



HAL
open science

Individual camera device identification from JPEG images

Tong Qiao, Florent Retraint, Rémi Cogranne, Thanh Hai Thai

► **To cite this version:**

Tong Qiao, Florent Retraint, Rémi Cogranne, Thanh Hai Thai. Individual camera device identification from JPEG images. *Signal Processing: Image Communication*, 2017, 52, pp.74-86. 10.1016/j.image.2016.12.011 . hal-01915647

HAL Id: hal-01915647

<https://hal.science/hal-01915647v1>

Submitted on 19 Mar 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Individual Camera Device Identification from JPEG Images

Tong Qiao^{a,*}, Florent Retraint, Rémi Cogranne^a, Thanh Hai Thai^b

^aICD, UMR 6281 CNRS, Troyes University of Technology, Troyes, France.

^bInstitute of Research and Development, Duy Tan University, Da Nang, Viet Nam.

Abstract

The goal of this paper is to investigate the problem of source camera device identification for natural images in JPEG format. We propose an improved signal-dependent noise model describing the statistical distribution of pixels from a JPEG image. The noise model relies on the heteroscedastic noise parameters, that relates the variance of pixels' noise with the expectation considered as unique fingerprints. It is also shown in the present paper that, non-linear response of pixels can be captured by characterizing the linear relation because those heteroscedastic parameters, which are used to identify source camera device. The identification problem is cast within the framework of hypothesis testing theory. In an ideal context where all model parameters are perfectly known, the Likelihood Ratio Test (LRT) is presented and its performance is theoretically established. The statistical performance of LRT serves as an upper bound of the detection power. In a practical identification, when the nuisance parameters are unknown, two generalized LRTs based on estimation of those parameters are established. Numerical results on simulated data and real natural images highlight the relevance of our proposed approach. While those results show a first positive proof of concept of the method, it still requires to be extended for a relevant comparison with PRNU-based approaches that benefit from years of experiences

Keywords: Digital forensics, device identification, improved signal-dependent noise model, nuisance parameters, hypothesis testing.

I. Introduction and Contributions

Digital image forensics has received an increasing attention in past decades. It is widely referred to as the art, or methods, for distinguishing between a realistic and faked image. With the development of information industry, digital images can be easily falsified by using a low-cost photo-editing software. Therefore, it is likely that unscrupulous individuals may be willing to alter and modify the content of digital images to distort the truth. Many such examples have been publicly exposed and may have major negative consequences. To restore the trust to digital images among the public, it is urgent to propose reliable and robust forensic tools against malicious image falsification.

a. State of the Art

Digital image forensics aims at addressing the two followings main problems: image origin identification and image content integrity; while the former aims at identifying the source camera model or device, the latter aims at detecting any sort of digital image manipulation (see details in [1, 2, 3]). To address these problems, researchers have developed a variety of image forensic algorithms to protect against attacks from image forgers. In general, it is proposed to classify all the forensic methodologies into two categories: active forensics and passive forensics. Active forensics involves forensic techniques which are based on prior-embedded relevant information such

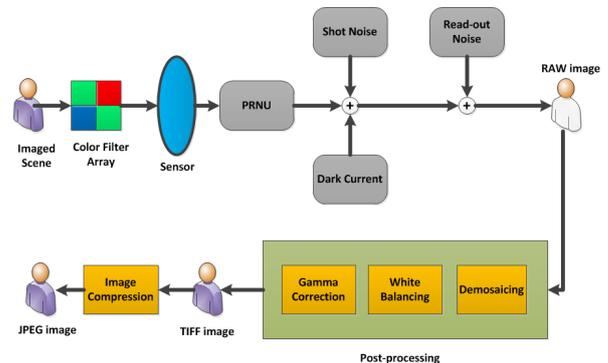


Figure 1: Illustration of a typical imaging pipeline within a digital camera.

as a digital watermark or signature. Since the embedding mechanism has to be available, active forensics has its limitation in a context of a widely adopted method. Therefore, passive forensics, which does not require any embedded information, is the current main focus of digital image forensics. In this context, the present paper mainly addresses the problem of image origin identification.

In general, when a photographer captures a digital image by using a camera, the image is stored with the header files (e.g. EXchangeable Image File, EXIF and Joint Photographic Experts Group, JPEG headers), which contains all recording and compression history. Thus, forensic investigators can potentially access to all the information on the inspected image by extracting its header files. Header files as extrinsic fingerprints can be used to solve the problem of identifying source cam-

*Corresponding author.

Email address: tong.qiao@utt.fr (Tong Qiao)

era. The original header files, however, can be easily removed or replaced due to post-camera operations, that are referred to as all possible operations that one can apply on a digital image after its acquisition; nowadays free and open source softwares for image header files are readily available such as exiftool software or libexif library. Besides, a mass of photos shared on the Internet, especially on the social web-site such as Facebook or Twitter or on photo-sharing platform such as Flickr web-site, do not always have their header files. Therefore, passive forensic investigators do not use header files as fingerprints for identifying source camera as it is usually considered as unreliable. In the image acquisition pipeline as illustrated in Fig. 1 (see details in [4, 5]), each step probably leaves some traces in the image. The passive forensics utilizes the traces as intrinsic fingerprints to acquire the concerning information of the camera source¹. The methods that investigate the problem of image origin identification aim at addressing the following questions:

- Which intrinsic/in-camera² fingerprint can be exploited for identifying with reliability and accuracy the image origin?
- How to extract accurately the proposed fingerprint from a single image or several images that may have a textured context and an important noise level in some areas?
- What are the statistical performance of the proposed detector, e.g. what is the probability that may classify two fingerprints extracted from different images as coming from the same source?

Many forensic detectors have been proposed, which can be generally formulated into two schemes: *supervised classification* and *unsupervised classification*, see [1, 2, 6] for a detailed review:

1. The methodologies in the category of supervised classification mainly rely on identifying unique traces of the imaging pipeline and exploiting the classification approach from a supervised machine. Many different traces (or adapted features set) have been proposed and many statistical learning frameworks have been used. Only a few works have been proposed to identify the camera device from early-acquisition stages such as lens distortion/aberration [7, 8, 9]. Forensic investigators can also utilize the traces left by the stage of post-processing for source camera identification. For instance, using the white balancing as a camera fingerprint, the algorithm proposed in [10] identified the device origin of a given image. The methodologies proposed by [11, 12, 13, 14] made use of Color Filter Array (CFA) and demosaicing algorithms to identify the camera model. Besides, the features extracted

by JPEG compression were also considered as influential factors to identify the camera origin [15]. Based on extracted features from the acquisition pipeline, these algorithms use supervised learning methods (such as Support Vector Machine, SVM) to identify source camera device from a digital image. However, the challenging problems are that a few manufacturers share the similar image processing techniques and most of the time partially the same components produced by a few manufacturers lead to the similarity when extracting fingerprints from given images, especially in the case of the same camera manufacturer and model. Moreover, the application of supervised learning is time-consuming. Besides, several problems such as accuracy of feature selection, robustness to a mismatch between training and testing sets, and the establishment of detection performance remain open since with any statistical learning approach the performance are only measured empirically on a validation dataset.

2. The methods in the category of unsupervised classification aim at identifying the unique noise as fingerprints of the acquisition device. Indeed, the level of noise that affects each pixel individually can be considered as a unique fingerprint for each camera device. The origin of noise level variation from pixel to pixel is due to imperfections during sensor manufacturing process and non-uniformity of photo-electronic conversion caused by inhomogeneity of silicon wafer (see [16, 4] for details). Hence, Sensor Pattern Noise (SPN) was generated, which can be extracted from a given image, and was first used for identifying source camera device [17]. Then, [18, 19, 20, 21, 22] improved the prior algorithm. It should be noted that the SPN includes two main components: Fixed Pattern Noise (FPN) and Photo-Response Non-Uniformity noise (PRNU). The FPN was represented by dark current, which was used in [23]. Since the FPN can be suppressed by subtracting a dark frame from the output image, it was proposed not to use the FPN any more for source camera identification. On the contrary, the PRNU is more robust than the FPN. Based on the PRNU, the algorithms proposed in [18, 19, 20, 24] directly investigated the problem of source camera identification. Although those PRNU-based detectors performed efficiently, the counter-forensic algorithm [25] challenged the credibility of their detection power. Therefore, the main challenge in this category is to extract relevantly reliable noise-based fingerprints from a given image. Besides, in the literature, few detectors investigate the hypothesis testing theory and are designed based on the statistical image model. Hence, the performance of detectors still remains analytically unestablished.

b. Contributions of the Paper

In a practical context, it is proposed to use the various sources of random noise (e.g. shot noise and read-out noise) as the unique fingerprints for individual device identification. In our prior research, based on the generalized signal-dependent noise model [26], a statistical test that can identify source camera

¹The term *source* means an individual camera device, a camera model, or a camera brand. Other sources such as cell-phone cameras, scanners are not addressed in this paper.

²The term *in-camera* is referred to as all the processing steps that are unavoidably carried out during image acquisition pipeline (see Fig. 1), which do not include post-camera operations.

model, from a given image, was designed in [27]. However, those prior works have the indisputable disadvantage to be unable of distinguishing different devices from the same camera model. By improving the signal-dependent noise model and extracting block fingerprints, we develop a new statistical test that aims at identifying the individual camera device from JPEG images. Hence, the main contributions of this paper are summarized below:

- By taking into account the impact of the main in-camera post-processing stage (such as white balancing, gamma correction) on the variance of each pixel and by studying the non-linearity of pixel's response, this paper proposes an improved signal-dependent noise model describing more accurately the relation between pixel's expectation and variance.
- This novel noise model is used in the present paper over blocks of several pixels to extract camera device fingerprints. Moreover, it is shown that the camera parameters describing our improved signal-dependent noise model have a linear relation, which can be used for designing our proposed tests.
- In an ideal context, that is when all the parameters are perfectly known, the Likelihood Ratio Test (LRT) is presented, and its properties are theoretically established. The statistical performance of LRT serves as an upper bound on the detection power of any test that aims at identifying source camera device using individual pixel's noise properties.
- In a practical context, it is proposed to design two Generalized Likelihood Ratio Tests (GLRT)s. In the first scenario, in which both camera parameters characterizing the improved signal-dependent noise model and model parameters describing the relation of camera fingerprints are known, and the statistical parameter of estimate errors is unknown, the first GLRT is designed. In another scenario, when both camera and model parameters are unknown, as well as the statistical parameter of estimate errors, the second GLRT is established.
- Numerical simulations show the sharpness of the theoretically established results and the good performance of our proposed tests. Comparisons with prior-art detector [19, 26] on the real image dataset show the efficiency of our proposed method for source camera device identification.

c. Organization of the Paper

This paper is organized as follows. Section II reviews the generalized signal-dependent noise model for a JPEG image, then improves this noise model. Based on this novel noise model, this section presents the parameters that are used as fingerprints for camera device identification, details their extraction over blocks of pixels, and presents the linear relation between those parameters. Based on the proposed linear model

for block fingerprints extracted from several JPEG images, the LRT is established and its statistical performance is presented in Section III. While the optimal LRT requires knowledge of camera fingerprints, model parameters and estimate error's parameter, Section IV presents the first practical GLRT for JPEG images that estimates the error's parameter, but still assumes both camera and model parameters known. Section V proposes even more practical GLRT that estimates the camera fingerprints, model parameters and error's parameter. Section VI presents numerical results of the proposed detectors on the simulated and real image dataset, and also presents comparison with current art. Finally, Section VII concludes this paper.

II. Proposed Camera Fingerprints for JPEG Images

a. Overview on Generalized Signal-Dependent Noise Model and Its Limitation

Digital image acquisition pipeline, depicted in Fig. 1, actually consists in various processing stages briefly described below (the reader interested may find more details in [4, 5]). First, light intensity measured at each pixel generates an electrical signal that is read out as a RAW format image. This RAW image is then subjected to several post-acquisition processes, such as demosaicing, white balancing and gamma correction, to issue a full resolution, colored and uncompressed image, that is referred to as a TIFF image in this paper. Besides, other processing operations, that widely differ among different camera models, such as camera denoising and edge enhancement are not discussed in this paper. Then, for storage, an image compression algorithm is applied very often using the JPEG standard and, hence, a JPEG image is finally output. In fact, our proposed model in this paper for JPEG image is evolved from the model designed for RAW data. Hence, in the following part, we first introduce a statistical model for RAW image in order to explain our proposed model for JPEG image.

Let us denote $\mathbf{Z} = \{z_i\}$, $i \in \{1, \dots, I\}$ a RAW image, made of I pixels. The noise corrupting a RAW image is usually described using the Poissonian-Gaussian noise model [28], which includes Gaussian read-out noise and Poissonian photo-counting shot noise, as follows:

$$z_i = \mu_{z_i} + \xi_{z_i} \sim \mathcal{N}(\mu_{z_i}, a\mu_{z_i} + b). \quad (1)$$

This model describes the pixel noise variance as an affine function of the pixel's expectation, characterized by the parameters (a, b) . Recently, this model has been extended in [29, 30] to describe a RAW image.

In fact, since demosaicing and white balancing are linear transformation, see details in [5], the application of those processes does not modify the Gaussian distribution of pixels up to the parameters (a, b) :

$$y_i = \mu_{y_i} + \eta_{y_i} \sim \mathcal{N}(\mu_{y_i}, \bar{a}\mu_{y_i} + \bar{b}) \quad (2)$$

where y_i denotes a pixel after white balancing and demosaicing, but before gamma correction, and μ_{y_i} the expectation of y_i which still characterizes the variance through the affine relation

$\tilde{a}\mu_{y_i} + \tilde{b}$. In addition, $\eta_{y_i} \sim \mathcal{N}(0, \tilde{a}\mu_{y_i} + \tilde{b})$ represents the signal-dependent noise with zero mean. Next, a usual processing that occurs, to adjust the brightness with former CTR (Cathode Ray Tube) display device, is the gamma correction, which is a pixel-wise non-linear operation defined as:

$$x_i = y_i^{\frac{1}{\gamma}} \quad (3)$$

where x_i denotes a TIFF pixel after gamma correction, and γ the correction factor. The statistical distribution of pixels after gamma correction is not very simple. However, based on the first order of Taylor series expansion of $(1+p)^{\frac{1}{\gamma}}$ at $p=0$, x_i can be approximated by:

$$x_i \approx \mu_i + \frac{1}{\gamma}\mu_i^{1-\gamma}\eta_i \sim \mathcal{N}(\mu_i, \sigma_i^2) \quad (4)$$

$$\text{with } \sigma_i^2 = \frac{1}{\gamma^2}\mu_i^{2-2\gamma}(\tilde{a}\mu_i^\gamma + \tilde{b}) \quad (5)$$

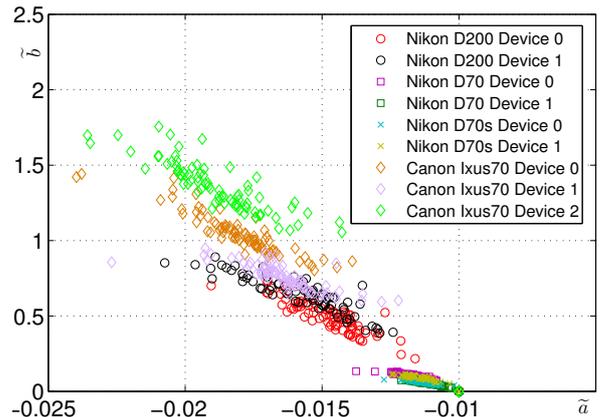
where μ_i and σ_i^2 represent the expectation and the variance of pixel x_i . Finally, an image in TIFF format is usually transformed to a JPEG one for storage. Thus, let us consider the quantization noise modelled as an additive noise that is uniformly distributed and independent of the input signal, see [31]. Then Equation (5) can be rewritten as follows:

$$\sigma_i^2 = \frac{1}{\gamma^2}\mu_i^{2-2\gamma}(\tilde{a}\mu_i^\gamma + \tilde{b}) + \frac{\Delta^2}{12} \quad (6)$$

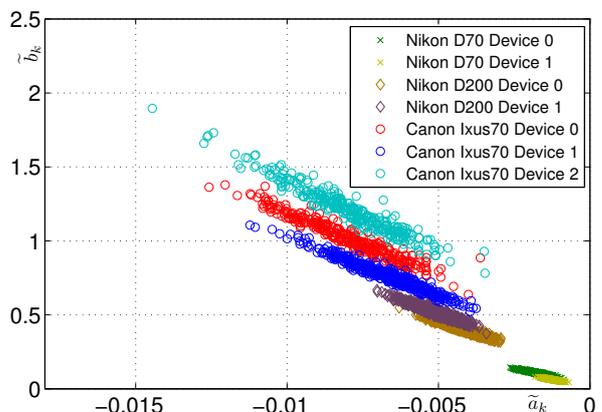
where $\frac{\Delta^2}{12}$ is the term modelling the variance of additive quantization noise. Besides, in this paper we assume, for clarity, that quantization step Δ equals 1. It should be noted that based on Equation (6), x_i of the generalized signal-dependent model (4) is redefined as the pixel from an image in JPEG format. It also should be noted that in this model, camera parameters (\tilde{a}, \tilde{b}) are constant for the whole inspected image.

Compared with the Poissonian-Gaussian noise model for a RAW image, the extra parameter γ of the model (4) is very difficult to estimate since it has a non-linear impact on pixel's statistical properties. Finally, it should also be noted that the parameters $(\tilde{a}, \tilde{b}, \gamma)$ estimated from JPEG images are relatively insensitive to ISO sensitivity parameters, unlike the fingerprints used for RAW images, see more details in [29, 30]. This allows the design of fingerprints and a statistical test for camera device identification which are independent of ISO sensitivity settings. It should be noted that in the present paper the pixel's value of the inspected image is used; hence we do not intend in this paper to model the DCT coefficient, see details in [32, 33, 34] for instance.

In our prior research [27], this generalized signal-dependent noise model (4) has been used for extracting the camera model fingerprints (\tilde{a}, \tilde{b}) from a single JPEG image. However, (\tilde{a}, \tilde{b}) can only be used to identify source camera model. Unfortunately they can not identify different devices of the same camera model. Fig. 2a provides a simple example showing the parameters (\tilde{a}, \tilde{b}) of the noise model (4) extracted from several images captured with a few devices of the same models. It is obvious that those estimated parameters overlap and, hence, lead to a



(a) Estimated camera parameters (\tilde{a}, \tilde{b}) proposed in [27]. Each result represents the parameters estimated from a single JPEG image.



(b) Estimated camera parameter $(\tilde{a}_k, \tilde{b}_k)$ proposed in this paper. Each result represents the parameters estimated from a single block.

Figure 2: Camera fingerprints comparison with several devices for each camera model. Natural JPEG images for *Nikon D70*, *Nikon D70s*, *Canon Ixus70* and *Nikon D200* are from Dresden dataset [35].

poor statistical test, whose detection performance is weak. In fact, even though in the test of [27] the average (\tilde{a}, \tilde{b}) is used for detection, the very similar (\tilde{a}, \tilde{b}) of each device unavoidably leads to unsatisfying detection results, see Section VI. Thus, in the following subsection IIb, it is proposed to improve the generalized signal-dependent noise model (4) and extract the novel device block fingerprints whose linear relationship is exposed for establishing the statistical tests.

b. Description of Improved Noise Model and Block Fingerprints

By challenging the assumption that all the noise corrupting all the pixels from a given digital image follows the same statistical distribution, modelled in (4), with the same distribution parameters, the improved signal dependent noise model proposes to take into account the non-uniformity of pixel's response by assuming that the parameters $(\tilde{a}_k, \tilde{b}_k)$ change for each pixel, or for each block of pixels for feasibility. This model is thus given

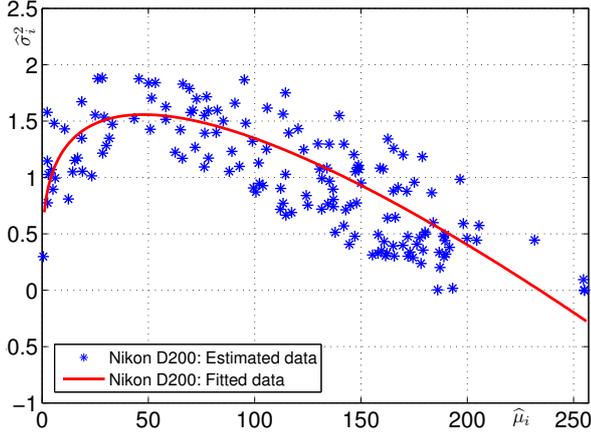


Figure 3: Scatter-plot of pixels' estimated expectation $\hat{\mu}_i$ and estimated variance $\hat{\sigma}_i^2$ in k -th block from several JPEG format images of *Nikon D200*.

by:

$$x_i \approx \mu_i + \frac{1}{\gamma} \mu_i^{1-\gamma} \eta_i \sim \mathcal{N}(\mu_i, \sigma_i^2) \quad (7)$$

$$\text{with } \sigma_i^2 = \frac{1}{\gamma^2} \mu_i^{2-2\gamma} (\bar{a}_k \mu_i^\gamma + \bar{b}_k) + \frac{\Delta^2}{12} \quad (8)$$

where $k \in \{1, \dots, K\}$ denotes the block index and block fingerprints (\bar{a}_k, \bar{b}_k) are estimated by using the same block, at the same location, from several JPEG images. Fig. 3 demonstrates on an example the accuracy of the proposed model and the relevance of the block fingerprints by showing the non-linear relationship (8) between pixel's expectation and variance in k -th block.

Let us describe the specific process of extracting the device block fingerprints from images as follows:

1. Let $\mathbf{X} = \{x_{i,j}\}$ be a single matrix representing a JPEG image of size $I \times J$, $i \in \{1, \dots, I\}$, $j \in \{1, \dots, J\}$ and let $\mathbf{X}^{(N)} = (\mathbf{X}_1, \dots, \mathbf{X}_N)$ be a set of N images. By using the BM3D denoising filter [36, 37], each image \mathbf{X}_n , $n \in \{1, \dots, N\}$ is first decomposed into two components: $\mathbf{X}_n^{\text{app}}$ an estimation of pixel's expectation $\mu_{i,j}$ and $\mathbf{X}_n^{\text{res}}$ the residual noise, that is: $\mathbf{X}_n^{\text{res}} = \mathbf{X}_n - \mathbf{X}_n^{\text{app}}$.
2. Due to the block artifact of JPEG image, it is proposed to decompose $\mathbf{X}_n^{\text{app}}$ into blocks of size 8×8 pixels, denoted $\{\mathbf{x}_{n,k}^{\text{app}}\}$ where $\mathbf{x}_{n,k}^{\text{app}} = \{x_{n,k,l}^{\text{app}}\}$, $k \in \{1, \dots, K\}$, $l \in \{1, \dots, 64\}$ with $K \approx \frac{I \times J}{64}$.
3. Similarly, we decompose $\mathbf{X}_n^{\text{res}}$ into blocks of size 8×8 pixels, denoted $\{\mathbf{x}_{n,k}^{\text{res}}\}$ where $\mathbf{x}_{n,k}^{\text{res}} = \{x_{n,k,l}^{\text{res}}\}$.
4. By computing the standard deviation of $\{\mathbf{x}_{n,k}^{\text{app}}\}$, it is proposed to set the threshold $\tau = 2$ to exclude non-homogeneous blocks in each image. In fact, those blocks correspond to highly textured areas for which estimations of pixel's expectation and of block variance are likely to be inaccurate.
5. For the k -th block of all N images, it is proposed to utilize the selected N_k with $N_k \leq N$ homogeneous blocks to estimate the expectation of $\mathbf{x}_{n,k}^{\text{app}}$ and the variance of $\mathbf{x}_{n,k}^{\text{res}}$.
6. Finally, based on the estimations of pixel's expectation and noise variance, the parameters of the proposed improved

signal-dependent noise model (\bar{a}_k, \bar{b}_k) are jointly estimated using a Least Square algorithm (LS) (see details in Appendix A). Those parameters are later used as fingerprints for the camera device identification problem cast with hypothesis testing theory, see Sections III-V.

Fig. 2 illustrates the comparison of the camera fingerprints extracted by using the algorithm of [27] and our proposed device block fingerprints. In Fig. 2a, each scatter point denotes parameters (\bar{a}, \bar{b}) estimated from a single image using the method proposed in [27]. By contrast, the proposed improved signal-dependent noise model parameters whose estimated values (\bar{a}_k, \bar{b}_k) are presented in Fig. 2b and used within the statistical test proposed in this paper as camera device fingerprints. Obviously, the proposed improved signal-dependent noise model is more relevant for extracting camera device fingerprints. This may be explained by the fact that this novel noise model takes into account the non-uniformity of pixels and hence estimates the proposed fingerprint by block; this leads to an increasing number of parameters for camera device fingerprints and, hence, helps distinguishing two camera devices of the same model.

c. Exposing Linear Relationship of Block Fingerprints

In fact, even though block fingerprints proposed in this paper, that are parameters (\bar{a}_k, \bar{b}_k) , allows the distinguishing of camera devices much more accurately than those proposed in [27], it is still very difficult to directly establish the optimal statistical test based on the improved signal-dependent noise model. This can be explained by the following main reasons: the variance of each image block is very difficult to estimate and, hence, can not always perfectly satisfies the improved signal-dependent noise model (7). Thus, we propose another solution for detection.

Based on the relationship between camera parameters (a, b) proposed in the the Poissonian-Gaussian noise model [28, 29]:

$$b = \omega^2 a^2 \quad (9)$$

where ω^2 is a depending parameter of the camera device, it is proposed to approximate Equation (9) due to the low variations of a by using the first order Taylor series at \bar{a} :

$$\begin{aligned} b &\approx \omega^2 \bar{a}^2 + 2\omega^2 \bar{a}(a - \bar{a}) \\ &= 2\omega^2 \bar{a}a - \omega^2 \bar{a}^2 \end{aligned} \quad (10)$$

where \bar{a} denotes the mean of a . Then, for clarity, let us rewrite Equation (10) as:

$$a \approx cb + d \quad (11)$$

where $c = 1/2\omega^2 \bar{a}$ and $d = \bar{a}/2$ approximately link the linear relationship between parameters (a, b) from the noise model (1) of RAW data. The white balancing, demosaicing and gamma correction operations do not change this linear relationship (see Equations (2) to (8)). Thus, the relationship is still valid for JPEG images. Fig. 4 illustrates the linear relationship between the parameters (\bar{a}_k, \bar{b}_k) for two *Canon Ixus70* devices. Based on the linear relationship of the parameters, let us replace \bar{a}_k by

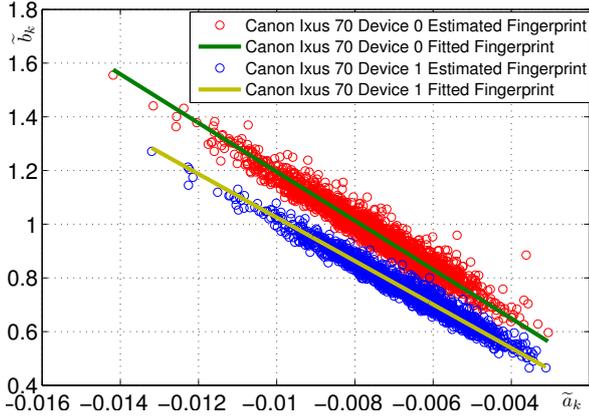


Figure 4: Illustration of linear relationship between the parameters (\bar{a}_k, \bar{b}_k) which are used as fingerprint for camera device identification; the results have been obtained following the process described in Section IIb fitted by using Least Square algorithm (LS).

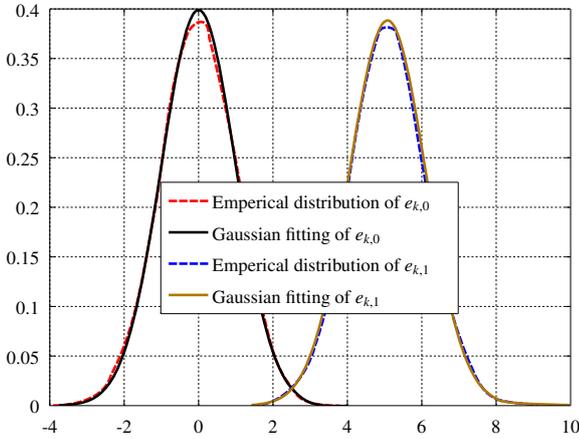


Figure 5: Empirical distribution of estimate error e_k .

$c\bar{b}_k + d$ and, hence, use $(c\bar{b}_k + d, \bar{b}_k)$ as the parameters that are used as fingerprints for camera device identification. As Fig. 4 illustrates, for images captured with the same device, the estimated parameters are well modelled using the proposed linear relation between \bar{a}_k and \bar{b}_k , which tends to confirm our assumption. Thus, let us define the estimate error by:

$$e_k = \bar{a}_k - (c\bar{b}_k + d). \quad (12)$$

where e_k denotes the estimate error in k -th block, c and d the linear model parameters which can be estimated by using LS algorithm. It should be noted that c and d are constant for the whole image. Then, the problem of identifying source individual device from JPEG images is solved by using the estimate error between camera fingerprint \bar{a}_k and its linear relation with \bar{b}_k . Because numerous sources of noise corrupt a digital image during its acquisition, and because within the proposed approach many parameters have to be estimated, and that each of those estimations generates errors, we can model accurately the statistical distribution of the error e_k , see (12), as the realization of a Gaussian random variable.

By using this statistical model for the error e_k , together with hypothesis testing theory, we first establish the optimal statis-

tical test for camera device identification and study its statistical performance. One can note that this optimal statistical test, namely the Likelihood Ratio Test (LRT), assumes that all distribution parameters are known, which are the framework adopted first in Section III. Then, for a practical application, it is proposed to investigate two GLRTs. In the first GLRT, it is proposed to study the case when the linear model parameters namely, (c, d) , of two devices and the camera fingerprints namely, (\bar{a}_k, \bar{b}_k) , of two devices are known, but the statistical distribution parameters of the error e_k are unknown. In such a case Section IV presents the proposed GLRT and especially studies the impact of distribution parameters on the statistical performance of this GLRT. In the second GLRT, presented in Section V, it is assumed that, under hypothesis \mathcal{H}_1 , neither the linear model parameters (c_1, d_1) nor the camera parameters $(\bar{a}_{k,1}, \bar{b}_{k,1})$ are known, the proposed GLRT deals with this practical case.

III. Likelihood Ratio Test for Two Simple Hypotheses Based on JPEG Format

a. Problem Statement

This section aims at presenting the optimal LRT and, more important, at studying its statistical performance; this statistical test is based on the error e_k , see (12). It is hardly possible to establish formally the statistical distribution of e_k ; however, because pixels within the same block share similar expectation, their distribution is well modelled by a Gaussian distribution and, hence, the estimation error e_k is also very accurately modelled by a Gaussian distribution. Hence in the rest of the paper, the error is modelled by a Gaussian distribution. This assumption has been confirmed empirically on a wide dataset such as shown in Fig. 5 which shows the comparison between the empirical error e_k and the Gaussian model under Hypothesis \mathcal{H}_0 . Each camera device j , $j \in \{0, 1\}$ is characterized by its linear parametric model with parameters (c_j, d_j) , which are obtained from the estimated camera device fingerprints $(\bar{a}_{k,j}, \bar{b}_{k,j})$. In fact, when testing an inspected image \mathbf{X} , two sets of errors e_k are known, that are $e_{k,0}$ based on parameters (c_0, d_0) ; $e_{k,1}$ based on parameters (c_1, d_1) . Hence, for simplicity, let us define $\mathbf{e}_k = (e_{k,0}, e_{k,1})$. Within the framework of the LRT, hence, the problem of identifying source camera device is reduced to a choice between the two following simple hypotheses:

$$\begin{cases} \mathcal{H}_0 : \{e_{k,0} = \bar{a}_k - (c_0\bar{b}_k + d_0) \sim \mathcal{N}(0, \sigma_0^2)\} \\ \mathcal{H}_1 : \{e_{k,1} = \bar{a}_k - (c_1\bar{b}_k + d_1) \sim \mathcal{N}(0, \sigma_1^2)\}, \end{cases} \quad (13)$$

where $\forall k = (1, \dots, K), (c_1, d_1) \neq (c_0, d_0), \sigma_0^2 \neq \sigma_1^2$.

Formally, a statistical test is a mapping $\delta : \mathbf{X} \rightarrow \delta(x)$, such that hypothesis \mathcal{H}_j is accepted if and only $\delta(\mathbf{X}) = \mathcal{H}_j$. For solving this statistical detection problem such as (13), it follows from the Neyman-Pearson lemma [38, Theorem 3.2.1] that the LRT is optimal in the sense described below. For definition, let

$$\mathcal{K}_{\alpha_\tau} = \{\delta : \mathbb{P}_0[\delta(\mathbf{X}) = \mathcal{H}_1] \leq \alpha_\tau\} \quad (14)$$

be the class of tests, solving problem (13), with an upper-bounded False Alarm Rate (FAR) α_τ . Here $\mathbb{P}_j[\cdot]$ is the probability under $\mathcal{H}_j, j \in \{0, 1\}$. Among all tests in $\mathcal{K}_{\alpha_\tau}$ the LRT is the most powerful test, which maximizes the detection power defined as:

$$\beta_\delta = \mathbb{P}_1[\delta(\mathbf{E}) = \mathcal{H}_1]. \quad (15)$$

In the following subsection, the LRT is first described in details and then its statistical performance is analytically established.

b. Optimal Detection Framework

For testing simple hypotheses the Neyman-Pearson Lemma [39, theorem 3.2.1] states that the LRT is the most powerful test in $\mathcal{K}_{\alpha_\tau}$ (14). Because of the statistical independence of pixels, and so is error \mathbf{e}_k , the LRT for camera device identification can be written:

$$\delta^{\text{lr}}(\mathbf{X}) = \begin{cases} \mathcal{H}_0 & \text{if } \Lambda^{\text{lr}}(\mathbf{E}) = \sum_{k=1}^K \Lambda^{\text{lr}}(\mathbf{e}_k) < \tau^{\text{lr}}, \\ \mathcal{H}_1 & \text{if } \Lambda^{\text{lr}}(\mathbf{E}) = \sum_{k=1}^K \Lambda^{\text{lr}}(\mathbf{e}_k) \geq \tau^{\text{lr}}, \end{cases} \quad (16)$$

where the decision threshold τ^{lr} is the solution of Equation $\mathbb{P}_{\mathcal{H}_0}[\Lambda^{\text{lr}}(\mathbf{E}) \geq \tau^{\text{lr}}] = \alpha_\tau$, to ensure that the FAR of the LRT equals α_τ . Based on the Gaussian distribution from Equation (13), it is proposed to define statistical parameters $\theta_0 = (0, \sigma_0^2)$ and $\theta_1 = (0, \sigma_1^2)$ under hypothesis \mathcal{H}_0 and \mathcal{H}_1 . Then the probability density function (pdf) is given as: \mathcal{P}_{θ_0} and \mathcal{P}_{θ_1} . Thus one can obtain the log Likelihood Ratio (LR) for one observation given by:

$$\Lambda^{\text{lr}}(\mathbf{e}_k) = \log \frac{\mathcal{P}_{\theta_1}[e_{k,1}]}{\mathcal{P}_{\theta_0}[e_{k,0}]}. \quad (17)$$

From the definition of (13), it is easy to rewrite the LR (17) as:

$$\Lambda^{\text{lr}}(\mathbf{e}_k) = \log \left(\frac{\sigma_0}{\sigma_1} \right) + \frac{1}{2} \left(\frac{e_{k,0}^2}{\sigma_0^2} - \frac{e_{k,1}^2}{\sigma_1^2} \right), \quad (18)$$

where the error $e_{k,j} = \tilde{a}_k - (c_j \tilde{b}_k + d_j)$, $j \in \{0, 1\}$ and block index $k \in \{1, \dots, K\}$.

c. Statistical Performance of LRT

The study of the proposed LRT is made easier thanks to some asymptotic theorems, which are relevant in this context as the number of blocks is chosen very large (each block is made of a small number of pixels). Let us denote $E_{\mathcal{H}_j}(\Lambda^{\text{lr}}(\mathbf{e}_k))$ and $V_{\mathcal{H}_j}(\Lambda^{\text{lr}}(\mathbf{e}_k))$ the expectation and the variance of the LR $\Lambda^{\text{lr}}(\mathbf{e}_k)$ under hypothesis $\mathcal{H}_j, j = \{0, 1\}$. Note that those moments can be analytically calculated, the details are provided in Appendix B. The Lindeberg's central limit theorem (CLT) [39, theorem 11.2.5] states that as the number of blocks K tends to infinity it

holds true that³:

$$\frac{\sum_{k=1}^K (\Lambda^{\text{lr}}(\mathbf{e}_k) - E_{\mathcal{H}_j}(\Lambda^{\text{lr}}(\mathbf{e}_k)))}{\left(\sum_{k=1}^K V_{\mathcal{H}_j}(\Lambda^{\text{lr}}(\mathbf{e}_k)) \right)^{1/2}} \xrightarrow{d} \mathcal{N}(0, 1), \quad j = \{0, 1\}, \quad (19)$$

where \xrightarrow{d} represents the convergence in distribution and $\mathcal{N}(0, 1)$ is the standard normal distribution with zero mean and unit variance. This theorem is of crucial interest to establish the statistical properties of the proposed test [40, 41, 42, 43, 44]. In fact, once the two first moments of the LR have been calculated analytically under hypothesis \mathcal{H}_0 , which again is detailed in Appendix B, one can normalize under hypothesis \mathcal{H}_0 the LR $\Lambda^{\text{lr}}(\mathbf{E})$ as follows:

$$\bar{\Lambda}^{\text{lr}}(\mathbf{E}) = \frac{\Lambda^{\text{lr}}(\mathbf{E}) - \sum_{k=1}^K E_{\mathcal{H}_0}(\Lambda^{\text{lr}}(\mathbf{e}_k))}{\left(\sum_{k=1}^K V_{\mathcal{H}_0}(\Lambda^{\text{lr}}(\mathbf{e}_k)) \right)^{1/2}}.$$

It is thus straightforward to define the normalized LRT with $\bar{\Lambda}^{\text{lr}}(\mathbf{E})$ by:

$$\bar{\delta}^{\text{lr}} = \begin{cases} \mathcal{H}_0 & \text{if } \bar{\Lambda}^{\text{lr}}(\mathbf{E}) < \bar{\tau}^{\text{lr}} \\ \mathcal{H}_1 & \text{if } \bar{\Lambda}^{\text{lr}}(\mathbf{E}) \geq \bar{\tau}^{\text{lr}}. \end{cases} \quad (20)$$

Thus, let us establish the statistical properties of the LRT (20).

Proposition 1. *Assuming that for the camera device identification problem as case within the two simple hypotheses (13), in which both parameters (c_j, d_j) and σ_j^2 are known, then for any $\alpha_\tau \in (0, 1)$ the decision threshold:*

$$\bar{\tau}^{\text{lr}} = \Phi^{-1}(1 - \alpha_\tau), \quad (21)$$

guarantees that the LRT is in the class $\mathcal{K}_{\alpha_\tau}$, see (14). Here Φ and Φ^{-1} respectively represent the cumulative distribution function (cdf) of the standard normal distribution and its inverse.

Proposition 2. *Assuming that for the camera device identification problem as case within the two simple hypotheses (13), in which both parameters (c_j, d_j) and σ_j^2 are known, for any decision threshold $\bar{\tau}^{\text{lr}}$, the power function associated with the proposed test $\bar{\delta}^{\text{lr}}$ (20) is given by:*

$$\beta_{\bar{\delta}^{\text{lr}}} = 1 - \Phi \left(\sqrt{\frac{v_0}{v_1}} \Phi^{-1}(1 - \alpha_\tau) + \frac{m_0 - m_1}{\sqrt{v_1}} \right). \quad (22)$$

where

$$m_j = \sum_{k=1}^K E_{\mathcal{H}_j}(\Lambda^{\text{lr}}(\mathbf{e}_k)) \quad (23)$$

$$v_j = \sum_{k=1}^K V_{\mathcal{H}_j}(\Lambda^{\text{lr}}(\mathbf{e}_k)), \quad j = \{0, 1\}. \quad (24)$$

³Note that we refer to the Lindeberg's CLT, whose conditions are easily verified in our case as the pixels have a bounded value.

Equations (21) and (22) emphasize the main advantage of normalizing the LR as described in relation (20): it allows to set any of threshold that guarantees a FAR independently from any distribution parameter. One can also note that, for practical application, it is considered in the present paper that a given image under investigation was captured with device 0. Hence the case “false alarm” occurs when an alarm is raised because the given image is authenticated as captured with another device 1. On the opposite, the “correct detection” which defines the power function, corresponds to the probability of accurately identifying an image captured from camera device 1.

IV. GLRT for JPEG Format with Knowing Camera Fingerprints and Linear Model Parameters

While the previous section studies the optimal LRT assuming that all the statistical parameters are known, this section proposes a less restrictive GLRT. It is indeed assumed in the present section that statistical parameters of the Gaussian distribution, see (13), are unknown. Hence it is proposed in this section to estimate those parameters and then, to design and study statically the GLRT which replaces the Gaussian distribution parameters with those estimates. Similar to the prior LRT, let us defines the GLRT $\widehat{\delta}_1^{\text{lr}}$ as follows:

$$\widehat{\delta}_1^{\text{lr}} = \begin{cases} \mathcal{H}_0 & \text{if } \widehat{\Lambda}_1^{\text{lr}}(\mathbf{E}) = \sum_{k=1}^K \widehat{\Lambda}_1^{\text{lr}}(\mathbf{e}_k) < \widehat{\tau}_1^{\text{lr}}, \\ \mathcal{H}_1 & \text{if } \widehat{\Lambda}_1^{\text{lr}}(\mathbf{E}) = \sum_{k=1}^K \widehat{\Lambda}_1^{\text{lr}}(\mathbf{e}_k) \geq \widehat{\tau}_1^{\text{lr}}. \end{cases} \quad (25)$$

$\widehat{\Lambda}_1^{\text{lr}}(\mathbf{e}_k)$ ensures that the proposed GLRT $\widehat{\delta}_1^{\text{lr}}$ is in the class $\mathcal{K}_{\alpha_\tau}$. The associated decision threshold $\widehat{\tau}_1^{\text{lr}}$ is defined as the solution of the following Equation:

$$\mathbb{P}_0[\widehat{\Lambda}_1^{\text{lr}}(\mathbf{E}) \geq \widehat{\tau}_1^{\text{lr}}] = \alpha_\tau. \quad (26)$$

Again, here α_τ is the prescribed probability of false alarm.

Again, the Generalized Likelihood Ratio (GLR) essentially consists in replacing the unknown parameters by the Maximum Likelihood (ML) estimation; hence, it follows from the definition of the LR (18), that the GLR is defined as:

$$\begin{aligned} \widehat{\Lambda}_1^{\text{lr}}(\mathbf{e}_k) &= \log \frac{\mathcal{P}_{\widehat{\theta}_1}[e_{k,1}]}{\mathcal{P}_{\widehat{\theta}_0}[e_{k,0}]} \\ &= \log \left(\frac{\widehat{\sigma}_0}{\widehat{\sigma}_1} \right) + \frac{1}{2} \left(\frac{e_{k,0}^2}{\widehat{\sigma}_0^2} - \frac{e_{k,1}^2}{\widehat{\sigma}_1^2} \right), \end{aligned} \quad (27)$$

where $e_{k,j} = \widetilde{a}_k - (c_j \widetilde{b}_k + d_j)$, with $j \in \{0, 1\}$, represents the estimate error and $(\widetilde{a}_k, \widetilde{b}_k)$ stand for the device fingerprint estimated from the inspected image \mathbf{X} . $\widehat{\theta}_0$ and $\widehat{\theta}_1$ denote respectively the estimates of statistical parameters θ_0 and θ_1 . It should be noted that in this GLR, the linear model parameters (c_j, d_j) and camera device fingerprints $(\widetilde{a}_{k,j}, \widetilde{b}_{k,j})$ are known. Our aim is

to estimate the Gaussian distribution parameter σ_j^2 . Using ML estimation, the variance is estimated by:

$$\widehat{\sigma}_j^2 = \frac{1}{K-1} \sum_{k=1}^K (\widetilde{a}_{k,j} - c_j \widetilde{b}_{k,j} - d_j)^2. \quad (28)$$

By invoking again the Lindeberg’s CLT [39, theorem 11.2.5] under hypothesis \mathcal{H}_j , $j \in \{0, 1\}$, immediately the statistical distribution of $\widehat{\Lambda}_1^{\text{lr}}(\mathbf{E})$ is given as:

$$\widehat{\Lambda}_1^{\text{lr}}(\mathbf{E}) \xrightarrow{d} \mathcal{N}(m_j^{(1)}, v_j^{(1)}) \quad (29)$$

where the expectation $m_j^{(1)}$ and the variance $v_j^{(1)}$ of the GLR $\widehat{\Lambda}_1^{\text{lr}}(\mathbf{E})$ are respectively defined by:

$$m_j^{(1)} = \sum_{k=1}^K E_{\mathcal{H}_j}(\widehat{\Lambda}_1^{\text{lr}}(\mathbf{e}_k)) \quad (30)$$

$$v_j^{(1)} = \sum_{k=1}^K V_{\mathcal{H}_j}(\widehat{\Lambda}_1^{\text{lr}}(\mathbf{e}_k)), \quad (31)$$

where the specific calculation of the expectation $E_{\mathcal{H}_j}(\widehat{\Lambda}_1^{\text{lr}}(\mathbf{e}_k))$ and the variance $V_{\mathcal{H}_j}(\widehat{\Lambda}_1^{\text{lr}}(\mathbf{e}_k))$ of $\widehat{\Lambda}_1^{\text{lr}}(\mathbf{e}_k)$ is expanded in Appendix C. Let us straightforwardly define the normalized GLR as follows:

$$\widehat{\Lambda}_1^{\star}(\mathbf{E}) = \frac{\widehat{\Lambda}_1^{\text{lr}}(\mathbf{E}) - m_j^{(1)}}{\sqrt{v_j^{(1)}}}. \quad (32)$$

It finally follows that the normalized GLRT, based on $\widehat{\Lambda}_1^{\star}(\mathbf{E})$, is given by:

$$\widehat{\delta}_1^{\star} = \begin{cases} \mathcal{H}_0 & \text{if } \widehat{\Lambda}_1^{\star}(\mathbf{E}) < \tau_1^{\star} \\ \mathcal{H}_1 & \text{if } \widehat{\Lambda}_1^{\star}(\mathbf{E}) \geq \tau_1^{\star}. \end{cases} \quad (33)$$

Hence, again, one can immediately establish the statistical properties of the GLRT (33).

Proposition 3. *Assuming that the pixels are modelled by the proposed improved signal-dependent noise model (8), when both linear model parameters (c_j, d_j) and camera device fingerprints $(\widetilde{a}_{k,j}, \widetilde{b}_{k,j})$ are known, and the unknown statistical parameter σ_j^2 is estimated as in (28), then any $\alpha_\tau \in (0, 1)$ the decision threshold of the proposed GLRT $\widehat{\delta}_1^{\star}$ is given by:*

$$\tau_1^{\star} = \Phi^{-1}(1 - \alpha_\tau). \quad (34)$$

Proposition 4. *Assuming that the pixels are modelled by the proposed improved signal-dependent noise model (8), when both linear model parameters (c_j, d_j) and camera device fingerprints $(\widetilde{a}_{k,j}, \widetilde{b}_{k,j})$ are known, and the unknown statistical parameter σ_j^2 is estimated as in (28), for any decision threshold τ_1^{\star} , the power function associated with test $\widehat{\delta}_1^{\star}$ (33) is given by*

$$\beta_{\widehat{\delta}_1^{\star}} = 1 - \Phi \left(\sqrt{\frac{v_0^{(1)}}{v_1^{(1)}}} \Phi^{-1}(1 - \alpha_\tau) + \frac{m_0^{(1)} - m_1^{(1)}}{\sqrt{v_1^{(1)}}} \right). \quad (35)$$

It is important to contact the optimal LRT presented in Section III that requires the knowledge of both noise model through camera device fingerprints $(\tilde{a}_{k,j}, \tilde{b}_{k,j})$ and the linear relation between those fingerprints characterized by (c_j, d_j) as well as the error \mathbf{e}_k statistical parameter, that is the expectation and the variance under both hypotheses. On the opposite, the practical GLRT proposed in this Section IV is essentially based on the estimation of noise model parameters, typically from a set of images from each device. Then the proposed GLRT consists of estimating the noise model parameters from given images, using those estimates to extract the error \mathbf{e}_k and testing, for each block, whether the error is more likely distributed according to what is expected under each hypothesis \mathcal{H}_0 or \mathcal{H}_1 . Besides, the impact of those estimations is taken into account in the calculation of the proposed GLRT statistical performance. In practice, before testing, we have known σ from BM3D denoising filter for decomposing an inspected image which controls the strength of denoising and γ which is the value of gamma correction.

V. GLRT for JPEG Format without Knowing Camera Fingerprints or Linear Model Parameters

While the GLRT proposed in Section IV can be used in many practical applications in which one has to decide if a given image was acquired either with device 0 or with device 1, it is also possible to imagine some practical case that does not fit within this framework. This would be typically the case if one has only access to a given device 0; and hence he may be interested to test whether a given image has been acquired with this known device, or with any other unknown device, for which the noise parameters are unknown. In such a practical context, the hypothesis \mathcal{H}_1 becomes composite. In this scenario, it is proposed to design a test which would allow us to identify whether or not the inspected JPEG image \mathbf{X} is taken by the camera device 0. Hence, it is important to note that the given image \mathbf{X} may have been acquired by any other unknown device characterized by noise parameters that may take any values. The proposed second GLRT, denoted as $\widehat{\delta}_2^{\text{lr}}$, if defined as follows:

$$\widehat{\delta}_2^{\text{lr}} = \begin{cases} \mathcal{H}_0 & \text{if } \widehat{\Lambda}_2^{\text{lr}}(\mathbf{E}) = \sum_{k=1}^K \widehat{\Lambda}_2^{\text{lr}}(\widehat{\mathbf{e}}_k) < \widehat{\tau}_2^{\text{lr}}, \\ \mathcal{H}_1 & \text{if } \widehat{\Lambda}_2^{\text{lr}}(\mathbf{E}) = \sum_{k=1}^K \widehat{\Lambda}_2^{\text{lr}}(\widehat{\mathbf{e}}_k) \geq \widehat{\tau}_2^{\text{lr}}, \end{cases} \quad (36)$$

where the decision threshold $\widehat{\tau}_2^{\text{lr}}$ is the solution of Equation

$$\mathbb{P}_0[\widehat{\Lambda}_2^{\text{lr}}(\mathbf{E}) \geq \widehat{\tau}_2^{\text{lr}}] = \alpha_\tau \quad (37)$$

to ensure that the GLRT is in the class $\mathcal{K}_{\alpha_\tau}$. The GLR $\widehat{\Lambda}_2^{\text{lr}}(\widehat{\mathbf{e}}_k)$ is defined by:

$$\widehat{\Lambda}_2^{\text{lr}}(\widehat{\mathbf{e}}_k) = \log \frac{\mathcal{P}_{\theta_1}[\widehat{e}_{k,1}]}{\mathcal{P}_{\theta_0}[e_{k,0}]} \quad (38)$$

$$= \log \left(\frac{\widehat{\sigma}_0}{\widehat{\sigma}_1} \right) + \frac{1}{2} \left(\frac{e_{k,0}^2}{\widehat{\sigma}_0^2} - \frac{\widehat{e}_{k,1}^2}{\widehat{\sigma}_1^2} \right), \quad (39)$$

where $\widehat{\mathbf{e}}_k$ consists of the estimate error $e_{k,0} = \tilde{a}_k - c_0 \tilde{b}_k - d_0$ and $\widehat{e}_{k,1} = \tilde{a}_k - \tilde{c}_1 \tilde{b}_k - \tilde{d}_1$. It should be noted that in this GLR, the linear model parameters (c_0, d_0) and camera device fingerprints $(\tilde{a}_{k,0}, \tilde{b}_{k,0})$ are known under hypothesis \mathcal{H}_0 ; (c_1, d_1) and $(\tilde{a}_{k,1}, \tilde{b}_{k,1})$ are unknown, which can be estimated by using several tested images. It is proposed to estimate parameters (c_1, d_1) by using LS algorithm, see Appendix A, and estimate $(\tilde{a}_{k,1}, \tilde{b}_{k,1})$ by using the proposed algorithm of block fingerprints estimation in subsection IIb. Then the variance $\widehat{\sigma}_j^2$ can be estimated by using Equation (28).

Once again one can use Lindeberg's CLT to establish the statistical distribution of the proposed normalized $\widehat{\Lambda}_2^{\star}(\mathbf{E})$, defined follows as:

$$\widehat{\Lambda}_2^{\star}(\mathbf{E}) = \frac{\widehat{\Lambda}_2^{\text{lr}}(\mathbf{E}) - m_j^{(2)}}{\sqrt{v_j^{(2)}}} \xrightarrow{d} \mathcal{N}(0, 1). \quad (40)$$

where the expectation $m_j^{(2)}$ and the variance $v_j^{(2)}$ of $\widehat{\Lambda}_2^{\text{lr}}(\mathbf{E})$ is defined by:

$$m_j^{(2)} = \sum_{k=1}^K E_{\mathcal{H}_j}(\widehat{\Lambda}_2^{\text{lr}}(\widehat{\mathbf{e}}_k)) \quad (41)$$

$$v_j^{(2)} = \sum_{k=1}^K V_{\mathcal{H}_j}(\widehat{\Lambda}_2^{\text{lr}}(\widehat{\mathbf{e}}_k)). \quad (42)$$

For clarity, the calculations of the expectation $E_{\mathcal{H}_j}(\widehat{\Lambda}_2^{\text{lr}}(\widehat{\mathbf{e}}_k))$ and the variance $V_{\mathcal{H}_j}(\widehat{\Lambda}_2^{\text{lr}}(\widehat{\mathbf{e}}_k))$ of $\widehat{\Lambda}_2^{\text{lr}}(\widehat{\mathbf{e}}_k)$ are detailed in Appendix D.

Finally, it is proposed to define the normalized GLRT with $\widehat{\Lambda}_2^{\star}(\mathbf{E})$ by:

$$\widehat{\delta}_2^{\star} = \begin{cases} \mathcal{H}_0 & \text{if } \widehat{\Lambda}_2^{\star}(\mathbf{E}) < \tau_2^{\star} \\ \mathcal{H}_1 & \text{if } \widehat{\Lambda}_2^{\star}(\mathbf{E}) \geq \tau_2^{\star}. \end{cases} \quad (43)$$

Again, it is proposed to establish the statistical properties of the GLRT (43) for testing a known camera device against any other unknown device.

Proposition 5. *Assuming that the pixels are modelled by the proposed improved signal-dependent noise model (8), when parameters (c_0, d_0) and $(\tilde{a}_{k,0}, \tilde{b}_{k,0})$ for device 0 are known, parameters (c_1, d_1) and $(\tilde{a}_{k,1}, \tilde{b}_{k,1})$ for device 1 are not known, and the unknown parameter σ_j^2 is estimated as in (28), then any $\alpha_\tau \in (0, 1)$ the decision threshold of the proposed GLRT $\widehat{\delta}_2^{\star}$ is given by:*

$$\tau_2^{\star} = \Phi^{-1}(1 - \alpha_\tau), \quad (44)$$

Proposition 6. *Assuming that the pixels are modelled by the proposed improved signal-dependent noise model (8), when parameters (c_0, d_0) and $(\tilde{a}_{k,0}, \tilde{b}_{k,0})$ for device 0 are known, parameters (c_1, d_1) and $(\tilde{a}_{k,1}, \tilde{b}_{k,1})$ for device 1 are not known, and the unknown parameter σ_j^2 is estimated as in (28), for any decision threshold $\widehat{\tau}_2^{\star}$, the power function associated with test $\widehat{\delta}_2^{\star}$ (43) is given by*

$$\beta_{\widehat{\delta}_2^{\star}} = 1 - \Phi \left(\sqrt{\frac{v_0^{(2)}}{v_1^{(2)}}} \Phi^{-1}(1 - \alpha_\tau) + \frac{m_0^{(2)} - m_1^{(2)}}{\sqrt{v_1^{(2)}}} \right). \quad (45)$$

The main advantages of the proposed GLRT $\widehat{\delta}_2^*$ are 1) that its statistical performance can be analytically established and, hence, one can easily compute the decision threshold that guarantees a prescribed FAR along with the ensuing power function and 2) thanks to the normalization the decision threshold only depends on the prescribed FAR α_τ . In practice, the parameters σ and γ for camera device 0 are both known; on the contrary, those two parameters for device 1 are estimated by using the algorithms proposed in [26] before applying source camera identification.

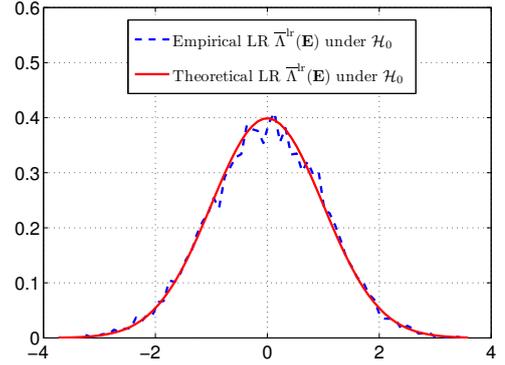
VI. Numerical Experiments

a. Results on Simulated Images for JPEG Format

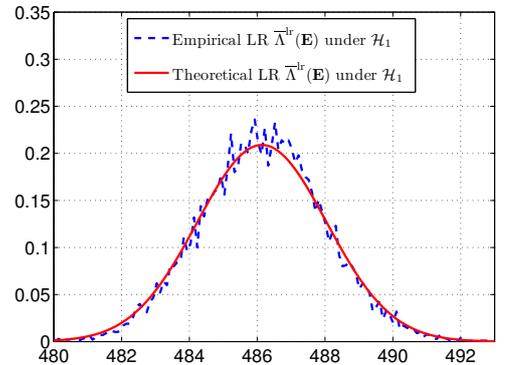
One of the main contributions of this paper is to show that hypothesis testing theory can be applied to design a statistical test with known statistical properties for identifying source camera device on simulated images.

To verify the sharpness of the theoretically established results, it is first proposed to use the Monte-Carlo simulation on a generated dataset. Camera device 0 is characterized by $(c_0, d_0) = (-0.0142, 0.0015)$ and device 1 characterized by $(c_1, d_1) = (-0.0132, 0.0018)$. Those values correspond to the estimated parameters for two different devices of camera model *Nikon D200* from the Dresden dataset [35]. With those settings, the statistical parameters of residuals $e_{k,0}$ and $e_{k,1}$ (see Equation (13)) following the Gaussian distribution are also known using estimates from Dresden dataset images. Here, the simulation reduces to a simple test between two Gaussian random variables with known and different variances. All the required parameters have been detailed, thus, one can now build two sets of random variables by repeating 10000 simulation with those distribution parameters. Then, the first set of 10000 simulated images from camera device 0 consisting of 5000 realization of random variables $(\widetilde{a}_0, \widetilde{b}_0)$ for each image; the second set of 10000 simulated images from device 1 consisting of 5000 realization of random variables $(\widetilde{a}_1, \widetilde{b}_1)$ for each image. Under hypothesis \mathcal{H}_0 and \mathcal{H}_1 respectively, Fig. 6 presents the comparison between empirical and theoretical distribution of the optimal LR $\overline{\Lambda}^{\text{lr}}(\mathbf{E})$. Under hypothesis \mathcal{H}_0 , the empirical distribution of the LR $\overline{\Lambda}^{\text{lr}}(\mathbf{E})$ approximately follows the standard normal distribution with zero mean and unit variance, which directly verifies the accuracy of the theoretically established statistical performance for the proposed LRT (20). Similarly, under hypothesis \mathcal{H}_1 , the empirical and theoretical distribution of the optimal LR $\overline{\Lambda}^{\text{lr}}(\mathbf{E})$ are nearly overlapped, which also verifies the correctness of the established statistical performance.

Another contribution of the present paper lies in the design of a statistical test that can warrant the prescribed FAR. Thus, it is proposed to compare the empirical and theoretical FAR α_τ of the optimal LR $\overline{\Lambda}^{\text{lr}}(\mathbf{E})$ as a function of the decision threshold $\overline{\tau}^{\text{lr}}$. The results of this comparison are presented in Fig. 7. This figure emphasizes that the proposed LRT (20) has in practice the ability of guaranteeing the prescribed FAR. In some cases ($\overline{\tau}^{\text{lr}} \geq 3$), it should be noted that the slight differences of two



(a) Under hypothesis \mathcal{H}_0 between the empirical distribution of the LR $\overline{\Lambda}^{\text{lr}}(\mathbf{E})$ and its theoretically established distribution (here Gaussian distribution with zero mean and unit variance).



(b) Under hypothesis \mathcal{H}_1 between the empirical distribution of the LR $\overline{\Lambda}^{\text{lr}}(\mathbf{E})$ and its theoretically established distribution (here Gaussian distribution with expectation $\frac{m_1 - m_0}{\sqrt{v_0}}$ and variance $\frac{v_1}{v_0}$, see subsection IIIc).

Figure 6: Comparison between empirical and theoretical distribution of $\overline{\Lambda}^{\text{lr}}(\mathbf{E})$.

curves are due to the inaccuracy of the CLT which can hardly model the tails of the distribution with accuracy.

b. Results on Real Images for JPEG Format

Another major contribution of this paper is to design the detectors with estimated parameters to identify source camera device in the practical cases. Therefore, it is proposed to verify on numerical experiments, using a real image dataset, the accuracy of both the GLRT $\widehat{\delta}_1^*$ (33), that aims at distinguishing between two camera devices whose parameters are known, as well as the accuracy of the second GLRT $\widehat{\delta}_2^*$ (43), whose goal is to distinguish between a known camera device and any other unknown device with unknown parameters.

To this end, it is proposed in this paper to exploit the reference images from Dresden dataset [35] for showing the accuracy of the proposed GLR tests as well as for comparison with the prior-art detectors. Note that the images from Dresden dataset are colored image and that in our tests, only the red color channel is selected. Besides, all the images are acquired with the highest available JPEG quality setting and maximum available resolution. Table 1 gives the specific number of each device in our experiments. It also displays the comparison of the

Table 1: Images statistic from Dresden dataset with mean estimated parameters (std: standard deviation)

Camera Model & Device	Alias	σ	γ	Resolution	No. images
<i>Nikon D200</i> # 0	N_D200_0	3.4000	0.8464	3872×2592	372
<i>Nikon D200</i> # 1	N_D200_1	3.3294	0.8591	3872×2592	380
<i>Nikon D70</i> # 0	N_D70_0	2.2365	0.7494	3008×2000	180
<i>Nikon D70</i> # 1	N_D70_1	2.1823	0.7024	3008×2000	189
<i>Nikon D70s</i> # 0	N_D70s_0	2.2322	0.7001	3008×2000	178
<i>Nikon D70s</i> # 1	N_D70s_1	2.0998	0.7506	3008×2000	189
<i>Canon Ixus70</i> # 0	C_I70_0	4.3573	0.8596	3072×2304	187
<i>Canon Ixus70</i> # 1	C_I70_1	4.0456	0.8425	3072×2304	194
<i>Canon Ixus70</i> # 2	C_I70_2	4.5535	0.8819	3072×2304	186
Σ 11	Σ 11	std: 1.0025	std: 0.0726	Σ 3	Σ 2055

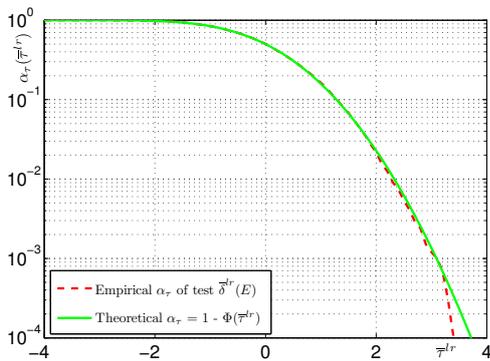


Figure 7: Comparison between the theoretical FAR α_τ and the empirical results, plotted as a function of the threshold τ^r .

estimated parameters where σ denotes the parameter of BM3D denoising filter [36, 37] which, roughly speaking, determines denoising strength, and the presented value of parameter γ represents the value of gamma correction mean parameter in the post-processing of digital imaging. It should be noted that in this paper, these two parameters are not the device fingerprint, and that they can be estimated accurately, for instance, by using the algorithm proposed in [26] before applying the identification step. In other words, parameters (σ, γ) are assumed to be known in our proposed practical test (33) and unknown in the test (43).

For the practical application of the proposed GLRT, the dataset available for each image is split into two subsets: one “*Learning Subset*” and one “*Testing Subset*”; those subsets are disjoint as none of images is used in both subsets. Images of “*Learning Subset*” are used to extract device fingerprints from each device; images of “*Testing Subset*” are used to identify the origin of a given image. The number of “*Learning Subset*” is set as 100 where the images are randomly selected with high quality factor ($QF \geq 70$). Besides, unlike the fingerprints of RAW images [29, 30], we have found that the ISO sensitivity has no impact on the proposed fingerprints and, hence, on the camera device identification methodology.

Experiments are realized on a large dataset to verify the sharpness of the proposed algorithms. It should be noted that in the proposed tests GLRT $\widehat{\delta}_1^*$ and $\widehat{\delta}_2^*$, we use multiple JPEG im-

ages together to identify the camera device. Then it is proposed to divide the “*Testing Subset*” into several overlapping groups of n images for each group, here of course $n < N$ with N the total number of images in the “*Testing Subset*”. If the value of n is set very small, we could lose the accuracy of the estimation, since our parameters estimation is based on LS algorithm which needs the large number of data (see Appendix A). To select the optimal number of each group, let us do the test with increasing the number n (see Table 2). Based on the detection power of the proposed test, therefore, it is proposed to set $n = 40$ in our test GLRT $\widehat{\delta}_1^*$ and $\widehat{\delta}_2^*$, which can guarantee the high detection power while decreasing the probability of missed detection.

First, let us highlight the relevance of the proposed GLRT $\widehat{\delta}_1^*$. In this scenario, the camera fingerprints $(\widetilde{a}_{k,j}, \widetilde{b}_{k,j})$ are known (or estimated) as well as the linear model parameters (c_j, d_j) , $j \in \{0, 1\}$ for all devices from the same camera model. Our goal is to identify the inspected images captured with device 0, corresponding here to hypothesis \mathcal{H}_0 , or with camera device 1, corresponding to hypothesis \mathcal{H}_1 . In [27], *Test one* is designed for identifying the inspected image from camera device 0 or 1, that deals with the same case as the proposed GLRT $\widehat{\delta}_1^*$. Thus, it is reasonable that let us compare the performance of our proposed test $\widehat{\delta}_1^*$ with *Test one* of [27]. As Table 3 illustrates, at the given very low FAR ($\alpha_\tau = 0.01$), our proposed method obviously outperforms *Test one*. The detection power of our proposed test $\widehat{\delta}_1^*$ remains 1 for each case. On the contrary, the average detection power of *Test One* only achieve 0.05. Based on the generalized signal-dependent noise model in [27], *Test one* can only identify two camera devices from different models. By contrast, our proposed test $\widehat{\delta}_1^*$ can identify source camera device of the same model.

Next, let us study the detection performance of the test $\widehat{\delta}_2^*$. In this scenario, we can not obtain the camera fingerprints $(\widetilde{a}_{k,1}, \widetilde{b}_{k,1})$ and linear model parameters (c_1, d_1) , $j \in \{0, 1\}$ for device 1, which are estimated from the group of “*Testing Subset*”. The goal of the test $\widehat{\delta}_2^*$ is to identify the inspected JPEG images acquired by device 0 or any other one. In [27], *Test two* is designed for identifying the inspected image from camera device 0 or any other one as well. Besides, the state-of-the-art detector of [19], based on the reference PRNU noise as camera finger-

Table 2: Detection performance comparison of the test $\widehat{\delta}_2^*$ at the given FAR $\alpha_\tau = 0.02$ with increasing the number n of tested images for each group, assuming that two devices 0 and 1 from the same camera model.

Camera Device	n = 10	n = 20	n = 30	n = 40
N_D200_0 vs N_D200_1	0.10	0.95	0.97	1.00
N_D70_0 vs N_D70_1	0.05	1.00	1.00	1.00
N_D70s_0 vs N_D70s_1	0.12	0.98	1.00	1.00
C_I70_0 vs C_I70_1	0.00	0.98	1.00	1.00
C_I70_0 vs C_I70_2	0.00	0.96	1.00	1.00
C_I70_1 vs C_I70_2	0.00	1.00	1.00	1.00
Average	0.05	0.98	1.00	1.00

Table 3: Detection performance comparison at the given FAR $\alpha_\tau = 0.01$, assuming that two devices 0 and 1 from the same camera model. This test is corresponding to the case of the GLRT (33).

Camera Device	Proposed Test $\widehat{\delta}_1^*$	<i>Test one</i> of [27]
N_D200_0 vs N_D200_1	1.00	0.02
N_D70_0 vs N_D70_1	1.00	0.01
N_D70s_0 vs N_D70s_1	1.00	0.03
C_I70_0 vs C_I70_1	1.00	0.16
C_I70_0 vs C_I70_2	1.00	0.03
C_I70_1 vs C_I70_2	1.00	0.07
Average	1.00	0.05

Table 4: Detection performance comparison at the given FAR $\alpha_\tau = 0.01$, assuming that two devices 0 and 1 from the same camera model. This test is corresponding to the case of the GLRT (43).

Camera Device	Proposed Test $\widehat{\delta}_2^*$	<i>Test two</i> of [27]	Test [19]
N_D200_0 vs N_D200_1	1.00	0.01	1.00
N_D70_0 vs N_D70_1	1.00	0.01	1.00
N_D70s_0 vs N_D70s_1	1.00	0.01	1.00
C_I70_0 vs C_I70_1	1.00	0.02	1.00
C_I70_0 vs C_I70_2	1.00	0.03	1.00
C_I70_1 vs C_I70_2	1.00	0.07	1.00
Average	1.00	0.03	1.00

prints, can also deal with this scenario. Then let us compare the detection perform of those three detectors for identifying source camera device from the same model. As Table 4 demonstrates, at the given very low FAR ($\alpha_\tau = 0.01$), the test $\widehat{\delta}_2^*$ is approximately perfect and exhibits roughly the detection power as the state-of-the-art detector of [19], as well as largely outperforming *Test two* whose detection power can only achieve 0.03. Besides, it should be noted that in the comparison with [19] the performance of both methods evolved very similarly with the increasing number of training images used to estimate noise parameters.

To compare with two tests of [27], our proposed two detectors largely outperform the algorithms from [27] (see Table 3

and 4). In fact, the camera fingerprints extracted from JPEG images in [27] are used for designing the detectors (for instance, *Test one* and *Test two*) which can only identify source camera model. Hence, the authors of [27] did not extract very distinguishable camera fingerprints, which is incapable of identifying source camera devices from the same model. Moreover, the camera fingerprints of [27] are extracted from one single JPEG image; our proposed fingerprints are extracted from multiple images. Therefore, our proposed fingerprints performs more distinguishable than the prior-art fingerprints. However, the limitation of our proposed detector is that the camera fingerprints are extracted from multiple images, which is not suitable for identifying one single JPEG image.

To extend the evaluation of our proposed detector, it is proposed to compare the performance of three instances from the same camera model *Canon Ixus70*. In this case, we detect if the inspected images acquired from the individual camera device is the same as the images under \mathcal{H}_0 . As Table 5 illustrated, even the extracted fingerprints from three instances are very close, the individual camera device can still be distinguished by using our proposed detector. It should be noted that the detection accuracy is denoted as the ratio of the number of inspected images selected as images under \mathcal{H}_0 to the total number of inspected images.

VII. Conclusion and Discussion

This paper studies the problem of identifying source camera device for JPEG images. In the most literature, PRNU-based detectors nearly dominate the research community of source camera identification. In this context, each camera device is characterized by its linear parametric model with parameters (c, d) , which are obtained from the estimated camera device fingerprints $(\tilde{a}_k, \tilde{b}_k)$ describing the proposed improved signal-dependent noise model. Then the problem of camera device identification is cast in the framework of hypothesis testing theory. Assuming that all the parameters are perfectly known, the statistical performance of the LRT is analytically established. In the practical cases, based on the estimated parameters, our designed two GLRTs perform very well at the prescribed FAR. Numerical simulations have showed positive results of the proposed method making a solid first proof of concept that still needs to be extended to get it up to the level of PRNU-based methods. The method proposed in this paper is however an interesting alternative as it is based on a very different approach.

Besides, the problem of multi-hypotheses can be addressed by modifying the criterion of optimality and designing a multi-hypotheses test for this criterion. Once the statistical distribution of hypotheses are known (cameras parameters) a mini-max approach can be used for instance to minimize the worst miss-classification probability. The multi-hypotheses testing framework can also be used to reject both hypothesis \mathcal{H}_0 and \mathcal{H}_1 , see [45, 46].

The main limitation of our proposed tests is that since the camera fingerprints $(\tilde{a}_k, \tilde{b}_k)$ are extracted based on multiple images, our designed tests can only deal with the case of testing a set of images (at least 40) together. In this scenario, if a forensic investigator aims at identifying the origin of the only one inspected image, our proposed tests will be invalid.

VIII. Acknowledgments

This work is funded by the China Scholarship Council (CSC) program and the region Champagne-Ardenne, IDENT project.

Appendix A. Block Fingerprints $(\tilde{a}_k, \tilde{b}_k)$ Estimation

In this appendix, it is proposed to detail the estimation of $(\tilde{a}_k, \tilde{b}_k)$ by using Least Square algorithm (LS). Then let us transform Equation (8) to:

$$p_{k,n} = q_{k,n}\tilde{a}_k + \tilde{b}_k \quad (\text{A.1})$$

where $p_{k,n} = (\sigma_{k,n}^2 - \frac{\Delta}{12}) \frac{\gamma}{\mu_{k,n}^{2-2\gamma}}$ where γ can be previously estimated using the algorithm proposed in [26] and $q_{k,n} = \mu_{k,n}$, the block index $k \in \{1, \dots, K\}$, the image index $n \in \{1, \dots, N\}$ with N denoting the number of JPEG images. Moreover, one can express Equation (A.1) in the form of vector like:

$$\mathbf{P} = \mathbf{Q} \cdot \begin{pmatrix} \tilde{a}_k \\ \tilde{b}_k \end{pmatrix} \quad (\text{A.2})$$

$$\text{where } \mathbf{P} = \begin{pmatrix} p_{k,1} \\ \vdots \\ p_{k,N} \end{pmatrix}, \mathbf{Q} = \begin{pmatrix} q_{k,1} & 1 \\ \vdots & \vdots \\ q_{k,N} & 1 \end{pmatrix}$$

Then, we can estimate the camera device fingerprints by:

$$\begin{pmatrix} \tilde{a}_k \\ \tilde{b}_k \end{pmatrix} = (\mathbf{Q}^T \mathbf{Q})^{-1} \mathbf{Q}^T \mathbf{P} \quad (\text{A.3})$$

where \mathbf{Q}^T denotes the transpose of \mathbf{Q} .

Appendix B. Expectation and Variance of LR $\Lambda^{\text{lr}}(\mathbf{e}_k)$

In this appendix, let us discuss the expectation $E_{\mathcal{H}_j}(\Lambda^{\text{lr}}(\mathbf{e}_k))$ and the variance $V_{\mathcal{H}_j}(\Lambda^{\text{lr}}(\mathbf{e}_k))$ of the LR $\Lambda^{\text{lr}}(\mathbf{e}_k)$ under hypothesis $\mathcal{H}_j, j = \{0, 1\}$. First, under hypothesis \mathcal{H}_0 with camera device 0 fingerprints $(\tilde{a}_{k,0}, \tilde{b}_{k,0})$, we can immediately obtain the following Gaussian distribution:

$$\begin{cases} e_{k,0} = \tilde{a}_{k,0} - (c_0 \tilde{b}_{k,0} + d_0) \sim \mathcal{N}(0, \sigma_0^2), \\ e_{k,1} = \tilde{a}_{k,0} - (c_1 \tilde{b}_{k,0} + d_1) \sim \mathcal{N}(\mu_{0,1}, \sigma_{0,1}^2). \end{cases} \quad (\text{B.1})$$

By putting the expression of the distribution from $e_{k,0}$ and $e_{k,1}$ respectively (B.1) into the expression LR $\Lambda^{\text{lr}}(\mathbf{e}_k)$ (18), the expectation $E_{\mathcal{H}_0}(\Lambda^{\text{lr}}(\mathbf{e}_k))$ and the variance $V_{\mathcal{H}_0}(\Lambda^{\text{lr}}(\mathbf{e}_k))$ can be calculated by:

$$\begin{aligned} E_{\mathcal{H}_0}(\Lambda^{\text{lr}}(\mathbf{e}_k)) &= \log\left(\frac{\sigma_0}{\sigma_1}\right) \\ &+ \frac{1}{2} \left(E_{\mathcal{H}_0} \left(\frac{e_{k,0}^2}{\sigma_0^2} \right) - E_{\mathcal{H}_0} \left(\frac{e_{k,1}^2}{\sigma_1^2} \right) \right) \\ &= \log\left(\frac{\sigma_0}{\sigma_1}\right) + \frac{1}{2} \left(1 - \frac{\sigma_{0,1}^2 + \mu_{0,1}^2}{\sigma_1^2} \right) \end{aligned} \quad (\text{B.2})$$

$$\begin{aligned} V_{\mathcal{H}_0}(\Lambda^{\text{lr}}(\mathbf{e}_k)) &= \frac{1}{4} \left(V_{\mathcal{H}_0} \left(\frac{e_{k,0}^2}{\sigma_0^2} \right) + V_{\mathcal{H}_0} \left(\frac{e_{k,1}^2}{\sigma_1^2} \right) \right) \\ &= \frac{1}{4} \left(2 + \frac{V_{\mathcal{H}_0}(e_{k,1}^2)}{\sigma_1^4} \right) \\ &= \frac{1}{2} \left(1 + \frac{\sigma_{0,1}^4 - 2\mu_{0,1}^2 \sigma_{0,1}^2}{\sigma_1^4} \right). \end{aligned} \quad (\text{B.3})$$

Table 5: Detection performance comparison at the given FAR $\alpha_\tau = 0.01$, assuming that two devices 0 and 1 from the same camera model. This test is corresponding to the case of the GLRT (43).

\mathcal{H}_0 Inspected	C_I70_0	C_I70_1	C_I70_2
C_I70_0	1.00	0	0
C_I70_1	0	1.00	0
C_I70_2	0	0	1.00

Similarly, under hypothesis \mathcal{H}_1 with camera device 1 fingerprints $(\tilde{a}_{k,1}, \tilde{b}_{k,1})$, we can immediately obtain the following Gaussian distribution:

$$\begin{cases} e_{k,0} = \tilde{a}_{k,1} - (c_0 \tilde{b}_{k,1} + d_0) \sim \mathcal{N}(\mu_{1,0}, \sigma_{1,0}^2), \\ e_{k,1} = \tilde{a}_{k,1} - (c_1 \tilde{b}_{k,1} + d_1) \sim \mathcal{N}(0, \sigma_1^2). \end{cases} \quad (\text{B.4})$$

The expectation $E_{\mathcal{H}_1}(\Lambda^{\text{lr}}(\mathbf{e}_k))$ and the variance $V_{\mathcal{H}_1}(\Lambda^{\text{lr}}(\mathbf{e}_k))$ can be expressed by:

$$E_{\mathcal{H}_1}(\Lambda^{\text{lr}}(\mathbf{e}_k)) = \log\left(\frac{\sigma_0}{\sigma_1}\right) + \frac{1}{2} \left(\frac{\sigma_{1,0}^2 + \mu_{1,0}^2}{\sigma_0^2} - 1 \right) \quad (\text{B.5})$$

$$\begin{aligned} V_{\mathcal{H}_1}(\Lambda^{\text{lr}}(\mathbf{e}_k)) &= \frac{1}{4} \left(2 + \frac{V_{\mathcal{H}_1}(e_{k,0}^2)}{\sigma_0^4} \right) \\ &= \frac{1}{2} \left(\frac{\sigma_{1,0}^4 - 2\mu_{1,0}^2 \sigma_{1,0}^2}{\sigma_0^4} + 1 \right). \end{aligned} \quad (\text{B.6})$$

Appendix C. Expectation and Variance of GLR $\widehat{\Lambda}_1^{\text{lr}}(\mathbf{e}_k)$

In this practical GLRT (33), we never know the statistical parameter of the errors \mathbf{e}_k , that is variance value. Thus, let us assume that the ML estimated variance $\widehat{\sigma}_j^2$ can be obtained by using Equation (28). In addition, under hypothesis \mathcal{H}_j of the GLRT $\widehat{\delta}_j^{\text{lr}}$, camera parameters $(\tilde{a}_{k,j}, \tilde{b}_{k,j})$ and linear model parameters (c_j, d_j) are perfectly known. Therefore, for simplicity, the expectation and variance of $\widehat{\Lambda}_1^{\text{lr}}(\mathbf{e}_k)$ can be obtained by putting $(\tilde{a}_{k,j}, \tilde{b}_{k,j})$ and $\widehat{\sigma}_j^2$ into Equations (B.2), (B.3), (B.5), (B.6), which are expressed by:

$$E_{\mathcal{H}_0}(\widehat{\Lambda}_1^{\text{lr}}(\mathbf{e}_k)) = \log\left(\frac{\widehat{\sigma}_0}{\widehat{\sigma}_1}\right) + \frac{1}{2} \left(1 - \frac{\widehat{\sigma}_{0,1}^2 + \widehat{\mu}_{0,1}^2}{\widehat{\sigma}_1^2} \right) \quad (\text{C.1})$$

$$V_{\mathcal{H}_0}(\widehat{\Lambda}_1^{\text{lr}}(\mathbf{e}_k)) = \frac{1}{4} \left(2 + \frac{V_{\mathcal{H}_0}(e_{k,1}^2)}{\widehat{\sigma}_1^4} \right) \quad (\text{C.2})$$

where under hypothesis \mathcal{H}_0 , $\widehat{\mu}_{0,1} = \frac{1}{K} \sum_{k=1}^K (\tilde{a}_{k,0} - c_1 \tilde{b}_{k,0} - d_1)$ and $\widehat{\sigma}_{0,1}^2 = \frac{1}{K-1} \sum_{k=1}^K (\tilde{a}_{k,0} - c_1 \tilde{b}_{k,0} - d_1 - \widehat{\mu}_{0,1})^2$ and $V_{\mathcal{H}_0}(e_{k,1}^2) = \frac{1}{K-1} \sum_{k=1}^K ((\tilde{a}_{k,0} - c_1 \tilde{b}_{k,0} - d_1)^2 - \frac{1}{K} \sum_{k=1}^K (\tilde{a}_{k,0} - c_1 \tilde{b}_{k,0} - d_1)^2)^2$;

$$E_{\mathcal{H}_1}(\widehat{\Lambda}_1^{\text{lr}}(\mathbf{e}_k)) = \log\left(\frac{\widehat{\sigma}_0}{\widehat{\sigma}_1}\right) + \frac{1}{2} \left(\frac{\widehat{\sigma}_{1,0}^2 + \widehat{\mu}_{1,0}^2}{\widehat{\sigma}_0^2} - 1 \right) \quad (\text{C.3})$$

$$V_{\mathcal{H}_1}(\widehat{\Lambda}_1^{\text{lr}}(\mathbf{e}_k)) = \frac{1}{4} \left(\frac{V_{\mathcal{H}_1}(e_{k,0}^2)}{\widehat{\sigma}_0^4} + 2 \right), \quad (\text{C.4})$$

where under hypothesis \mathcal{H}_1 , $\widehat{\mu}_{1,0} = \frac{1}{K} \sum_{k=1}^K (\tilde{a}_{k,1} - c_0 \tilde{b}_{k,1} - d_0)$ and $\widehat{\sigma}_{1,0}^2 = \frac{1}{K-1} \sum_{k=1}^K (\tilde{a}_{k,1} - c_0 \tilde{b}_{k,1} - d_0 - \widehat{\mu}_{1,0})^2$ and $V_{\mathcal{H}_1}(e_{k,0}^2) = \frac{1}{K-1} \sum_{k=1}^K ((\tilde{a}_{k,1} - c_0 \tilde{b}_{k,1} - d_0)^2 - \frac{1}{K} \sum_{k=1}^K (\tilde{a}_{k,1} - c_0 \tilde{b}_{k,1} - d_0)^2)^2$.

Appendix D. Expectation and Variance of GLR $\widehat{\Lambda}_2^{\text{lr}}(\mathbf{e}_k)$

In the case of GLRT (43), the unknown parameters (c_1, d_1) and $(\tilde{a}_{k,1}, \tilde{b}_{k,1})$ can be estimated by using several tested images. Let us estimate parameters (c_1, d_1) by using LS algorithm, see Appendix A, and estimate $(\tilde{a}_{k,1}, \tilde{b}_{k,1})$ by using the proposed algorithm of block fingerprints estimation in subsection IIb. Then, for clarity, it is proposed to replace the unknown parameters by their corresponding estimates and put them into the calculations of expectation and variance of GLR $\widehat{\Lambda}_1^{\text{lr}}(\mathbf{e}_k)$, see Equation (C.1)~(C.4). Finally, it is very easy to obtain the expectation and variance of GLR $\widehat{\Lambda}_2^{\text{lr}}(\mathbf{e}_k)$.

References

- [1] H. Farid, Image forgery detection, *Signal Processing Magazine*, IEEE 26 (2) (2009) 16–25.
- [2] H. T. Sencar, N. Memon, *Digital image forensics: There is more to a picture than meets the eye*, Springer, 2012.
- [3] M. C. Stamm, M. Wu, K. Liu, *Information forensics: An overview of the first decade*, *IEEE Access* 1 (1) (2013) 167–200.
- [4] J. Nakamura, *Image sensors and signal processing for digital still cameras*, CRC press, 2005.
- [5] R. Ramanath, W. E. Snyder, Y. Yoo, M. S. Drew, Color image processing pipeline, *Signal Processing Magazine*, IEEE 22 (1) (2005) 34–43.
- [6] M. Kirchner, T. Gloe, Forensic camera model identification, in: A. Ho, S. Li (Eds.), *Handbook of Digital Forensics of Multimedia Data and Devices*, Chapter 9, John Wiley & Sons Ltd, 2004.
- [7] K. San Choi, E. Y. Lam, K. K. Wong, Source camera identification using footprints from lens aberration, in: *Electronic Imaging 2006*, International Society for Optics and Photonics, 2006, pp. 60690J–60690J.
- [8] S. Bayram, H. T. Sencar, N. Memon, Efficient techniques for sensor fingerprint matching in large image and video databases (2010). doi:10.1117/12.845737. URL <http://dx.doi.org/10.1117/12.845737>
- [9] M. Goljan, J. Fridrich, Estimation of lens distortion correction from single images (2014). doi:10.1117/12.2036804. URL <http://dx.doi.org/10.1117/12.2036804>
- [10] Z. Deng, A. Gijssenij, J. Zhang, Source camera identification using auto-white balance approximation, in: *Computer Vision (ICCV)*, 2011 IEEE International Conference on, IEEE, 2011, pp. 57–64.
- [11] S. Bayram, H. Sencar, N. Memon, I. Avciabas, Source camera identification based on cfa interpolation, in: *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, Vol. 3, IEEE, 2005, pp. III–69.
- [12] H. Cao, A. C. Kot, Accurate detection of demosaicing regularity for digital image forensics, *Information Forensics and Security*, *IEEE Transactions on* 4 (4) (2009) 899–910.
- [13] M. Kharrazi, H. T. Sencar, N. Memon, Blind source camera identification, in: *Image Processing, 2004. ICIP'04. 2004 International Conference on*, Vol. 1, IEEE, 2004, pp. 709–712.

- [14] A. Swaminathan, M. Wu, K. R. Liu, Nonintrusive component forensics of visual sensors using output images, *Information Forensics and Security, IEEE Transactions on* 2 (1) (2007) 91–106.
- [15] K. S. Choi, E. Y. Lam, K. K. Wong, Source camera identification by jpeg compression statistics for image forensics, in: *TENCON 2006. 2006 IEEE Region 10 Conference, IEEE, 2006*, pp. 1–4.
- [16] J. R. Janesick, *Scientific charge-coupled devices*, Vol. 117, SPIE press Bellingham, 2001.
- [17] J. Lukas, J. Fridrich, M. Goljan, Digital camera identification from sensor pattern noise, *Information Forensics and Security, IEEE Transactions on* 1 (2) (2006) 205–214.
- [18] M. Chen, J. Fridrich, M. Goljan, J. Lukás, Determining image origin and integrity using sensor noise, *Information Forensics and Security, IEEE Transactions on* 3 (1) (2008) 74–90.
- [19] M. Goljan, J. Fridrich, T. Filler, Large scale test of sensor fingerprint camera identification, in: *IS&T/SPIE Electronic Imaging, International Society for Optics and Photonics, 2009*, pp. 72540I–72540I.
- [20] C.-T. Li, Y. Li, Color-decoupled photo response non-uniformity for digital image forensics, *Circuits and Systems for Video Technology, IEEE Transactions on* 22 (2) (2012) 260–271.
- [21] C.-T. Li, Source camera identification using enhanced sensor pattern noise, *Information Forensics and Security, IEEE Transactions on* 5 (2) (2010) 280–287.
- [22] X. Kang, Y. Li, Z. Qu, J. Huang, Enhancing source camera identification performance with a camera reference phase sensor pattern noise, *Information Forensics and Security, IEEE Transactions on* 7 (2) (2012) 393–402.
- [23] K. Kurosawa, K. Kuroki, N. Saitoh, Ccd fingerprint method-identification of a video camera from videotaped images, in: *Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on*, Vol. 3, IEEE, 1999, pp. 537–540.
- [24] T. Filler, J. Fridrich, M. Goljan, Using sensor pattern noise for camera model identification, in: *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, IEEE, 2008, pp. 1296–1299.
- [25] T. Gloe, M. Kirchner, A. Winkler, R. Böhme, Can we trust digital image forensics?, in: *Proceedings of the 15th international conference on Multimedia, ACM, 2007*, pp. 78–86.
- [26] T. H. Thai, F. Retraint, R. Cogranne, Generalized signal-dependent noise model and parameter estimation for natural images, *Signal Processing* 114 (2015) 164–170.
- [27] T. H. Thai, F. Retraint, R. Cogranne, Camera model identification based on the generalized noise model in natural images, *Digital Signal Processing* 48 (2016) 285–297.
- [28] A. Foi, M. Trimeche, V. Katkovnik, K. Egiazarian, Practical poissonian-gaussian noise modeling and fitting for single-image raw-data, *Image Processing, IEEE Transactions on* 17 (10) (2008) 1737–1754.
- [29] T. H. Thai, R. Cogranne, F. Retraint, Camera model identification based on the heteroscedastic noise model, *Image Processing, IEEE Transactions on* 23 (1) (2014) 250–263.
- [30] T. Qiao, F. Retraint, R. Cogranne, T. H. Thai, source device identification based on raw images, *ICIP*.
- [31] B. Widrow, I. Kollar, M.-C. Liu, Statistical theory of quantization, *IEEE Transactions on Instrumentation and Measurement* 45 (2) (1996) 353–361.
- [32] E. Y. Lam, J. W. Goodman, A mathematical analysis of the dct coefficient distributions for images, *Image Processing, IEEE Transactions on* 9 (10) (2000) 1661–1666.
- [33] T. H. Thai, R. Cogranne, F. Retraint, Statistical model of natural images, in: *Proc. IEEE. Int. Conf. Image Process.*, 2012, pp. 2525–2528.
- [34] T. H. Thai, R. Cogranne, F. Retraint, Statistical model of quantized DCT coefficients: Application in the steganalysis of Jsteg algorithm, *IEEE Trans. Image Process.* 23 (5) (2014) 1980–1993.
- [35] T. Gloe, R. Böhme, The dresden image database for benchmarking digital image forensics, *Journal of Digital Forensic Practice* 3 (2-4) (2010) 150–159.
- [36] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, Image denoising by sparse 3-d transform-domain collaborative filtering, *Image Processing, IEEE Transactions on* 16 (8) (2007) 2080–2095. doi:10.1109/TIP.2007.901238.
- [37] M. Lebrun, An analysis and implementation of the bm3d image denoising method, *Image Processing On Line* (2012).
- [38] R. J. Lehman, E., *Testing statistical hypotheses*, in: *Second Edition*, Springer, 2005.
- [39] E. L. Lehmann, J. P. Romano, *Testing statistical hypotheses*, Springer, Germany, 2006.
- [40] R. Cogranne, F. Retraint, An asymptotically uniformly most powerful test for LSB matching detection, *Information Security and Forensics, IEEE Transactions on* 8 (3) (2013) 464–476.
- [41] R. Cogranne, C. Zitzmann, L. Fillatre, F. Retraint, I. Nikiforov, P. Cornu, A cover image model for reliable steganalysis, in: *Information Hiding, Springer, 2011*, pp. 178–192.
- [42] T. Qiao, C. Zitzmann, R. Cogranne, F. Retraint, Detection of JSteg algorithm using hypothesis testing theory and a statistical model with nuisance parameters, in: *Proceedings of the 2nd ACM workshop on Information hiding and multimedia security, ACM, 2014*, pp. 3–13.
- [43] C. Zitzmann, R. Cogranne, L. Fillatre, I. Nikiforov, F. Retraint, P. Cornu, Hidden information detection based on quantized laplacian distribution, in: *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, IEEE, 2012, pp. 1793–1796.
- [44] R. Cogranne, J. Fridrich, Modeling and extending the ensemble classifier for steganalysis of digital images using hypothesis testing theory, *Information Forensics and Security, IEEE Transactions on* 10 (12) (2015) 2627–2642. doi:10.1109/TIFS.2015.2470220.
- [45] R. Cogranne, J. Fridrich, Modeling and extending the ensemble classifier for steganalysis of digital images using hypothesis testing theory, *IEEE Transactions on Information Forensics and Security* 10 (12) (2015) 2627–2642.
- [46] B. Baygun, A. O. Hero, Optimal simultaneous detection and estimation under a false alarm constraint, *IEEE Transactions on Information Theory* 41 (3) (1995) 688–703.