Digital Camera Identification from Images – Estimating False Acceptance Probability

Miroslav Goljan

Dept. of Electrical and Computer Engineering, SUNY Binghamton,
Binghamton, NY 13902-6000, USA
mgoljan@binghamton.edu

Abstract. Photo-response non-uniformity noise present in output signals of CCD and CMOS sensors has been used as fingerprint to uniquely identify the source digital camera that took the image. The same fingerprint can establish a link between images according to their common source. In this paper, we review the state-of-the-art identification method and discuss its practical issues. In the camera identification task, when formulated as a binary hypothesis test, a decision threshold is set on correlation between image noise and modulated fingerprint. The threshold determines the probability of two kinds of possible errors: false acceptance and missed detection. We will focus on estimation of the false acceptance probability that we wish to keep very low. A straightforward approach involves testing a large number of different camera fingerprints against one image or one camera fingerprint against many images from different sources. Such sampling of the correlation probability distribution is time consuming and expensive while extrapolation of the tails of the distribution is still not reliable. A novel approach is based on cross-correlation analysis and peak-to-correlation-energy ratio.

1 Introduction

Digital cameras became affordable commodity for almost everyone. Tens of millions of them have been produced and sold every year. Billions of images are taken and stored in digital form. Along with the scene content, they contain auxiliary data in file headers. But even if the file header is stripped off, the pixels data contain some traces of signal and image processing that can be used in forensics analysis. Image forensics aims to reveal information about the source camera, its brand and model, camera setting, amount of zoom, exposure, time and date, to detect image forgeries and manipulations, reverse-engineer cameras and more. For example, the work of Khanna *et al.* addresses the problem of classification of imaging sensor types [1], [2], Swaminthan *et al.* recognizes color filter arrays and interpolation methods [3], Popescu and Farid introduced a large number of image forensic tools [4] that can reveal forgeries. Forensic analysis of this kind is in an early stage of development but increasing interest of research community speeds up the progress. One of the most reliable methods was proposed by Lukáš *et al.* [9] and further explored by Chen *et al.* [10] and others [11] that is capable of identifying the exact digital camera the image was taken with

(source identification). There are some situations, when such information is a vital piece of evidence in crime investigation. One of them is child pornography where linking photos to the suspect camera can provide a strong evidence for prosecution or steer the investigation. Applications like this require very low probability of wrong accusation. This paper addresses the problem of false camera identification and aims to improve error control while lowering the cost and demand on computations.

One of the challenging aspects is the large number of cameras that have to be uniquely distinguished in an analogy with human fingerprinting. A list of requirements on a camera identifier (*camera fingerprint*) is the following

- high dimensionality (to cover the large number of cameras)
- uniqueness (no two cameras have the same fingerprint)
- stability over time and typical range of physical conditions under which cameras operate
- robustness to common image processing, including brightness, contrast, and gamma adjustment, filtering, format conversions, resampling and JPEG compression
- universality (virtually all digital cameras have it).

On the one hand, this list may not be complete; on the other hand, some requirements may be relaxed if necessary.

In digital image watermarking, an invisible signal (watermark) is inserted in the image to carry some information. This information can be used for owner identification, for an evidence of image authenticity and integrity, for media fingerprinting, or to carry auxiliary data inseparably from the image pixels for other applications. There is a trade-of between watermark robustness and the amount of data it can carry. Image forensics, in contrast to image watermarking, cannot plant fingerprints into existing images. The only option is to explore existing signals that are produced in cameras during image acquisition and on-board signal processing. Fortunately, Photo-Response Non-Uniformity (PRNU) of imaging sensors (CCD, CMOS, and their modern derivatives) is an ideal source of such fingerprinting watermark that is already inherently present in almost all pictures imaging sensors produce. PRNU is caused by material impurities and imperfections in CCD and CMOS manufacturing. Dark content images, such as those taken at night with low light exposure, are not as much affected by PRNU while they may contain dark current [13]. Both these signals resemble noise, and, together, they do exhibit the desirable properties listed above.

Once we know that a fingerprint exists, using it for camera sensor identification (CSI) consists of two tasks. One is *fingerprint estimation*, for which we may have the functional camera or a set of images that were positively taken with that camera. The second is *fingerprint detection* or testing the hypothesis that the camera fingerprint is present in the image under investigation. We shortly reiterate the estimation and detection parts in the next section. We choose the simplest form without improvements found in recent publication of Chen *et al.* [12]. We will discuss the detector properties, the problem of setting detection threshold, and our new approach in Sections 3 and 4. Experiments are presented in Section 5. The summary concludes the paper.

2 Fingerprint Estimation and Detection

Because all images from one camera should contain the same fingerprint, a naïve but a fair representation of the camera fingerprint is a pixel-wise average of a number of images taken with the camera. However, simple averaging is not the best choice. First, such average would contain a significant portion of images' content, so it is better to work with noise residuals. Second, simple averaging does not take into account the fact that PRNU is modulated by the amount of light that falls on the imaging sensor. The maximum likelihood estimate derived by Chen *et al.* [12] for the multiplicative model of PRNU [13] is therefore a better camera fingerprint. In the next paragraph, we adopt necessary notation and concepts from the referenced paper [12].

2.1 Estimation

Let the grayscale (or one color – red, green, or blue) image be represented with an $m \times n$ matrix $\mathbf{I}[i, j]$, $i = 1, \ldots, m$, $j = 1, \ldots, n$, of integers (pixel intensities) ranging from 0 to 255. To eliminate the image content and to increase signal-to-noise ratio for the fingerprint as the signal of interest, the first step is noise extraction. Noise residual $\mathbf{W}$ is obtained by Wiener filtering the image in wavelet domain. We prefer the filter according to [14] for its performance and high speed implementation. Its computational complexity is linear in the number of pixels. (For example, applying the filter to a 4 Mpixel gray scale image on a PC with Pentium 4, 3.4 GHz takes around 2.25 sec, 8 Mpixel image around 4.5 sec.) The only parameter of the filter is variance σ^2 of the stationary noise that is being separated. Denoting the denoised image as $F(\mathbf{I})$, $\mathbf{W} = \mathbf{I} - F(\mathbf{I})$. We start with a multiplicative model of PRNU $\mathbf{K}$,

$$\mathbf{W} = a\mathbf{I}\mathbf{K} + \boldsymbol{\Xi}, \tag{1}$$

where a is a constant, $\boldsymbol{\Xi}$ is a noise term representing all random noises. Throughout this paper, matrices and vectors are in bold font and the operations between them, such as multiplication or ratio of matrices, will always be understood as element-wise. This model is simpler than the one in the reference [12], where an attenuation matrix is considered instead of the constant a. With this complexity reduction in the model, the performance of CSI may slightly drop. What we will gain by that is a much simpler detection part later (no need for a correlation predictor and for more images associated with it) and the system will be more amenable to error analysis. At this point, we omit the property of dark current that behaves the same way as PRNU for fixed image intensities. We will address this issue partially later with an attenuation function.

The maximum likelihood estimation of PRNU $\hat{\mathbf{K}}$ (together with dark current) from a set of N images $\mathbf{I}_1, \ldots, \mathbf{I}_N$ originated from one camera is given by formula (2).

$$\hat{\mathbf{K}} = \frac{1}{S}\sum_{k=1}^{N} \mathbf{W}_k \mathbf{I}_k, \quad S = \sum_{k=1}^{N} (\mathbf{I}_k)^2, \tag{2}$$

up to a multiplicative constant. The noise residuals in the formula are calculated independently as $\mathbf{W}_k = \mathbf{I}_k - F(\mathbf{I}_k)$ for each $k = 1, \ldots, N$. At this point, $\hat{\mathbf{K}}$ becomes our *camera fingerprint.*

Images that are brighter than others contribute more to the sum in (2) because PRNU is more pronounced in them. If we have the camera at hand the best images we can obtain for PRNU estimation are images of high luminance, (obviously) no saturation and uniform content (flat fields). The mean square error of this estimate increases with N.

2.2 Detection

The very basic scenario for the camera identification problem is the following. All we have is an image in a format produced by a digital camera, not further processed in other ways than lossless conversions, and one camera fingerprint obtained by the estimation in Section 2.1. Let $\mathbf{I}$ denote the image matrix that represents pixel values. The question is whether or not the image originated from the camera. At this point, we assume that the question is equivalent to deciding whether or not the image contains the camera fingerprint. This leads to a binary hypothesis test. As we did before, we apply the host signal rejection by noise extraction from the image data $\mathbf{I}$.

Let $\mathbf{W} = \mathbf{I} - F(\mathbf{I})$ be the noise residual of the image. The binary hypothesis test contains noise-only hypothesis H_0 and fingerprint presence hypothesis H_1,

$$\begin{aligned} H_0&: \ \mathbf{W} = \boldsymbol{\Xi}, \\ H_1&: \ \mathbf{W} = \mathbf{I}\hat{\mathbf{K}} + \boldsymbol{\Xi}. \end{aligned} \tag{3}$$

The optimal detector under the assumption that the noise term $\boldsymbol{\Xi}$ is a sequence of i.i.d. random variables with unknown variance is the normalized correlation

$$\rho = corr\left(\mathbf{I}\hat{\mathbf{K}}, \mathbf{W}\right). \tag{4}$$

The decision is obtained by comparing ρ to a decision threshold ρ_{th}.

In Neyman-Pearson hypothesis approach, the decision threshold is set so that the false acceptance probability will not exceed a certain level α. False acceptance (FA) occurs when hypothesis H_0 is true but we decide H_1, while false rejection (FR) occurs when we accept H_0 when H_1 is true. In our scenario, FA occurs when the camera fingerprint is declared to be present in the image while it is not, FR occurs when the presence of the fingerprint is missed. To satisfy the desired level α for FA probability, P_{FA}, we need to estimate the pdf f_0 of the test statistics (4) under H_0, $f_0(x) = \Pr(x|H_0)$. This typically requires evaluation of (4) for a large amount of images coming from other cameras than the fingerprint $\mathbf{K}$. Relations between probability of false acceptance P_{FA} or false rejection P_{FR} and the threshold ρ_0 are given by equations (5).

$$P_{\text{FA}} = \int_{x>\rho_0} f_0(x)dx, \ P_{\text{FR}} = \int_{x\le\rho_0} f_1(x)dx. \tag{5}$$

3 System Improvements and Generalization – Previous Art

The first publication on PRNU based camera identification by Lukáš *et al.* in 2005 [15] spurred a lot of research and publications [16], [12], [17]. Some adjusted the method to address other devices. Its applicability to scanners was tested by Khanna *et al.* [18], Sankur *et al.* studied cell phone cameras that produce highly compressed JPEG images [19], [20] is devoted to camcorders. Other work deals with various image processing generalizing the detection part of CSI for cropped and scaled images, including digitally zoomed ones [11], and for printed images involving very small rotation and nonlinear distortion [21]. PRNU often survives image processing, such as JPEG compression, noise adding, filtering, or gamma correction, unless the PSNR is too low. Surprisingly, no improvement has been done at the noise extraction stage, i.e., the SNR between the fingerprint and the noise residual **W** seems to be hard to improve. This is namely true when computational complexity is an issue. Other characteristics than PRNU have also been explored, dust specs on the sensor protective glass of SLR cameras [22], optical deficiencies of lenses [23], or CFA artifacts [24], [3].

The demand for minimizing error rates has been the motivation for some previous work. Chen *et al.* [10] introduced a correlation predictor and modified the fingerprint detection in terms of prediction error. By modeling the prediction error and consequently pdf $f_1(x) = \Pr(x|\mathrm{H}_1)$ as identically distributed Gaussian variables, this construction allows for satisfying estimation of the FR probability. Only this innovative modification takes into account the prior knowledge of image **I** and corresponds to a slightly different scenario in which one image is fixed and the camera fingerprints are what is being randomly chosen.

As noted in Section 2, to evaluate false acceptance probability as a function of the threshold on the normalized correlation (4) one may need a large amount L of "randomly chosen" images and calculate ρ for each of them and every time a new fingerprint is taken into the hypothesis test. The FA probability is then estimated by the false alarm rate, which is the function of the decision threshold x,

$$FAR(x) = \frac{1}{L}\sum_{i,\rho_i \geq x}^{L} 1 . \tag{6}$$

The problem with such sampling becomes apparent when we are after a very small probability α <<1. The amount of images needed is of the order of $L \approx 1/\alpha$. The smaller the number L, the less reliable the FA probability estimate is. Early attempts to model this probability with Generalized Gaussian pdf had limited success [9]. The biggest obstacle was instability of the shape of samples (6), a bias and skewness in their distribution. More light into this problem was shed in [10] in the section about preprocessing $\hat{\mathbf{K}}$. The noise term in (1) almost always contains periodic signals and structured noise responsible for small positive correlation between noise residuals of images from different cameras. After realizing the reason for this effect, the importance of separating such unwanted signals from the PRNU estimate became eminent.

Filler *et al.* proposed in his recent study [25] to characterize $\hat{\mathbf{K}}$ from (2) with a set of features and utilize them in a different forensic application – identification of camera brands and camera models. Once we admit that image noise residuals systematically contain a signal that is the same or similar in all images from more cameras the hypothesis test changes its character. We will come back to this in the next section.

One significant structured noise with periodicities is called *linear pattern* of the camera fingerprint and is defined through *zero-mean operation* on $\hat{\mathbf{K}}$. Elements of $ZM(\hat{\mathbf{K}})$ for all i=1,.. , m and j=1,.. , n are

$$ZM(\hat{\mathbf{K}})[i,j] = \hat{\mathbf{K}}[i,j] - \frac{1}{m}\sum_{i=1}^{m}\hat{\mathbf{K}}[i,j] - \frac{1}{n}\sum_{j=1}^{n}\hat{\mathbf{K}}[i,j] + \frac{1}{mn}\sum_{i=1,j=1}^{m,n}\hat{\mathbf{K}}[i,j], \tag{6}$$

which makes the mean of every column and every row of the matrix equal to zero. The linear pattern is

$$LP(\hat{\mathbf{K}}) = \hat{\mathbf{K}} - ZM(\hat{\mathbf{K}}). \tag{7}$$

It is the color interpolation (demosaicing) in cameras equipped with color filter arrays (CFA) and row-wise and column-wise operations in signal processing of the imaging sensor what is responsible for *linear pattern*. A slight rounding error due to limited bit-depth in missing colors computations is all it takes to cause this phenomenon. Typically, the energy of the linear pattern $LP(\hat{\mathbf{K}})$ is about an order smaller than the energy of $ZM(\hat{\mathbf{K}})$. However, it does influence pdf f_0 and FA probability markedly as we will also see in the next sections. Suddenly, pdf f_0 behaves nicely, Gaussian model fits very well. We can estimate the FA probability by fitting the model through a smaller number of sampled data (6) and computing the right tail probability of the Gaussian pdf known as the Q-function.

Another preprocessing step proposed earlier is Wiener filtering in the Fourier domain in order to suppress any high magnitudes in the spectrum. The exact type of images, amount of lossy compression, or type of cameras for which this filtering really improves CSI, is yet to be determined. The importance of removing the linear pattern from camera fingerprints have been emphasized in publications ([10],[12]). Despite of that, it is still being omitted in some papers ([26][Bayram]) resulting in poor performance of camera identification for some cameras and image formats or their processing.

From now on, $ZM(\hat{\mathbf{K}})$ is what we call the *camera fingerprint*.

4 Normalized Correlation Abandoned

We want to choose the FA probability for our tests, determine the threshold ρ_0 once, and evaluate the decision possibly many times. There is no problem with this plan

unless the camera fingerprint estimate changes. Such change will require re-evaluation of the threshold, which can be time consuming or even infeasible. This problem is not new. In the work on CIS for cropped and scaled images [11], we handled similar problem. Having a fixed FA probability, it was not guaranteed that the threshold for resized images had to stay the same. The need for a more stable relation between FA probability and the decision threshold led to the introduction of Peak to Correlation Energy (PCE) ratio as a replacement for normalized correlation detector. We demonstrate that PCE is much more suitable detection statistic, even for the basic problem of camera identification, than the normalized correlation. This may be surprising when the correlation was derived as the optimal detector. But assumptions on the model, which our hypothesis test (3) is based on, may not be satisfied. The assumption of independence of Gaussian variables (the noise term $\boldsymbol{\Xi}$) is one culprit. Properties of PCE are especially useful when a periodic signal common to images from various cameras (like the linear pattern) enter the image noise residuals $\mathbf{W}$ as well as $\mathbf{W}_1, \ldots, \mathbf{W}_k$.

It can be shown that the expected sample variance of the normalized correlation between two identically distributed independent random signals of length k is inversely proportional to k (for large k). It is also now true for correlation between noise residuals and camera fingerprint (multiplied by $\mathbf{I}$) under hypothesis H_0 since they are close to being perfectly independent. Thus, a change of the fingerprint size from k_1 pixels to k_2 pixels will cause an expected change of the threshold from ρ_{th} to $\rho_{\text{th}}\sqrt{k_1/k_2}$. We could therefore normalize ρ by the same factor and keep the threshold unchanged if it is just the camera resolution what changes. Our advocacy for PCE comes from elsewhere. We show that the introduction of a periodic signal (like the linear pattern) in both the image noise residue and the camera fingerprint increases correlation ρ, possibly triggering a FA (if ρ_{th} is not adjusted), while PCE drops in such situation (affecting the threshold for PCE very little).

First, we introduce notation and definitions on one-dimensional vectors of real numbers $\mathbb{R}^n$. Their later generalization to two-dimensional matrices is straightforward: one more index is added and a sum over one index is replaced with a sum over both indices. The following definitions apply to centered vectors. Vector $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ is centered if the sum of all elements is zero. If a vector is not centered, its sample mean must be subtracted from all its elements before applying these definitions. Such approach makes formulas simpler compared to general definitions that do not assume before-hand centralization.

Let $\boldsymbol{a}$ and $\boldsymbol{b}$ be two centered vectors in $\mathbb{R}^n$, and operation $\oplus$ is modulo n addition in $\mathbf{Z}_n$. The *circular cross-correlation* is defined as

$$c(k) = \frac{1}{n}\sum_{i=1}^{n} a_i b_{i\oplus k}, \quad k = 0, \ldots, n-1. \tag{8}$$

Normalized circular cross-correlation between $\boldsymbol{a}$ and $\boldsymbol{b}$ is

$$C(k) = \frac{\sum_{i=1}^{n} a_i b_{i\oplus k}}{\sqrt{\sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2}}, \quad k = 0,...,n-1\,. \tag{9}$$

It is the scalar product of **a** and **b** circularly shifted by offset *k* divided by the norms of **a** and **b**.

Peak to Correlation Energy (PCE) ratio is the squared correlation divided by sample variance of the circular cross-correlations,

$$PCE_0(\mathbf{a},\mathbf{b}) = \frac{c^2(0)}{\frac{1}{n-|\mathcal{A}|}\sum_{k,k\notin\mathcal{A}} c^2(k)}\,. \tag{10}$$

where $\mathcal{A}$ is a small square area around zero where a peak correlation is expected for correlated vectors and $|\mathcal{A}|$ is its cardinality. Index 0 at *PCE* is meant to distinguish it from definition in reference [11] where it includes a search for signal shift (or cropping). In that paper, PCE has the maximum over all circular cross-correlations in its numerator. Here, we basically follow the definition according to Kumar and Hassebrook [27]. We point out that PCE does not change if correlation *c* is replaced with normalized correlation *C*. The denominator from (9) cancels out when substituted into (10).

From the Central Limit Theorem, the cross-correlation values for independent vectors follow the Gaussian distribution. We demonstrate in Figure 1 that cross-correlations between $\mathbf{I}\hat{\mathbf{K}}$ and **W** for $k\notin\mathcal{A}$ are also well approximated using the Gaussian distribution. Connecting PCE with P_{FA} then needs the following assumption.

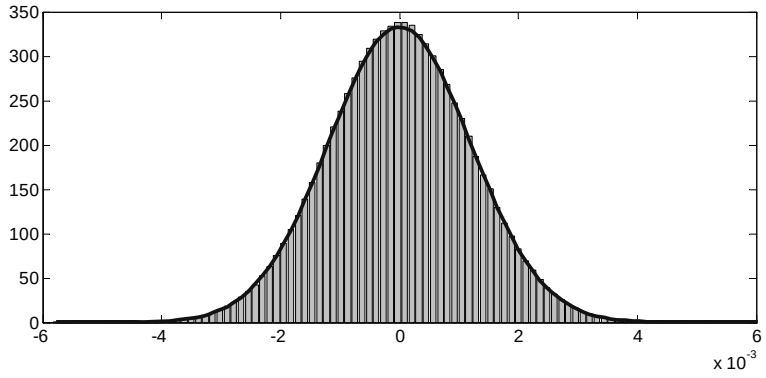

Figure 1. Gaussian fit of cross-correlations for a 1-Mpixel image.

Assumption: The mean of squared normalized correlation between image noise residual and other than correct camera fingerprints (hypothesis H_0) can be estimated as the

mean squared correlation of image noise residual and the correct but shifted fingerprint modulated by the image intensities.

Although this assumption may not be perfectly satisfied, it offers a good insight into experimental evaluation of CSI. All such correlations appear in the denominator of PCE_0 definition (10). Our new detection statistic based on PCE becomes

$$\varphi = PCE_0\left(ZM\left(\mathbf{I}\hat{\mathbf{K}}\right), \mathbf{W}\right). \tag{11}$$

Assuming the pdf in Figure 1 is zero mean, the new detection threshold φ_{th} has the following analytical relationship to the probability of false alarm,

$$\varphi_{\text{th}} = \left[Q^{-1}\left(P_{FA}\right)\right]^2, \tag{12}$$

where Q is the complementary cumulative density function of a normal random variable $N(0,1)$, i.e., Q^{-1} is a scaled inverse error function. Notice that this threshold does not depend on signals length (number of pixels), while ρ_{th} does. Computation complexity is not an issue. The cross-correlation (8) is implemented via Fast Fourier Transform and the normalization as it is in the normalized correlation, or in (9), is not needed. Another saving comes when removing the linear pattern from the signals. Zero-mean operation needs to be applied just on one of the two signals because the following holds,

$$PCE_0\left(ZM\left(\mathbf{I}\hat{\mathbf{K}}\right), ZM\left(\mathbf{W}\right)\right) = PCE_0\left(ZM\left(\mathbf{I}\hat{\mathbf{K}}\right), \mathbf{W}\right). \tag{13}$$

The proof is easy, all that has to be show is that all $c(k)$ in (10) are equal on both sides of (13). The same equation does not hold for the normalized correlation due to changing vector norms, i.e. $corr(ZM(\mathbf{a}), ZM(\mathbf{b})) \neq corr(ZM(\mathbf{a}), \mathbf{b})$. We thus explained why zero-mean preprocessing is not applied to $\mathbf{W}$ in (11).

One of the most important properties of PCE in our Camera ID application is its response to the presence of weak periodic signals. Such signals appear in one or another form in all images. They are artifacts of signal readout and image processing increasing correlation between image noises from one camera brand or model or when the same or similar imaging sensors are built in two cameras.

Let $\mathbf{a}$, $\mathbf{b}$, $\mathbf{z}$ be i.i.d. realizations of Gaussian random variables, $\mathbf{a}, \mathbf{b}, \mathbf{z} \in \mathbb{R}^n$. Then $PCE_0(\mathbf{a}+\mathbf{z}, \mathbf{b}+\mathbf{z}) > PCE_0(\mathbf{a}, \mathbf{b})$. The same inequality is true for the normalized correlation. If $\mathbf{z}$ is a periodic signal (we rename it as $\mathbf{s}$) the situation is different with PCE. Let $l = m/n$ be an integer, $\mathbf{s} = (r_1, r_2, \ldots, r_m)^l \in \mathbb{R}^n$, $m>1$. The correlation is not affected by the fact that $\mathbf{s}$ is periodic. On the other hand, PCE drops whenever the period m is not too large, and when the circular cross-correlation $c(k)$ peaks for more values of k. Then more likely, $PCE_0(\mathbf{a}+\mathbf{s}, \mathbf{b}+\mathbf{s}) < PCE_0(\mathbf{a}, \mathbf{b})$. PCE decreases with decreasing m and the drop is larger for larger variance of $\mathbf{s}$. This is a desirable property because signals with small periods (such as below 100) cannot be as unique as those with no periodicity and thus should not be part of camera fingerprints. If any periodicities are present, we wish they do not trigger positive identification.

In an ideal case, we may be able to estimate P_{FA} for one single hypothesis test by inverting (12),

$$P_{FA} \approx Q\left(\sqrt{\varphi}\right). \tag{14}$$

However, a presence of some unknown weak signals that may be hidden in fingerprints of different cameras (still causing false alarms) would cause an error in the estimate (14). Large experimental tests reveal that (14) is a good practical estimate if we adjust it conservatively by a correction factor. But a cleaner solution lies in further de-correlation of all camera fingerprints – our future research topic.

5 Experiments

We have run several experiments to support our theoretical results. The first one compares behavior of the normalized correlation and the PCE detection statistics when evaluated for every $i = 1, \ldots, N$ during estimation of camera fingerprint from i images of a "flat" scene. We ran the experiments twice, first with zero-mean (6) preprocessing and then without it and repeated them for two cameras, Canon G2, never compressed 4-Mpixel images of a twilight sky, $N=100$, and Fuji E550, JPEG compressed 6-Mpixel images of a blue sky, $N=80$.

In the plots, Figures 2-5 (left), we see that without zero-mean preprocessing the normalized correlation is slightly larger and is always steadily increasing with N. At the same time Figures 2-5 (right) show that PCE is much smaller when zero-mean preprocessing did not remove the linear pattern and it is not increasing for N larger than some N_0, $N_0 \approx 30$ for never compressed images and $N_0 \approx 1$ for JPEGs.

After scaling down the y-axis 10× in Figure 5 (right) we see (Figure 6) that PCE does not grow with increasing N. This means that we would not have a better identification in terms of FAR with increasing number of images we are estimating the fingerprint from at all (!). This is a very surprising observation that we can now explain.

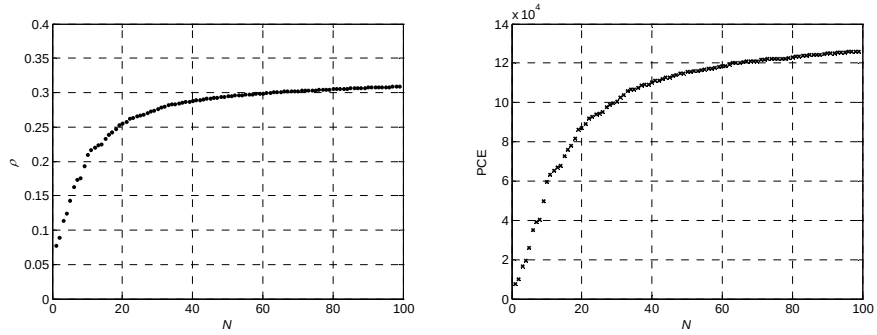

Figure 2. Correlations ρ (left) and PCE φ (right) of one image with a Canon G2 camera fingerprint estimated from N uncompressed images of cloudless sky.

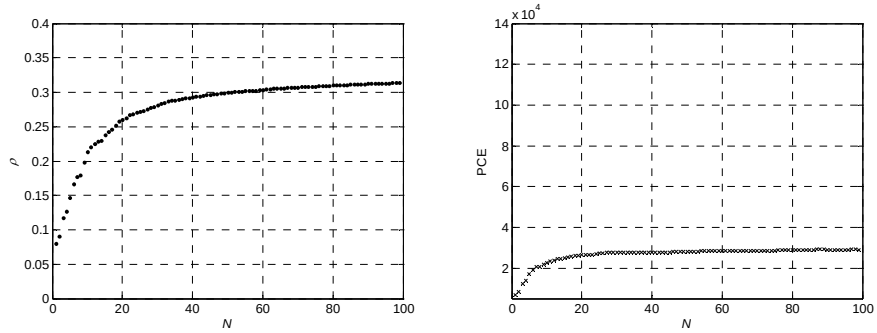

Figure 3. Without zero-mean preprocessing. Correlations ρ (left) and PCE φ (right) of one image with a Canon G2 camera fingerprint estimated from N uncompressed images of cloudless sky.

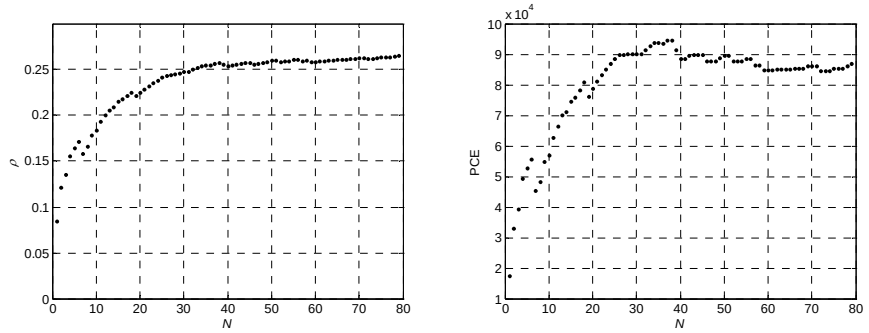

Figure 4. Correlations ρ (left) and PCE φ (right) of one image with a Fuji E550 camera fingerprint estimated from N JPEG compressed images of blue sky.

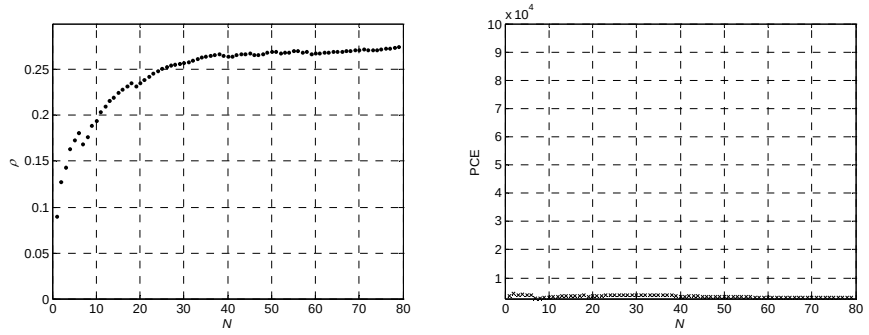

Figure 5. Without zero-mean preprocessing. Correlations ρ (left) and PCE φ (right) of one image with a Fuji E550 camera fingerprint estimated from N JPEG compressed images of blue sky.

The estimate $\hat{\mathbf{K}}$ from (2) contains at least 3 kinds of signals that do not average out with increasing N (while random noise does). It is the PRNU **K** coming from the sensor, true linear pattern **L** from demosaicing, and averaged JPEG compression artifacts **J**. The last one is likely enhanced by the same gradient in sky shots; the im-

ages are brightest in their upper-right corner and darkest in the lower-left corner. As N increases, all three signals gain SNR. When testing the hypothesis (3), we have increasing detection of **K** (good) but also increasing detection of **L** (bad) and **J** (bad). Moreover, PRNU follow Gaussian-like pdf while **L** (and similarly **J**) is limited within a small range of rounding errors and is closer to uniformly distributed random variable. The gain in SNR for **L** and **J** becomes much higher once the mean square error (MSE) of the estimates falls below a certain level. At the same time, the estimate of **K** is improving at the same pace. This is how we explain the deterioration of the camera fingerprint as N exceeded 38. We have to understand that images with different content, as well as with different compression factors, may contain less similar artifacts **J** and the deterioration would not show up. This is what we see in Figure 3. Zero-mean preprocessing is very effective but other filtering is necessary to better remove JPEG compression artifacts **J**.

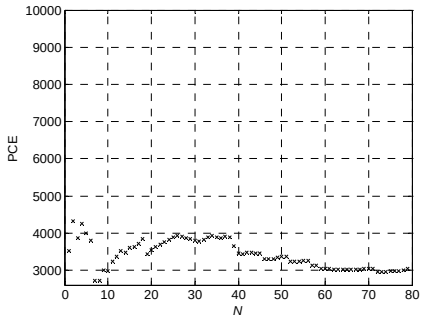

Figure 6. Scaled plot from Figure 5.

These experimental examples show three things.

a) The normalized correlation does not tell much about camera identification performance in terms of detection errors. The threshold for correlation has to be adjusted every time the camera fingerprint is changed.

b) PCE is likely behaving in relation with FA probability. The threshold for PCE can stay fixed even though the camera fingerprint changes.

c) Processing of the camera fingerprint with the zero-mean operation (6) is highly important. Further filtering when JPEG compression is present is desirable. Wiener filtering in the Fourier domain as proposed earlier by Chen *et al*. may be the right answer.

The second experiment was to verify the relation (14). We employed a large scale test with 100,050 images downloaded from the Flickr image sharing web site. Images were in their native resolution, camera fingerprints estimated from randomly chosen 50 (only) images. The signals went through RGB-to-luminance conversion before correlating. Beside zero-mean preprocessing and filtering the camera fingerprints in Fourier domain, we included two correction steps: intensity attenuation function in the term $\mathbf{I}\hat{\mathbf{K}}$ replacing it with $att(\mathbf{I})\hat{\mathbf{K}}$ and cutting off all saturated pixels from the images. These pixels were identified automatically using a simple thresholding filter

combined with a constraint on the minimum number 2 of such pixels in the closest neighborhood. The parameter for the denoising filter F was $\sigma^2 = 9$.

For each of the 667 cameras, 150 images from different cameras were randomly chosen to calculate PCE statistics φ. The resulting histogram is in Figure 7 (left). To see how it compares to the ideal case, we evaluated PCE for 100,050 pairs of random signals with normal Gaussian pdf and plotted next to it in Figure 7 (right).

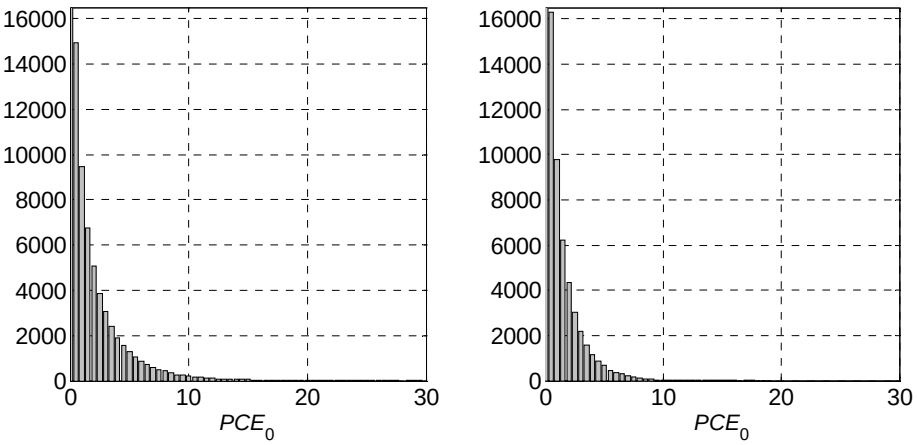

Figure 7. Histogram of PCE for full size images under H_0 (left), for simulated Gaussian distributed correlations (right).

Comparing the two histograms, the tail is heavier in Figure 7 (left). It suggests that our decision threshold φ_{th} may have to be adjusted by adding a small correction factor. As the conclusion of this section, we argue that the threshold φ_{th} does not have to be re-evaluated for every camera fingerprint. A certain correction factor may be needed to improve the estimate (14). Other measures may include additional filtering of the noise residuals. This will be a subject of our following-up research.

6 Summary

After reviewing the method of camera sensor identification by unique photo-response non-uniformity, we propose to replace the normalized correlation detector with peak to correlation energy ratio. This way, the detection threshold will not vary with varying signal length, different cameras and their on-board image processing nearly as much as for normalized correlation detector. We estimate the probability of FA directly from the threshold set on PCE, which reduces otherwise high demand for large testing needed to set up the threshold for normalized correlation used before.

We show that the linear pattern strongly limits the performance of camera sensor identification if not removed from camera fingerprint estimates. Larger normalized correlation may not necessarily mean smaller probability of FA even if evaluated for the same camera.

Acknowledgements

The work on this paper was supported by the AFOSR grant number FA9550-06-1-0046. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation there on. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of Air Force Research Laboratory, or the U.S. Government.

References

[1] Khanna, N., Mikkilineni, A.K., Chiu, G.T.C, Allebach, J.P., and Delp, E.J.: "Forensic Classification of Imaging Sensor Types." *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX*, vol. **6505**. San Jose, CA (2007) 0U–0V.

[2] Gou, H., Swaminathan, A., and Wu, M.: "Robust Scanner Identification based on Noise Features," *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX*, vol. **6505**. San Jose, California (2007) 0S–0T.

[3] Swaminathan, A., Wu, M., and Liu, K.J.R.: "Nonintrusive Component Forensics of Visual Sensors Using Output Images." *IEEE Transactions on Information Forensics and Security*, vol. **2**(1) (2007) 91–106.

[4] Propescu, A.C., and Farid, H.: "Statistical Tools for Digital Forensic," in J. Fridrich (ed.): *6th International Workshop on Information Hiding*, LNCS, vol. **3200**, Springer-Verlag, Berlin-Heidelberg, New York (2004) 128–147.

[5] Popescu, A.C., and Farid, H.: "Exposing Digital Forgeries by Detecting Traces of Resampling", *IEEE Transactions on Signal Processing*, vol. 53(2) (2005) 758–767.

[6] Popescu, A.C., and Farid, H.: "Exposing Digital Forgeries in Color Filter Array Interpolated Images." *IEEE Transactions on Signal Processing*, vol. **53**(10) (2005) 3948–3959.

[7] Farid, H.: "Exposing Digital Forgeries in Scientific Images." *Proc. ACM Multimedia & Security Workshop*. Geneva, Switzerland (2006) 29–36.

[8] Popescu, A.C., and Farid, H.: "Exposing Digital Forgeries by Detecting Duplicated Image Regions." *Technical Report*, TR2004-515. Dartmouth College, Computer Science (2004).

[9] Lukáš, J., Fridrich, J., and Goljan, M.: "Digital Camera Identification from Sensor Pattern Noise." *IEEE Transactions on Information Forensics and Security*, vol. **1**(2) (2006) 205–214.

[10] Chen, M., Fridrich, J., and Goljan, M.: "Digital Imaging Sensor Identification (Further Study)." *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX*, vol. **6505**. San Jose, California (2007) 0P–0Q.

[11] Goljan, M., and Fridrich, J., "Camera Identification from Scaled and Cropped Images," In E. J. Delp et al. editors, *Security, Forensics, Steganography, and Watermarking of Multimedia Contents X*, vol. **6819** (2008) 68190E.

[12] Chen, M., Fridrich, J., Goljan, M., and Lukáš, J.: "Determining Image Origin and Integrity Using Sensor Noise," with J. Fridrich, M. Chen, and J. Lukáš, *IEEE Transactions on Information Security and Forensics*, vol. **3**(1) (2008) 74–90.

[13] Healey, G. and Kondepudy, R.: "Radiometric CCD Camera Calibration and Noise Estimation." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. **16**(3) (1994) 267–276.

[14] Mihcak, M.K., Kozintsev, I., and Ramchandran, K.: "Spatially Adaptive Statistical Modeling of Wavelet Image Coefficients and its Application to Denoising." *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. **6**. Phoenix, Arizona (1999) 3253–3256.

[15] Lukáš, J., Fridrich, J., and Goljan, M.: "Determining Digital Image Origin Using Sensor Imperfections," *Proc. SPIE, Image and Video Communications and Processing*, vol. 5685, San Jose, California (2005) 249–260.

[16] Lukáš, J., Fridrich, J., and Goljan, M.: "Detecting Digital Image Forgeries Using Sensor Pattern Noise." *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents VIII*, vol. **6072**. San Jose, California (2006) 0Y1–0Y11.

[17] Goljan, M., Chen, M., Fridrich, J., "Identifying Common Source Digital Camera From Image Pairs," *Proc. ICIP 2007*. San Antonio, Texas (2007).

[18] Khanna, N., Mikkilineni, A.K., Chiu, G.T.C., Allebach, J.P. and Delp, E.J.: "Scanner Identification Using Sensor Pattern Noise." *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX*, vol. **6505**. San Jose, CA (2007) 1K–1.

[19] Sankur, B., Celiktutan, O., and Avcibas, I.: "Blind Identification of Cell Phone Cameras." *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX*, vol. **6505**. San Jose, California (2007) 1H–1I.

[20] Chen, M., Fridrich, J., and Goljan, M.: "Source Digital Camcorder Identification Using CCD Photo Response Non-uniformity." *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX*, vol. **6505**. San Jose, California (2007) 1G–1H.

[21] Goljan, M., Fridrich, J., and Lukáš, J., "Camera Identification from Printed Images," In E. J. Delp et al. editors, *Security, Forensics, Steganography, and Watermarking of Multimedia Contents X*, vol. **6819** (2008) 68190I.

[22] Dirik, A.E., Sencar, H.T., Husrev T., Memon, N.: "Source Camera Identification Based on Sensor Dust Characteristics," *Proc. IEEE Workshop on Signal Processing Applications for Public Security and Forensics*, Washington, DC (2007) 1–6.

[23] Choi, K.S., Lam, E.Y., Wong, K.K.Y.: "Automatic source camera identification using the intrinsic lens radial distortion," *Optics Express*, vol. **14**(24) (2006) 11551–1565.

[24] Bayram, S., Sencar, H.T., Memon, N., Avcibas, I. : "Source camera identification based on CFA interpolation," *Proc. ICIP 2005. IEEE International Conference on Image Processing* (2006) 69–72.

[25] Filler, T., and Fridrich, J.: "Using Sensor Pattern Noise for Camera Model Identification," *Proc. ICIP'08*, San Diego, California (2008) 12–15.

[26] [Bayram] Sutcu, Y.; Bayram, S.; Sencar, H.T.; Memon, N. "Improvements on Sensor Noise Based Source Camera Identification," *Proc IEEE, International Conference on Multimedia and Expo*, (2007) 24–27.

[27] Kumar, B.V.K.V., and Hassebrook, L., "Performance measures for correlation filters", Applied Optics, vol. **29**(20) (1990) 2997–3006.