0.023 | 5.6e–2 | 0.014/0.016 | 8.7e–2 |
| 2 | 0.066/0.069 | 2.3e–4 | 0.036/0.036 | 1.2e–2 | 0.020/0.020 | 4.5e–2 |
| 3 | 0.107/0.111 | 3.8e–7 | 0.050/0.050 | 1.8e–3 | 0.025/0.025 | 1.9e–2 |
| 4 | 0.137/0.136 | 4.6e–9 | 0.068/0.066 | 1.7e–4 | 0.036/0.036 | 2.5e–2 |
| 5 | 0.134/0.145 | 8.2e–10 | 0.068/0.068 | 1.2e–4 | 0.034/0.034 | 3.5e–3 |
| 6 | 0.139/0.141 | 2.0e–9 | 0.067/0.069 | 1.1e–4 | 0.035/0.036 | 2.5e–3 |
| 7 | 0.250/0.236 | 5.5e–17 | 0.113/0.108 | 2.7e–7 | 0.051/0.049 | 1.7e–4 |
| 8 | 0.281/0.271 | 0 | 0.119/0.110 | 1.9e–7 | 0.048/0.050 | 1.4e–4 |

| Olympus C 765-1 | RAW | | JPEG 90 | | JPEG 75 | |
|---|---|---|---|---|---|---|
| | True/Pred. | FRR | True/Pred. | FRR | True/Pred. | FRR |
| 1 | 0.019/0.020 | 8.6e–2 | 0.015/0.016 | 1.2e–1 | 0.012/0.012 | 1.8e–1 |
| 2 | 0.028/0.027 | 4.3e–2 | 0.022/0.021 | 5.7e–2 | 0.016/0.015 | 1.1e–1 |
| 3 | 0.059/0.052 | 3.7e–3 | 0.039/0.033 | 1.1e–2 | 0.021/0.018 | 7.1e–2 |
| 4 | 0.052/0.053 | 3.6e–3 | 0.029/0.032 | 1.0e–2 | 0.019/0.020 | 4.7e–2 |
| 5 | 0.073/0.069 | 7.1e–4 | 0.046/0.044 | 2.0e–3 | 0.029/0.026 | 1.6e–2 |
| 6 | 0.079/0.075 | 4.1e–4 | 0.061/0.055 | 3.3e–4 | 0.047/0.040 | 1.1e–3 |
| 7 | 0.102/0.098 | 4.5e–5 | 0.076/0.071 | 2.3e–5 | 0.058/0.050 | 1.4e–4 |
| 8 | 0.178/0.180 | 2.6e–8 | 0.124/0.109 | 2.9e–8 | 0.079/0.072 | 8.1e–7 |

| Olympus C 765-2 | RAW | | JPEG 90 | | JPEG 75 | |
|---|---|---|---|---|---|---|
| | True/Pred. | FRR | True/Pred. | FRR | True/Pred. | FRR |
| 1 | 0.015/0.014 | 1.4e–1 | 0.013/0.012 | 1.7e–1 | 0.009/0.009 | 2.5e–1 |
| 2 | 0.043/0.041 | 2.9e–3 | 0.033/0.028 | 1.1e–2 | 0.023/0.020 | 3.6e–2 |
| 3 | 0.040/0.047 | 1.2e–3 | 0.029/0.028 | 1.0e–2 | 0.011/0.018 | 6.0e–2 |
| 4 | 0.062/0.058 | 2.2e–4 | 0.033/0.036 | 2.4e–3 | 0.016/0.023 | 1.9e–2 |
| 5 | 0.074/0.078 | 4.9e–6 | 0.054/0.058 | 1.6e–5 | 0.040/0.040 | 2.6e–4 |
| 6 | 0.080/0.077 | 7.0e–6 | 0.040/0.047 | 2.1e–4 | 0.025/0.030 | 4.0e–3 |
| 7 | 0.092/0.083 | 2.3e–6 | 0.050/0.042 | 6.7e–4 | 0.020/0.023 | 1.9e–2 |
| 8 | 0.130/0.127 | 5.7e–10 | 0.090/0.085 | 1.6e–8 | 0.054/0.050 | 1.6e–5 |

| Sigma SD-9 | RAW | | JPEG 90 | | JPEG 75 | |
|---|---|---|---|---|---|---|
| | True/Pred. | FRR | True/Pred. | FRR | True/Pred. | FRR |
| 1 | 0.047/0.038 | 4.4e–2 | 0.024/0.021 | 1.2e–1 | 0.014/0.011 | 2.4e–1 |
| 2 | 0.068/0.075 | 3.5e–3 | 0.032/0.034 | 4.6e–2 | 0.018/0.020 | 1.1e–1 |
| 3 | 0.088/0.083 | 2.1e–3 | 0.030/0.035 | 4.1e–2 | 0.019/0.024 | 7.9e–2 |
| 4 | 0.100/0.097 | 8.7e–4 | 0.038/0.036 | 4.0e–2 | 0.022/0.021 | 9.7e–2 |
| 5 | 0.133/0.130 | 9.5e–5 | 0.062/0.054 | 1.0e–2 | 0.031/0.032 | 3.9e–2 |
| 6 | 0.143/0.135 | 6.6e–5 | 0.056/0.057 | 8.7e-3 | 0.030/0.035 | 2.9e–2 |
| 7 | 0.170/0.183 | 3.0e–6 | 0.076/0.070 | 3.7e-3 | 0.042/0.035 | 2.9e–2 |
| 8 | 0.176/0.158 | 1.5e–5 | 0.077/0.079 | 2.0e-3 | 0.040/0.035 | 3.1e–2 |

| Olympus C3030 | RAW | | JPEG 90 | | JPEG 75 | |
|---|---|---|---|---|---|---|
| | True/Pred. | FRR | True/Pred. | FRR | True/Pred. | FRR |
| 1 | 0.045/0.058 | 1.5e–2 | 0.044/0.031 | 3.8e–2 | 0.021/0.019 | 5.3e–2 |
| 2 | 0.081/0.088 | 1.2e–3 | 0.058/0.062 | 5.3e–4 | 0.038/0.033 | 2.6e–3 |
| 3 | 0.102/0.113 | 1.2e–4 | 0.061/0.064 | 3.9e–4 | 0.035/0.032 | 3.2e–3 |
| 4 | 0.140/0.131 | 2.0e–5 | 0.076/0.074 | 7.7e–5 | 0.036/0.035 | 1.5e–3 |
| 5 | 0.143/0.146 | 4.4e–6 | 0.077/0.078 | 3.7e–5 | 0.036/0.038 | 7.4e–4 |
| 6 | 0.171/0.156 | 1.7e–6 | 0.094/0.089 | 8.4e–6 | 0.044/0.039 | 4.8e–4 |
| 7 | 0.212/0.190 | 3.5e–8 | 0.103/0.096 | 1.5e–6 | 0.045/0.043 | 1.6e–4 |
| 8 | 0.209/0.206 | 4.4e–9 | 0.104/0.103 | 3.3e–7 | 0.046/0.047 | 3.7e–5 |

In the tables (Table 1) we display the true value of the test statistics $\rho$ (16), its predicted value $\hat{\rho} = \sum_{b=1}^{M} \beta_b \hat{\rho}_b$ , and the FRR (at FAR=$10^{-5}$) for eight selected images. We always selected the worse image in the set (No. 1), the best image (No. 8), and 6 other in-between representative images. Fig. 2 shows the eight images from Cannon G2. We note that the worst images were always heavily textured images, often combined with very bright (saturated) and dark areas. The PRNU in such images is largely attenuated and difficult to detect.
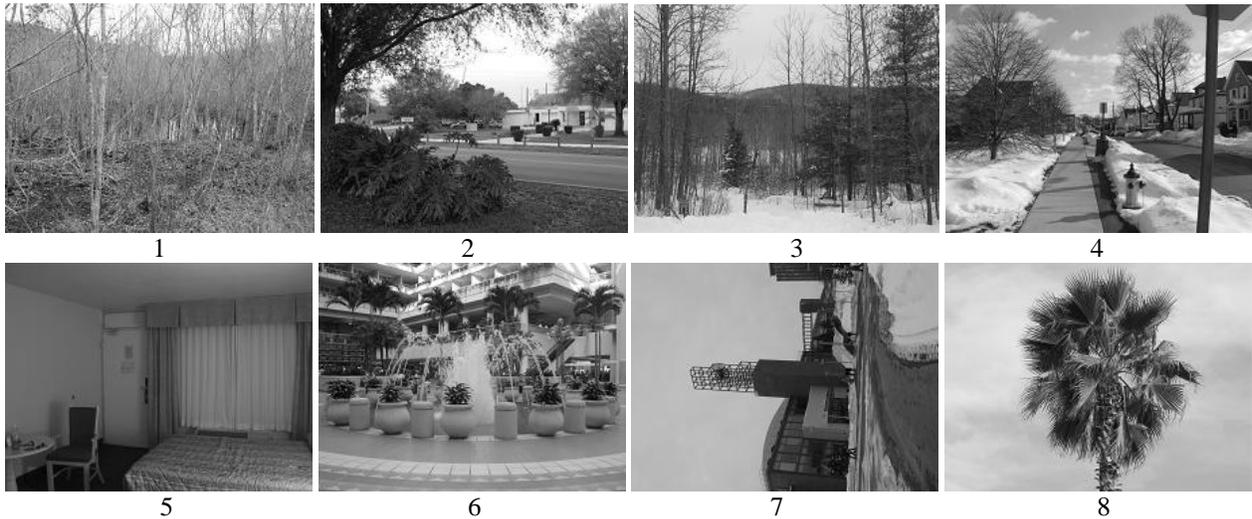


**Figure 2**. Eight Canon G2 images used in experiments.

## 7.    PRNU PRE-PROCESSING

Besides the PRNU, the signal $\hat{\mathbf{K}}$ estimated in Section 3.1 will contain all components that are systematically present in every image, such as artifacts due to color interpolation, on-sensor signal transfer [9], and JPEG compression (blockiness). These artifacts are not unique to the sensor and are shared among cameras of the same brand or cameras sharing the same imaging sensor design[**]. This may cause the estimated PRNUs between two different cameras to be slightly correlated, which increases the false identification rate and decreases the reliability of the camera ID process.

Thus, to further improve the accuracy of the camera ID algorithm, we suppress these artifacts by pre-processing the estimated PRNU before it is used for identification. We measure the success of removing the unwanted artifacts by

---

[**] This observation was already made in [5] but was not further investigated.

calculating the correlation between PRNUs from different cameras (of the same brand or sharing the same sensor design), aiming for the correlation to be as close to zero as possible.

We identified three main causes responsible for the artifacts.

**1. Color interpolation.** Cameras equipped with a CFA (color filter array) only capture one color at each pixel, while the remaining colors must be interpolated from neighboring pixels. Although there exists a large number of interpolation algorithms, all involve some common elements. In particular, after quantization the sensor raw output is first adjusted for gain (because of varying sensitivity of silicone to light of different wavelengths). The resulting colors are then interpolated. The asymmetry of most color interpolation algorithms combined with slightly offset gains might introduce small but measurable biases into the interpolated colors. Because CFAs contain a periodic structure, these biases will show up as a periodic bias in column and row averages of the estimated PRNU. Thus, we can remove such artifacts by zeroing out the means of rows and columns of the estimated PRNU.

**2. Row-wise and column-wise operation of sensors and processing circuits.** The row-wise and column-wise character of operations of digital imaging sensors and/or image processing circuits [9][††] also introduces a bias into each column and row. This linear pattern is also removed by zeroing out columns and rows.

**3. JPEG blockiness artifacts.** Strong JPEG compression, especially in digital camcorders, causes blockiness artifacts that can propagate into the estimated PRNU [10]. These artifacts manifest themselves as peaks and ridges in the magnitude of the PRNU in the Fourier domain.

We note that the first two artifacts are specific to the sensor design, the color filter array (CFA), and color interpolation. Thus they might be potentially useful for identification of the camera brand or model. We again leave this idea to our future work and now focus on methods for removing the artifacts.

To suppress the artifacts, we pre-process the estimated PRNU by zeroing out the means of rows and columns and by further filtering/masking of the PRNU in the Fourier domain. In column and row zeroing, we first subtract column averages from each pixel in each column (for each color channel separately) and then subtract row averages from every pixel in the row. This makes the column and row averages of the estimated PRNU equal to zero. We denote this matrix operation as $ZM(\hat{\mathbf{K}})$. The *linear pattern*, a potentially a useful forensic entity by itself, is the difference $LP(\hat{\mathbf{K}}) = \hat{\mathbf{K}} - ZM(\hat{\mathbf{K}})$. The linear pattern is weak compared to $\hat{\mathbf{K}}$ with SNR below –10 for compact or SLR cameras and it can be stronger for cheap cameras, e.g., cell-phone cameras. Figure 3 shows the enhanced linear pattern estimated from Canon G2 camera. This camera has the Bayer CFA with periodicity 2 along the rows and columns, which clearly shows up in the liner pattern.
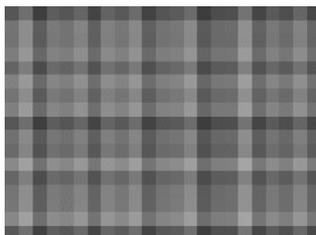


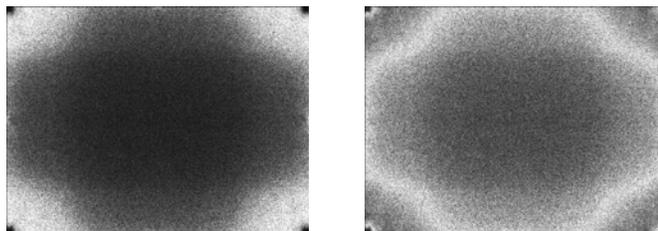**Figure 3.** Linear pattern of Canon G2 (detail).        **Figure 4** Fourier transform of the PRNU after zero-meaning.

To remove the JPEG blockiness and remaining periodic patterns from the zero-meaned PRNU, we apply the Wiener filter in the Fourier domain and only keep the noise component: $WF(ZM(\hat{\mathbf{K}})) = F^{-1}\{F(ZM(\hat{\mathbf{K}})) - W(F(ZM(\hat{\mathbf{K}})))\}$,

---

[††] Gamal *et al.* [9] model the readout process as a first order isotropic autoregressive process.

where W is the Wiener filter with variance determined from the magnitude of the Fourier transform $|F(ZM(\hat{\mathbf{K}}))|$ and window size 3×3. The resulting PRNU has a much flatter frequency spectrum than $ZM(\hat{\mathbf{K}})$ (see Figure 4).

**Table 2.** Correlations between $\hat{\mathbf{K}}_{G2}$ for Canon G2 and $\hat{\mathbf{K}}_{S40}$ for Canon S40 before and after pre-processing.

|  | Red | Green | Blue |
|---|---|---|---|
| $\hat{\mathbf{K}}_{G2}$ vs. $\hat{\mathbf{K}}_{S40}$ | 0.02505 | 0.01341 | 0.01971 |
| $ZM(\hat{\mathbf{K}}_{G2})$ vs. $ZM(\hat{\mathbf{K}}_{S40})$ | 0.01224 | 0.01127 | 0.00768 |
| $WF(ZM(\hat{\mathbf{K}}_{G2}))$ vs. $WF(ZM(\hat{\mathbf{K}}_{S40}))$ | 0.00133 | 0.00076 | 0.00046 |

In Table 2 we show the correlations between differently processed PRNUs estimated from two Canon cameras, G2 and S40. Both cameras share the same CCD sensor. Both PRNU estimates were generated from 20 images (high quality JPEGs in case of G2 and uncompressed images for S40).

Another example shown in Table 3 involves two 1.3 Mp cell phone cameras LG VX8100 and Samsung A900. Only 10 images of sky and gray wall were the source for $\hat{\mathbf{K}}_{LG}$ and 125 similar images for $\hat{\mathbf{K}}_{SA}$. The zero-meaning efficiently reduced the correlation between the PRNU estimates. Further filtering in the Fourier domain removed the blockiness caused by JPEG compression.

**Table 3.** Correlations between $\hat{\mathbf{K}}_{LG}$ for LG VX8100 and $\hat{\mathbf{K}}_{SA}$ Samsung A900 before and after pre-processing.

|  | Red | Green | Blue |
|---|---|---|---|
| $\hat{\mathbf{K}}_{LG}$ vs. $\hat{\mathbf{K}}_{SA}$ | 0.01635 | –0.00580 | 0.03441 |
| $ZM(\hat{\mathbf{K}}_{LG})$ vs. $ZM(\hat{\mathbf{K}}_{SA})$ | –0.00107 | 0.00141 | –0.00199 |
| $WF(ZM(\hat{\mathbf{K}}_{LG}))$ vs. $WF(ZM(\hat{\mathbf{K}}_{SA}))$ | 0.00023 | 0.00100 | –0.00121 |

## 8. CONCLUSIONS

In this paper, we present an improved method for camera ID based on joint estimation and detection of the camera photo-response non-uniformity (PRNU) in images. The method can be used whenever there is a need to answer the question whether or not a given image was taken with a specific camera that is either in our possession or images provably taken by that camera are available.

The method uses the same principles as the approach proposed by Lukas et al. [5] but uses more advanced signal estimation and detection methods. First, the available images are denoised and the PRNU is estimated from the denoising residuals. In contrast to the intuitive approach reported in [5] in which the residuals were averaged, we start with a simplified linearized model of camera output and formulate the problem as parameter estimation in noisy observations. Under the assumption that the corrupting noise sources are Gaussian, the maximum likelihood estimator is MVU. The detection of presence of the PRNU in a given image then amounts to detection of a known signal attenuated by local image properties in independent Gaussian noise with unequal variances. The attenuation factor and the noise variance are estimated from a specially constructed predictor of normalized correlations on small image blocks. The optimal detector for this detection problem is the normalized correlation of pre-whitened signals (generalized matched filter).

Using the Neyman-Pearson hypothesis testing, we estimate the probability of false rejection (falsely deciding that the camera did not take the image when it did) when setting the probability of false identification to $10^{-5}$. The correlation predictor enables us to estimate the probability of false rejection for images of *similar content* (images producing

similar values of the test statistics). This leads to much more accurate error estimates because the reliability of the camera identification method is known to strongly depend on image content.

Finally, we report some new results regarding the PRNU estimation. Besides the PRNU, the estimated signal will contain all the components that are systematic (present in every image), artifacts due to JPEG compression (blockiness), column (or row) artifacts due to on-sensor signal transfer, and artifacts due to color interpolation. These artifacts slightly increase the value of the test statistics among cameras of the same brand or cameras sharing the same imaging sensor design. This increases the false identification rate and decreases the reliability of the camera ID process. To suppress the effect of these artifacts, we proposed to pre-process the estimated PRNU by zeroing out the means of rows and columns and further filtering the PRNU in the Fourier domain. We note that some of these artifacts are specific to the sensor design, the CFA, and color interpolation, and thus might be potentially useful for identification of the camera brand or model.

Our future research will be focused to applying the developed methodology to digital forgery detection.

## ACKNOWLEGEMENT

## REFERENCES

1. M. Kharrazi, H. T. Sencar, and N. Memon, "Blind Source Camera Identification," *Proc. ICIP' 04*, Singapore, October 24–27, 2004.
2. A.C. Propescu and H. Farid, "Statistical Tools for Digital Forensic," in J. Fridrich (ed.): *6th International Workshop on Information Hiding*, LNCS vol. 3200, Springer-Verlag, Berlin-Heidelberg, New York, pp. 128–147, 2004.
3. A. Swaminathan, M. Wu, and K.J.R. Liu, "Non-intrusive Forensic Analysis of Visual Sensors Using Output Images," *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing* (ICASSP'06), May 2006.
4. K. Kurosawa, K. Kuroki, and N. Saitoh, "CCD Fingerprint Method – Identification of a Video Camera from Videotaped Images," *Proc of ICIP'99*, Kobe, Japan, pp. 537–540, October 1999.
5. J. Lukáš, J. Fridrich, and M. Goljan, "Digital Camera Identification from Sensor Pattern Noise," *IEEE Transactions on Information Security and Forensics*, vol. 1(2), pp. 205–214, June 2006.
6. G. Healey and R. Kondepudy, "Radiometric CCD Camera Calibration and Noise Estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16(3), pp. 267–276, March, 1994.
7. M.K. Mihcak, I. Kozintsev, and K. Ramchandran, "Spatially Adaptive Statistical Modeling of Wavelet Image Coefficients and its Application to Denoising," *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Phoenix, Arizona, vol. 6, pp. 3253–3256, March 1999.
8. S.M. Kay, *Fundamentals of Statistical Signal Processing*, Volume II, Detection theory, Prentice Hall, 1998.
9. A. El Gamal, B. Fowler, H. Min, and X. Liu, "Modeling and Estimation of FPN Components in CMOS Image Sensors." *Proceedings of the SPIE*, *Solid State Sensor Arrays: Development and Applications II*, vol. 3301-20, San Jose, CA, pp. 168–177, January 1998.
10. M. Chen, J. Fridrich, and M. Goljan, "Source Digital Camcorder Identification Using CCD Photo Response Nonuniformity," *Proc. SPIE Electronic Imaging*, *Security, Steganography, and Watermarking of Multimedia Contents IX*, vol. 6505, San Jose, California, January 28 – February 1, 2007.