FAST SEQUENTIAL FORENSIC DETECTION OF CAMERA FINGERPRINT

F. Pérez-González[⋆] *M. Masciopinto*[⋆] *I. González-Iglesias*[⋆] *P. Comesaña*[⋆]

[⋆] Signal Theory and Communications Department, University of Vigo
E. E. Telecomunicación, Campus-Lagoas Marcosende, Vigo 36310, Spain

ABSTRACT

Two sequential camera fingerprint detection methods are proposed. Sequential tests implement a log-likelihood ratio test in an incremental way, thus enabling a reliable decision with a minimal number of observations. One of our methods adapts Goljan *et al.*'s to sequential operation. The second, which offers better performance in terms of average number of test observations, is based on treating the alternative hypothesis as a doubly stochastic model. Finally, we validate the performance of our methods with experiments and compare them with the state of the art in fast camera fingerprint detection.

***Index Terms*—** camera fingerprint, fast detection, photoresponse non-uniformity, sequential test

1. INTRODUCTION

The PhotoResponse Non-Uniformity (PRNU) is a spatial pattern that acts as fingerprint or unique identifier of an image camera device. The PRNU is caused by minute imperfections in the image sensor manufacturing process, which remain constant over the life of the device. Since it is a variation in pixel responsivity, the PRNU is active when the device is illuminated [1]. Despite the PRNU is generally an extremely weak signal, given enough image samples and using signal processing techniques, it is possible to estimate it and use it as a robust fingerprint [2]. This fingerprint can be useful to address several challenges in image forensics, such as source identification, device linking, fingerprint matching, or forgery detection [3].

The problem of fast camera fingerprint detection has been tackled in previous works, often relying on a so-called *fingerprint digest* instead of the whole PRNU. Works [4, 5] focus on the problem of whether a given image was taken by a camera whose fingerprint is in a large database of fingerprints, while [6] faces the problem of common source identification between images in a large database. In [7], a fast algorithm based on both fingerprint binarization and digesting was also proposed to deal with large fingerprint databases.

However, to the best of our knowledge, PRNU detection methods are neither specialized in large image databases (as opposed to large fingerprint databases) nor able to operate sequentially, that is, capable of analyzing small blocks of image pixels until enough reliability on the hypothesis test is accumulated, thus guaranteeing that the minimum number of blocks is used to achieve a target accuracy. This is extremely valuable when checking very large amounts of images for the presence of a specific PRNU from a given device. Of particular interest to us is the fast PRNU matching in huge image databases used by the police to investigate child pornography and other cybercrime forensic cases. As an important practical use, a reliable method is needed to search for images taken by a pedophile's camera impounded by the police in huge databases, social networks or even across the Internet. This is the topic of the European Project NIFTY [8] under which our work has been carried out.

In this paper, we revisit the PRNU estimation and detection problems, and propose a fast algorithm for camera fingerprint detection. Firstly, we derive a detector that improves the classical PRNU detection. Next, a sequential algorithm working on blocks of pixels sorted by their PRNU magnitude (i.e., the *fingerprint digest* approach [4]) is proposed, performing the PRNU detection in a very fast way. Finally, the proposed sequential detector is compared with the fingerprint digest method in [4].

The paper is organized as follows: Sect. 2 reviews PRNU estimation and detection, drawing connections with existing methods, and proposing an improved detector. The fast sequential fingerprint detection algorithm based on the improved detector is presented in Sect. 3. Then, Sect. 4 shows experimental results on a dataset containing images from several devices, while Sect. 5 gives our conclusions.

Notation: Vectors are represented in boldface. The mth component of $\mathbf{x}$ is denoted as x_m. The scalar product of vectors $\mathbf{x}$ and $\mathbf{y}$ is denoted by $\langle \mathbf{x}, \mathbf{y} \rangle$, while $\mathbf{x} \circ \mathbf{y}$ and $||\mathbf{x}||$ denote the sample-wise product and the Euclidean norm, respectively.

2. MODEL

We assume that the sensor output at pixel (i, j), $y(i, j)$, can be written as $y(i, j) = [1 + k(i, j)]x(i, j) + n(i, j)$ [9], where $x(i, j)$ is the incident light intensity, $k(i, j)$ is the (possibly gamma-corrected) PRNU, and $n(i, j)$ subsumes a number of noise sources including dark current, shot noise, read-out noise and quantization noise. As $x(i, j)$ is generally unknown, it is reasonable to obtain an estimate $\hat{x}(i, j)$ from $y(i, j)$ by applying some denoising procedure and accounting for demosaicing [10]. In such case, we can write

$$y(i, j) = [1 + k(i, j)] \cdot [\hat{x}(i, j) + r(i, j)] + n(i, j), \qquad (1)$$

where $r(i, j)$ is the denoising and demosaicing residue. For simplicity, we assume that $n(i, j) \sim \mathcal{N}(0, \sigma_n^2)$ and $r(i, j) \sim \mathcal{N}(0, \sigma_{i,j}^2)$, where the latter are mutually independent. For compactness, we also introduce the *shifted PRNU*, which is $\kappa(i, j) \doteq [1 + k(i, j)]$.

2.1. PRNU estimation

From the model in (1), it is possible to formulate the PRNU estimation problem. We recall that in this case we have L available images taken with the same device from which we want to estimate the PRNU at every pixel. Assuming PRNU sample-wise independence, we can solve the estimation problem independently for each

This work was partially funded by the Spanish Ministry of Economy and Competitiveness and the European Regional Development Fund (ERDF) under projects TACTICA and COMPASS (TEC2013-47020-C2-1-R), by the Galician Regional Government and ERDF under projects "Consolidation of Research Units" (GRC2013/009), REdTEIC (R2014/037) and AtlantTIC, and by the EU under project NIFTy (HOME/2012/ISEC/AG/INT/4000003892).

pixel. Let then y_m and $\hat{x}_m$, $m = 1, \cdots, L$, denote respectively the observation and the denoised image for an arbitrary pixel of the mth available image. Also let $\mathbf{y}$, $\hat{\mathbf{x}}$ be the vectors formed by stacking the L respective samples for the pixel under analysis. Then, the log-likelihood becomes

$$L(\kappa, \hat{\mathbf{x}}, \mathbf{y}) = -\frac{1}{2}\left[\sum_{m=1}^{L} \log(2\pi\sigma_{e,m}^2) + \sum_{m=1}^{L} \frac{(\kappa\hat{x}_m - y_m)^2}{\sigma_{e,m}^2}\right], \tag{2}$$

where $\sigma_{e,m}^2 \doteq \kappa^2\sigma_m^2 + \sigma_n^2$, with σ_m^2 the variance of the estimation residue in image m for the analyzed pixel.

Taking the derivative of (2) with respect to κ and equating to zero, it is possible to write a (nonlinear) equation that gives the maximum likelihood estimate (MLE) of κ.

A simpler approach consists in neglecting the first term in (2). This gives a minimum weighted-MSE solution, namely

$$\hat{\kappa} = \arg\min_{\kappa} \sum_{m=1}^{L} \frac{(\kappa\hat{x}_m - y_m)^2}{\sigma_{e,m}^2}. \tag{3}$$

When $\sigma_m^2 = \sigma_r^2$ for all $m = 1, \cdots, L$, then taking the derivative with respect to κ and setting to zero, we obtain that $\hat{\kappa}$ must be a solution to the equation

$$\hat{\kappa}^2\langle\hat{\mathbf{x}}, \mathbf{y}\rangle\sigma_r^2 + \left(||\hat{\mathbf{x}}||^2\sigma_n^2 - ||\mathbf{y}||^2\sigma_r^2\right)\hat{\kappa} - \langle\hat{\mathbf{x}}, \mathbf{y}\rangle\sigma_n^2 = 0. \tag{4}$$

Assuming that $\sigma_n^2 \gg \sigma_r^2$, the solution to (4) becomes $\hat{\kappa} = \langle\hat{\mathbf{x}}, \mathbf{y}\rangle/||\hat{\mathbf{x}}||^2$ or equivalently, in terms of the PRNU $\hat{k}$,

$$\hat{k} = \frac{\langle(\mathbf{y} - \hat{\mathbf{x}}), \hat{\mathbf{x}}\rangle}{||\hat{\mathbf{x}}||^2}, \tag{5}$$

which in fact resembles Chen *et al.*'s estimator $\hat{k} = \langle(\mathbf{y} - \hat{\mathbf{x}}), \mathbf{y}\rangle/||\mathbf{y}||^2$ in [9], as $\hat{\mathbf{x}} \approx \mathbf{y}$.

2.2. PRNU detection

Once a PRNU estimate is available, it can be used for camera fingerprint detection purposes. Given a set of L images which have been taken from the same camera with PRNU $\mathbf{k}_0$, and a test image $\mathbf{y}_t$, both arranged in vector form,[1] we want to decide whether $\mathbf{y}_t$ has been taken from that camera or, in other words, if the PRNU $\mathbf{k}_0$ is present in $\mathbf{y}_t$. As customary, we can formulate a binary hypothesis test with the following two hypotheses:

- H_0: Image $\mathbf{y}_t$ does not contain the PRNU $\mathbf{k}_0$,
- H_1: Image $\mathbf{y}_t$ contains the PRNU $\mathbf{k}_0$.

We assume the existence of an unbiased estimate $\hat{\mathbf{k}}$ of $\mathbf{k}_0$ obtained using the method proposed in the previous section, and we denote by $\hat{\mathbf{x}}_t$ the image vector obtained from $\mathbf{y}_t$ after denoising and demosaicing. We have derived the corresponding distributions under each hypothesis; this allows us to write Neyman-Pearson's generic detector for known $\mathbf{k}_0$, and later replace the needed statistics by their estimates, as in the *Generalized Likelihood Ratio Test* (GLRT).

Thus, when H_1 holds, we can see that the difference $y_t(i,j) - \hat{x}_t(i,j)$ for the (i,j)th pixel is Gaussian with mean $k_0(i,j)\cdot\hat{x}_t(i,j)$ and variance $\sigma_H^2 = [1 + k_0(i,j)]^2\sigma_r^2 + \sigma_n^2 \approx \sigma_r^2 + \sigma_n^2$. When H_0 holds, the PRNU must be treated as unknown. Modeling it as a zero-mean random variable, it follows that $y_t(i,j) - \hat{x}_t(i,j)$ is approximately Gaussian with zero mean and variance approximately σ_H^2, because the influence of the variance of the PRNU in the total variance is negligible. From this, the likelihood-ratio test becomes

$$\frac{\langle(\mathbf{y}_t - \hat{\mathbf{x}}_t), \mathbf{k}_0 \circ \hat{\mathbf{x}}_t\rangle}{\sigma_H^2} - \frac{||\mathbf{k}_0 \circ \hat{\mathbf{x}}_t||^2}{2\sigma_H^2} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{6}$$

for some threshold η that is chosen so as to produce the desired probability of false positive.

Implementation of (6) faces two practical problems: 1) The true PRNU $\mathbf{k}_0$ is unknown; 2) σ_H^2 is unknown. Hence, focusing on the first term of (6) produces one (generally, non-sufficient) statistic that is very similar to which has been proposed by Goljan *et al.* [11]:[2]

$$u \doteq \langle(\mathbf{y}_t - \hat{\mathbf{x}}_t), \hat{\mathbf{k}} \circ \hat{\mathbf{x}}_t\rangle, \tag{7}$$

and, as we argue above, the means for the respective hypotheses are $E\{u|H_0\} = 0$ and $E\{u|H_1\} = E\{||\mathbf{k}_0 \circ \mathbf{x}_t||^2\}$.

Let us define the shift operator $\Delta_{(q_1,q_2)}$ that applied to a vector $\mathbf{x}$ representing an image, outputs the vector corresponding to a right circular shift of (q_1, q_2) pixels of such image. Then, following [11], an estimate of the variance of statistic u can be obtained as

$$\hat{\sigma}_u^2 = \frac{1}{M - |A|}\sum_{(q_1,q_2)\notin A} \langle\Delta_{(q_1,q_2)}(\mathbf{y}_t - \hat{\mathbf{x}}_t), \hat{\mathbf{k}} \circ \hat{\mathbf{x}}_t\rangle^2, \tag{8}$$

where M is the number of available pixels, A is an *exclusion set* defined as those (q_1, q_2) in a neighborhood (w.r.t. circular shifts) of the origin $(0,0)$, and $|A|$ denotes its cardinality.

Since $E\{u|H_1\}$ is difficult to obtain accurately, it may be reasonable to assume that under H_1 the statistic u has a positive but unknown mean. Then, from Karlin-Rubin theorem [12], the test

$$u' \underset{H_0}{\overset{H_1}{\gtrless}} \eta_2, \tag{9}$$

where $u' \doteq u/\hat{\sigma}_u$, is the uniformly most powerful test for a given probability of false positive P_F. However, notice that even though the test threshold η_2 can be set since P_F is computable, we cannot find the detection probability P_D as the mean under H_1 is unknown. This has important implications for the tests discussed in Sect. 3.

2.3. Improved detector

Although $\mu_{u,1} \doteq E\{u'|H_1\}$ is not known, we have experimentally found that it can be modeled as a normal random variable whose parameters depend on the statistic

$$v \doteq ||\hat{\mathbf{k}} \circ \hat{\mathbf{x}}_t||^2/\hat{\sigma}_u, \tag{10}$$

so we will denote by $\mu(v)$ and $\sigma^2(v)$ the mean and variance of $\mu_{u,1}$, respectively. Thus, $\mu_{u,1} \sim \mathcal{N}(\mu(v), \sigma^2(v))$. Notice that the statistic v resembles the second term in (6), but obtained from computable quantities. However, the laws $\mu(v)$ and $\sigma^2(v)$ are device-dependent, so they must be learned during the PRNU extraction phase, which can be done concurrently with the estimation of $\mathbf{k}_0$, as it is described in Sect. 2.4. Furthermore, although the distribution of u' under H_0 can be modeled by a Gaussian, a slight improvement is afforded by employing a zero-mean generalized Gaussian distribution with scale parameter α_0 and shape parameter c_0, which can be reliably estimated using images from different cameras [13] (see Sect. 2.4). With all these considerations, the test becomes

$$\left(\frac{|u'|}{\alpha_0}\right)^{c_0} - \frac{(u' - \mu(v))^2}{2\sigma^2(v)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_3. \tag{11}$$

[1] From now on, vectors contain pixel values from a single image.

[2] Goljan *et al.* use $\mathbf{y}_t$ instead of $\hat{\mathbf{x}}_t$ in the second term of the scalar product.

2.4. Model training

At this point, it is necessary to estimate the model parameters for both hypotheses, which is done concurrently with the estimation of $\mathbf{k}_0$. For H_1, $\mu(v)$ and $\sigma^2(v)$ must be estimated for the target camera, while for H_0 the parameters α_0 and c_0 are estimated from the universe of available images.

Let $\mathbf{y}_{tr}$ be one of the L available images for training. Then, $\hat{\mathbf{k}}$ is estimated from the remaining $L-1$ images by using (5), and pairs of (u', v) values are obtained from $\mathbf{y}_{tr}$ and $\hat{\mathbf{k}}$ by taking blocks of pixels with the same size as that used in the hypothesis test. This process is repeated for each of the L images in the training set to produce a collection of (u', v) pairs that is used to estimate $\mu(v)$ and $\sigma^2(v)$. This is done by binning the values of v and for each bin calculating the mean and variance of the corresponding set of u' values.

3. SEQUENTIAL TEST FOR FAST PRNU DETECTION

The Sequential Probability Ratio Test (SPRT) was proposed by A. Wald in [14]. On a hypothesis testing problem, the main purpose of the SPRT is to minimize the expected number of observations to achieve error probabilities less than some pre-fixed target probabilities of misdetection (P_M^*) and false positive (P_F^*).

Let $x_1, x_2, \ldots, x_N$ be i.i.d. observations, with N its number, and $f(x_j|H_i)$ the probability density function of the jth observation conditioned on the ith hypothesis ($i = 0, 1$). In a SPRT, the likelihood ratio is compared with two thresholds, and the test accepts one hypothesis or continues processing another observation (a.o.) while

$$B \underset{\text{a.o.}}{\overset{H_0}{\gtrless}} \prod_{j=1}^{n} \frac{f(x_j|H_1)}{f(x_j|H_0)} \underset{\text{a.o}}{\overset{H_1}{\lessgtr}} A, \tag{12}$$

where n is the number of processed observations which is set to 1 at the beginning and increased step by step until the test accepts either hypothesis, or n reaches the number of available observations N.

The thresholds are chosen so as to control the error probabilities on each hypothesis. Following Wald's approximation [14], the relations between P_M^*, P_F^*, and the thresholds are

$$A \leq \frac{1 - P_M^*}{P_F^*}, \quad B \geq \frac{P_M^*}{1 - P_F^*}, \tag{13}$$

where equalities are usually a good choice in practice.

Therefore, the hypothesis testing problem discussed in Sect. 2.2 and, hence, our proposed detector in (11), can be transformed into a SPRT by taking logarithms in (12). Then, the resulting SPRT is

$$\eta_B(n) \underset{\text{a.o.}}{\overset{H_0}{\gtrless}} \sum_{j=1}^{n} D_j \underset{\text{a.o.}}{\overset{H_1}{\lessgtr}} \eta_A(n), \tag{14}$$

where the thresholds are $\eta_A(n) \doteq \log(A) - n \cdot \log(2\alpha_0\Gamma(1/c_0))$ and $\eta_B(n) \doteq \log(B) - n \cdot \log(2\alpha_0\Gamma(1/c_0))$, and

$$D_j \doteq \left(\frac{|u'_j|}{\alpha_0}\right)^{c_0} - \frac{(u'_j - \mu(v_j))^2}{2\sigma^2(v_j)} - \log\left(c_0\sqrt{2\pi\sigma^2(v_j)}\right). \tag{15}$$

Figure 1 summarizes the proposed algorithm for fast camera fingerprint detection. Let $\mathcal{I} \doteq \{1, \cdots, M\}$ be the set of vector indices. Then, the values of $\hat{\mathbf{k}}$ are sorted by decreasing magnitude, thus inducing a permutation $\pi : \mathcal{I} \to \mathcal{I}$ which is then applied to any given test image $\mathbf{y}_t$ and its denoised counterpart $\hat{\mathbf{x}}_t$. With some abuse of notation to keep it simple, henceforth we denote by $\hat{\mathbf{k}}$, $\mathbf{y}_t$, $\hat{\mathbf{x}}_t$, the reordered vectors. Those vectors reordered according to π are divided

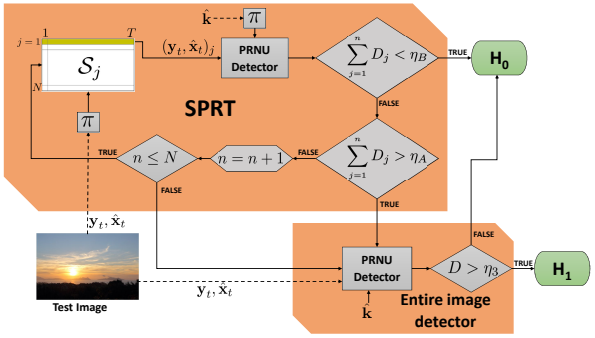

Fig. 1. Sequential test implementation.

into successive blocks $\mathcal{S}_j$, $j = 1, \cdots, N$, with T pixels each. The jth pair of statistics (u'_j, v_j) is computed by using in (7) and (10) only those pixels in block $\mathcal{S}_j$. If n reaches N without a decision being taken, then the entire image is analyzed with a non-sequential test. In addition, those images classified as H_1 by the SPRT are retested using the whole image, in order to achieve the smallest P_M.

3.1. Fixed-parameter PRNU Sequential Test

Here we propose an SPRT based on the well-known detection method by Goljan *et al.* in [11]. Such method sets the detection threshold on the basis of a target P_F, but entirely disregards hypothesis H_1 and, consequently, P_M. See (9). However, knowledge of P_M is necessary to implement a SPRT as is apparent from (13).

In contrast, our detector in (11) overcomes this issue, as now the distribution of u' under H_1 is well defined. In order to quantify how much is gained by learning $\mu(v)$ and $\sigma^2(v)$, we have studied the performance of our detector in (11) when $\mu(v)$ and $\sigma^2(v)$ are independent of v. The results are reported in Sect. 4.

3.2. Fast variance estimation

The variance estimator $\hat{\sigma}_u^2$ in (8) is taken from [11], which is quite time-consuming as it contains two nested sums (corresponding to the scalar product and the averaging over the spatial shifts). Since one of the aims of the methods proposed in this paper is to reduce the detection time, we propose the following simpler estimator

$$\hat{\sigma}_u^{'2} = \frac{1}{M}||\hat{\mathbf{k}} \circ \hat{\mathbf{x}}_t||^2||(\mathbf{y}_t - \hat{\mathbf{x}}_t)||^2. \tag{16}$$

In terms of camera fingerprint detection performance, there is no significant difference between both variance estimation methods, since the Areas Under Curve (AUCs) for the respective Receiver Operating Characteristic (ROC) curves are practically identical (differences show up in the 4th significant digit).

However, if M is the total number of pixels in the image, the estimator in (16) is $M - |A| \approx M$ times faster, which even for small images results in huge savings.

4. EXPERIMENTAL RESULTS

The image database for the experimental results is a collection of pictures from different sources. The database is composed of TIFF images coming from our own cameras, the Dresden image database [15], and the Raise database [16].

For PRNU extraction, $L = 50$ images are randomly selected for those devices in Table 1 with more than 50 available images. Extraction is done as described in Sect. 2.1. The experimental results

Table 1. Cameras used in SPRT experiments.

Camera Model	Sensor	Native resolution	Devices	Number of images	Database
Canon 600D	22.3x14.9mm CMOS	5184x3456	1	241	Own
Canon1100D	23.2x14.7mm CMOS	4272x2848	3	316/122/216	Own
Nikon D60	23.6x15.8mm CCD	3872x2592	1	197	Own
Nikon D70	23.7x15.6mm CCD	3008x2000	2	43/43	Dresden
Nikon D70S	23.7x15.6mm CCD	3008x2000	2	43/47	Dresden
Nikon D90	23.6x15.8mm CMOS	4288x2848	1	250	Raise
Nikon D200	23.6x15.8mm CCD	3872x2592	2	48/43	Dresden
Nikon D3000	23.6x15.8mm CCD	3872x2592	1	230	Own
Nikon D3200	23.2x15.4mm CMOS	6016x4000	1	250	Own
Nikon D5100	23.6x15.6mm CMOS	4928x3264	1	250	Own
Nikon D7000	23.6x15.6mm CMOS	4928x3264	1	250	Raise

Table 2. Results for both SPRTs. $P_D^* = 0.98$ and $P_F^* = 0.05$.

Device	**SPRT $\mu(v)$ and $\sigma(v)$**				**SPRT with fixed μ and σ^2**			
	P_D	$\bar{n}^d_{H_1}$	P_F	$\bar{n}^d_{H_0}$	P_D	$\bar{n}^d_{H_1}$	P_F	$\bar{n}^d_{H_0}$
Canon 600D	1	1	0.001	1.51	1	1	0.0006	1.75
Canon 1100D #1	0.998	1.03	0.014	2.76	1	1.03	0.020	3.51
Canon 1100D #2	1	1.46	0.034	6.94	0.997	1.44	0.053	7.57
Canon 1100D #3	0.99	1.06	0.024	3.66	1	1.06	0.031	5.32
Nikon D60	0.981	1.71	0.031	6.60	0.984	1.68	0.034	7.06
Nikon D90	0.853	2.34	0.027	7.18	0.858	2.69	0.020	7.08
Nikon D3000	0.99	1.02	0.015	3.65	1	1.01	0.013	3.71
Nikon D3200	1	1	0.009	3.72	1	1	0.008	4.07
Nikon D5100	0.995	1.56	0.035	9.45	0.992	1.65	0.035	8.46
Nikon D7000	0.988	1.75	0.025	4.75	0.991	1.74	0.029	5.47

sequentially pick for hypothesis H_1 each device with more than 50 images, and H_0 all the images from the remaining devices.

In order to make the results independent of the specific choice of L images and also to increase the number of test images corresponding to H_1, the reported results are the average of 5 different random selections of the L images, using the remaining images of each selection to test H_1.

In all the experiments, the denoised images $\hat{\mathbf{x}}_t$ are obtained using the filter described in [10]. In addition, the estimated PRNU in (5) is postprocessed to remove the artifacts discussed in [9], so mean-removal and Wiener filtering in the Fourier domain are applied.

For each test image, non-overlapping successive blocks $\mathcal{S}_j$ of size 1024 pixels, $j = 1, \cdots, N$, are taken from the reordered versions $\mathbf{y}'_t$ and $\hat{\mathbf{x}}'_t$. The maximum number of observations N is fixed to 256 because we have experimentally found that for TIFF images a size of 512×512 should be enough for successful PRNU detection. The results obtained after the SPRT described in Sect. 3 are shown in Table 2, where $\bar{n}^d_{H_0}$ and $\bar{n}^d_{H_1}$ denote the average number of observations that the SPRT needs in order to make a decision for H_0 and H_1, respectively, and $d = 1, \cdots, D$ indexes the devices in Table 2. The parameters for H_0 were set to $\alpha_0 = 1.22$ and $c_0 = 1.84$ for all the experiments after applying maximum likelihood estimation. The SPRT observation track for some images is shown in [17].

As Table 2 shows, the target probabilities $P_D^* \doteq 1 - P_M^*$ and P_F^* are achieved, with the exception of Nikon D90. We conjecture that this is due to Nikon D90 images showing a high saturation degree, as confirmed by the metadata. Moreover, the improved detector based on $\mu(v)$ and $\sigma^2(v)$ offers a small gain in terms of $\bar{n}_{H_0}$ steps, so it is slightly faster for the same target probabilities.

It is important to quantify the computational savings achieved by the SPRT with respect to a full-image test. If O_F is the cost of classifying a full image, and O_S that of our SPRT detector, both are directly proportional to the respective number of pixels. The total number of pixels is M, where as for the SPRT detector is $M' = \bar{n}T$, where $\bar{n} = \bar{n}_0 \cdot p_{H_0} + \bar{n}_1 \cdot p_{H_1}$, with $\bar{n}_0 \doteq 1/D \sum_d \bar{n}^d_{H_0}$, $\bar{n}_1 \doteq 1/D \sum_d \bar{n}^d_{H_1}$, and $p_{H_0}, p_{H_1} = (1 - p_{H_0})$ the prior probabilities of hypotheses H_0, H_1, respectively.

Table 3. Fixed-length digest vs. SPRT. Numbers marked with † refer to $P_D < P_D^* = 0.98$.

Device	**SPRT** $(\mu(v), \sigma^2(v))$		**Fixed-length Digest**			
	ℓ_d	P_D	$P_D(\ell_d)$	$P_D(\bar{\ell})$	$P_D(\ell^+)$	$P_D(\ell^-)$
Canon 600D	1542	1	1	1	1	1
Canon 1100D #1	2809	0.998	0.998	0.998	1	0.993
Canon 1100D #2	7051	1	0.994	0.994	0.997	$0.958^\dagger$
Canon 1100D #3	3722	0.999	0.996	0.999	1	0.99
Nikon D60	6709	0.981	$0.969^\dagger$	$0.962^\dagger$	0.98	$0.935^\dagger$
Nikon D90	7303	$0.853^\dagger$	$0.861^\dagger$	$0.849^\dagger$	$0.869^\dagger$	$0.82^\dagger$
Nikon D3000	3711	0.999	1	1	1	0.998
Nikon D3200	3782	1	1	1	1	1
Nikon D5100	9597	0.995	0.99	0.981	0.993	$0.92^\dagger$
Nikon D7000	4834	0.988	$0.974^\dagger$	$0.97^\dagger$	0.989	$0.88^\dagger$

Hence, when testing a large database with the SPRT (see Fig. 1), the computational cost is proportional to $O_S + [P_D^* p_{H_1} + P_F^* p_{H_0}] \cdot O_F$, where the term in brackets is the probability that the test gives a (true or false) positive. Then, the saving is given by

$$\frac{O_S}{O_F} = P_D^* \cdot p_{H_1} + P_F^* \cdot p_{H_0} + \bar{n}T/M. \tag{17}$$

For a database with images of size $M = 2000 \times 3000$ pixels and $p_{H_1} = 0.01$, a sequential detector with blocks of size $T = 1024$, and the P_D^*, P_F^* values of Table 2, the ratio in (17) is ~ 0.3. Moreover, the computation of the estimated variance following (16) produces an additional reduction factor of $1/M \approx 1.6 \cdot 10^{-7}$.

Finally, in order to compare the proposed SPRT detector in Sect. 2.3 with the fixed-length digest in [4], the length of an equivalent digest for the latter is calculated as $\ell_d = \lceil T \cdot (\bar{n}^d_{H_1} \cdot p_{H_1} + \bar{n}^d_{H_0} \cdot p_{H_0}) \rceil$ for the dth device in Table 2, and P_D is obtained by setting P_F as given by the SPRT. Thus, both algorithms are set to operate with the same computational cost for each device and the same P_F, and hence can be compared in terms of P_D. The results are reported in Table 3, together with the average ($\bar{\ell}$), the maximum (ℓ^+), and the minimum (ℓ^-) digest lengths across all cameras. These values are, respectively, $\bar{\ell} = 5106$, $\ell^+ = 9597$, and $\ell^- = 1542$.

As seen in Table 3, for the same computational cost, our test offers a slight advantage in terms of P_D. However, such comparison would not fully convey the real advantage of our method, as in a practical scenario, the digest length in [4] must be set beforehand. Hence, if a too-small length (e.g., ℓ^-) is used, the computational cost is smaller than ours, but the target P_D^* is not achieved for half of the cameras. Conversely, if a large length (e.g., ℓ^+) is chosen, a P_D similar to that of our test is achieved, but the average computational cost increases by $(\ell^+ - \bar{\ell})/\bar{\ell} \cdot 100 \approx 88\%$. In contrast, our SPRT is flexible and smart so as to stop as soon as P_D^* and P_F^* are achieved, and without prior knowledge of the power of the device fingerprint.

5. CONCLUSIONS

In this paper we have shown how Wald's sequential test can be implemented for PRNU detection purposes, with the advantage of enabling a very fast test that makes a reliable decision with a minimum number of observations.The proposed tests are especially useful when very large image databases must be searched for device verification, since they automatically adapt to the target detection probabilities without prior knowledge of the power of the device fingerprint, and resulting on a reduced computational cost to achieve those target probabilities.

6. REFERENCES

[1] Gerald C Holst, *CCD Arrays, Cameras, and Displays*, SPIE Optical Engineering Press Bellingham, WA, 2nd edition, 1998.

[2] Jan Lukáš, Jessica Fridrich, and Miroslav Goljan, "Digital camera identification from sensor pattern noise," *IEEE Trans. Inf. Forensics Security*, vol. 1, no. 2, pp. 205–214, Jun. 2006.

[3] Jessica Fridrich, "Digital image forensics," *IEEE Signal Process. Mag.*, vol. 26, no. 2, pp. 26–37, Mar. 2009.

[4] Miroslav Goljan, Jessica Fridrich, and Tomáš Filler, "Managing a large database of camera fingerprints," in *Proc. SPIE, Electronic Imaging, Media Forensics and Security XII*, Jan. 2010, vol. 7541, pp. 08 01–12.

[5] Yongjian Hu, Chang-Tsun Li, Zhimao Lai, and Shangfan Zhang, "Fast camera fingerprint search algorithm for source camera identification," in *Proc. 5th Int. Symp. Commun. Control Signal Processing (ISCCSP)*, May 2012, pp. 1–5.

[6] Floris Gisolf, Pelle Barens, Ewald Snel, Anwar Malgoezar, Mark Vos, Arjan Mieremet, and Zeno Geradts, "Common source identification of images in large databases," *Forensic science international*, vol. 244, pp. 222–230, Nov. 2014.

[7] Sevinç Bayram, Hüsrev Taha Sencar, and Nasir Memon, "Efficient sensor fingerprint matching through fingerprint binarization," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 4, pp. 1404–1413, Aug. 2012.

[8] "NIFTY website," Online: `http://research.ncl.ac.uk/nifty/`.

[9] Mo Chen, Jessica Fridrich, Miroslav Goljan, and Jan Lukáš, "Determining image origin and integrity using sensor noise," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 1, pp. 74–90, Mar. 2008.

[10] M Kivanc Mihcak, Igor Kozintsev, and Kannan Ramchandran, "Spatially adaptive statistical modeling of wavelet image coefficients and its application to denoising," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Mar. 1999, vol. 6, pp. 3253–3256.

[11] Miroslav Goljan, Jessica Fridrich, and Tomáš Filler, "Large scale test of sensor fingerprint camera identification," in *Proc. SPIE, Electronic Imaging, Media Forensics and Security XI*, Feb. 2009, vol. 7254, pp. 0I 1-0I 12.

[12] Samuel Karlin and Herman Rubin, "The theory of decision procedures for distributions with monotone likelihood ratio," *The Annals of Mathematical Statistics*, vol. 27, no. 2, pp. 272–299, Jun. 1956.

[13] K.A. Birney and T.R. Fischer, "On the modeling of DCT and subband image data for compression," *IEEE Trans. Signal Process.*, vol. 4, no. 2, pp. 186–193, Feb. 1995.

[14] A. Wald, "Sequential tests of statistical hypotheses," *The Annals of Mathematical Statistics*, vol. 16, no. 2, pp. 117–186, Jun. 1945.

[15] Thomas Gloe and Rainer Böhme, "The 'Dresden Image Database' for benchmarking digital image forensics," in *Proc. of the 25th Symp. On Applied Computing (ACM SAC 2010)*, Mar. 2010, vol. 2, pp. 1585–1591.

[16] D.T. Dang-Nguyen, C. Pasquini, V. Conotter and G. Boato, "Raise a raw images dataset for digital image forensics," in *Proc. 6th ACM Multimedia Systems Conference*. ACM, Mar. 2015, pp. 219–224.

[17] Fernando Pérez-González, Iria González-Iglesias, Miguel Masciopinto, and Pedro Comesaña, "Fast sequential forensic camera identification," arXiv:1510.03730, Oct. 2015.