On the Sensor Pattern Noise Estimation in Image Forensics: A Systematic Empirical Evaluation

Mustafa Al-Ani, and Fouad Khelifi, *Member, IEEE*

***Abstract*— Extracting a fingerprint of a digital camera has fertile applications in image forensics, such as source camera identification and image authentication. In the last decade, Photo Response Non_Uniformity (PRNU) has been well established as a reliable unique fingerprint of digital imaging devices. The PRNU noise appears in every image as a very weak signal, and its reliable estimation is crucial for the success rate of the forensic application. In this paper, we present a novel methodical evaluation of 18 state-of-the-art PRNU estimation/enhancement techniques that have been proposed in the literature in various frameworks. The techniques are classified and systematically compared based on their role/stage in the PRNU estimation procedure, manifesting their intrinsic impacts. The performance of each technique is extensively demonstrated with over 2.2 million test images to conclude this case-sensitive study. The experiments have been conducted on our created database and a public image database, the 'Dresden image database'.**

***Index Terms*—Authentication, camera identification, digital forensics, photo response non-uniformity (PRNU), sensor pattern noise (SPN).**

I. Introduction

A. Background

Nowadays, digital cameras have increasingly become affordable and available for almost everyone in the society, and hence millions of pictures are being taken, transmitted and saved digitally on a daily basis. In the file headers of these digital images, there is useful information about the source camera, time and data, camera settings, exposure, etc. However, this information can be easily stripped off and tampered, and hence it cannot be used as a trustworthy source for sensitive issues, such as courtrooms and criminal evidence.

Digital cameras leave traces in the pixel data of their images. Researchers have found and extracted traces and features of different types and origins to use for various image forensic analyses. Forensic applications in general demand a substantially high accuracy, and one of the most reliable features that can provide such accuracy was first exploited by Lukas et al. [1] and it is known as the Photo Response Non_Uniformity (PRNU). It results from the variations of the sensor pixels at collecting the light energy (this is due to imperfections in the manufacturing of the pixels' physical dimensions as well as the non homogeneity that is naturally present in the silicon in sensors). The variations in quantum efficiency among pixels can be captured and denoted with a matrix $\mathbf{K} \in \mathbb{R}^{M \times N}$, where $M \times N$ are the dimensions of the sensor. $\mathbf{K}$ follows a (zero-mean) white Gaussian distribution. When an imaging sensor is illuminated with light intensity $\mathbf{Y} \in \mathbb{R}^{M \times N}$, in the absence of other noise sources, the sensor generates a signal $\mathbf{Y} + \mathbf{KY}$. (The product of the matrices herein is elementwise.)

With the described underlying mechanism of generating the aforementioned non-uniformity, a unique pattern of spatial noise that is fixed for an individual camera is integrated in every image. In contrast to other sources of random noises, this noise is of a deterministic nature and cannot be eliminated by averaging (‘pattern noise’ is the term used in the literature to describe such systematic noise). However, whilst other sources of noise are added and the generated signal is gamma corrected, colour de-mosaicked and corrected, de-noised and subjected to few other operations in the pipeline of digital cameras, the PRNU noise can still survive for estimation [2]. In general, forensic applications of PRNU fingerprint fall in two categories:

1) Image Origin Identification: There are various applications under this category. The most popular applications are source camera identification and source camera verification. For the former, the main goal is to identify the exact camera that was used to take a query image among other cameras provided to the analyst. In verification, however, the forensic analyst aims to determine whether an image was taken by a certain camera or not. In both cases, the cameras or sets of images taken by the cameras are available to the analyst. Another application, known as fingerprint matching, is to link a set of images to another set among a large database. This scenario could be met when a set of malicious images become available to the analyst to search a public database to find images taken by the same camera. More applications include video clip linking in which the aim is to attribute a set of images to another set where the set is taken from a video clip. The problem of image origin has been attempted in the literature using different approaches, such as [3]-[7]. However, a key characteristic of PRNU fingerprint is that it serves as an intrinsic feature that can represent the

This work is supported by the EPSRC Research Grant (EP/L006812/1).

Mustafa Al-Ani is with the Faculty of Science and Technology, University of Westminster, 115 New Cavendish Street, London, W1W 6UW, UK. Email: m.alani1@westminster.ac.uk.

Fouad Khelifi is with the Department of Computer Science and Digital Technology, University of Northumbria, Newcastle, NE2 1E, UK. Email: fouad.khelifi@northumbria.ac.uk.

individual imaging device sensor. So that, it is not only possible to identify and differentiate device models of the same make, but also individual devices of the same model. Other methods such as [8] incorporate the PRNU ingredient with another camera model identification approach to refine the final results. PRNU fingerprint has also been used in scanner identifications in [9] and [10].

2) Image Forgery Detection: The PRNU can serve as a kind of watermark for image forgery detection. The idea is that forgery operations such as object copy-move will change or destroy the PRNU in the forgery area. This is met when the forged object comes from another camera or the spatial location is changed when compared to the original image. Indeed, some malicious acts may preserve the PRNU and cannot be detected using this approach. There are numerous image forgery detection techniques in the literature [11]-[15], each has its assumption on the nature of manipulation. With no universal approach, the practical solution is to incorporate a set of approaches of different principles to detect forgeries of various types.

Many studies have confirmed the reliability of the PRNU as a fingerprint for image forensic analysis, such as [16] where a large-scale experiment was conducted—over one million images were tested spanning 6896 individual cameras covering 150 models. The promising results of the PRNU fingerprint have drawn the attention of many research groups in the last decade. However, the success rates of its forensic applications are dedicated by the quality of the estimation of this weak signal. Correspondingly, many techniques appeared with the objective of improving the success rate of a PRNU forensic application through implicitly or explicitly enhancing the quality of the PRNU estimation. The efficiency of a proposed technique is usually demonstrated in a numerical experiment of a particular forensic application in a comparison with few other techniques. Although these concerned techniques go more or less in the same direction, they target different stages of the PRNU estimation procedure. Thus, the impacts of many of these improved PRNU estimation techniques are relatively unapparent, even for experts in the field. This calls for a study that collects and categorises all these techniques in order to systematically evaluate, compare and manifest their contributions. This paper is the first and most complete effort to classify *all* the techniques proposed in the literature based on their role/stage in the PRNU estimation procedure. And, under each category those techniques are reviewed and numerically evaluated through intensive experiments. The relatively large-scale experiments are crucial because of the high variations in the performances of the methods that can be seen among cameras and images.

B. The standard Procedure of PRNU Estimation

The PRNU of a camera can be estimated using L of its images: $\mathbf{I}_l \in \mathbb{R}^{M\times N}$, $l = 1, \dots, L$. Each image is de-noising first to separate the "original" image from its noise:

$$\hat{\mathbf{r}} \coloneqq \mathbf{I} - F(\mathbf{I}) \,. \tag{1}$$

where $F(.)$ is a filtering operation. The noise residual $\hat{\mathbf{r}}$ contains the PRNU noise and other types of random noises that cannot be used in forensic applications. The noise residuals of the images are then typically averaged to suppress the random noises, and estimate the reference PRNU that serves as the camera fingerprint. For further refinement of the PRNU estimate, diverse additional signal processing strategies have been proposed and adopted in the literature.

C. Paper Evaluation System and Outline

The diagram in Fig. 1 summarises the key stages of the PRNU estimation. The various techniques proposed in the literature are categorised in this paper according to these stages. The paper is organised as follows. We start with describing the numerical evaluation framework and the benchmark database in Section II. After we describe the PRNU-focused camera output model in Section III, we theoretically and numerically analyse the methods under each stage of the PRNU estimation procedure in each section respectively. In Section IV, we analyse several *de-noising operations*, and provide our findings. We investigate the standard and the alternative approaches of *noise residual combining* in Section V. Section VI covers the various pre/post-calculation *PRNU enhancement techniques*. In Section VII, the various *compact PRNU representation methods* are studied and evaluated. *Alternative enhancement strategies* are examined in Section VIII. In Section IX, we look over the different similarity measures that are used to compare and link the estimated PRNUs in forensic applications. In Section X, conclusions are drawn.

D. Notation

We remark that throughout the paper the boldface font denotes matrices. Unless otherwise stated, operations among/on matrices such as product, raising to a power, ratio, and summation are elementwise. Also, (m, n), $m = 1, \dots, M, n = 1, \dots N$, where $M \times N$ are the dimensions of the sensor, represent the pixel positions, and they are used as indices for matrices to designate their (pixel) components.

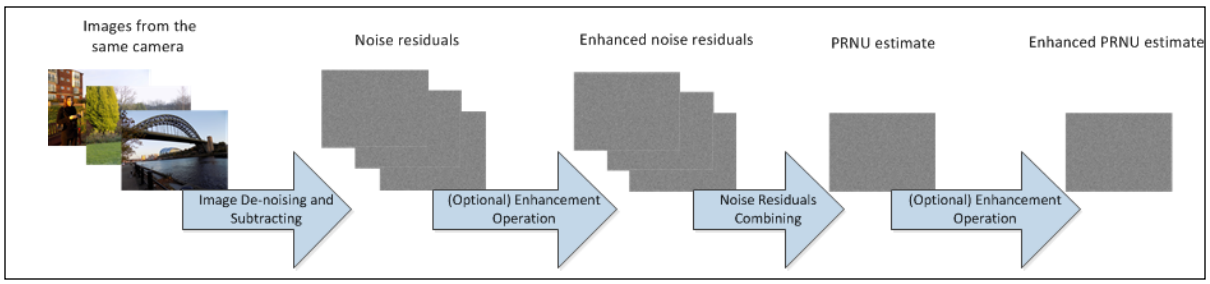

Fig. 1. A diagram of the typical PRNU estimation procedure.

II. Numerical Evaluation Framework

A. Performance Metric

As we mentioned earlier, the quality of the PRNU estimation is the performance characteristic that we seek to evaluate. To quantify the quality of the estimated PRNU, we measure the similarity between PRNUs extracted from the same cameras, i.e. intraclass PRNU estimates, as well as the dissimilarity between PRNU estimates of different camera sources, interclass PRNUs. The standard correlation coefficient, i.e. the normalized cross-correlation, is used as a similarity measure.

The estimated PRNU can contain different types of contaminations; camera-model specific noise can be one of them. The presence of this noise in estimated PRNUs can potentially increase the dissimilarity between some interclass PRNU estimates, as opposed to more accurate and unique PRNU estimates with less camera model noises. Accordingly, the numerical results would be influenced by the choice of the interclass cameras in the experiments. Therefore, in our evaluation, we set the interclass experiments to PRNUs extracted from the same camera model.

Ideally speaking, the distribution of the correlation coefficients between interclass PRNU estimates should be concentrated around zero, whereas the intraclass PRNU estimates should provide a correlation close to one. Nonetheless, the high impurity of the estimated PRNUs causes the two distributions to come near each other and overlap. An example of the distributions of the intraclass and interclass correlation coefficients of the PRNU estimates for a certain camera is shown in Fig. 2. These distributions could be modelled for each camera as generalized Gaussian distributions. However, the correlations cannot be described precisely using this model or any other model; accordingly we cannot exploit the model's parameters as reliable performance metrics. Nonetheless, the "distance" between the two distributions echoes the quality of the estimated PRNU signals. Therefore, the rate of overlap would be considerable for a performance marker. To this end, we draw and exploit the ROC curve to discriminate the interclass and intraclass correlation coefficients and drive two well-established performance measures. Since most of our studied techniques are developed to achieve low false positive rates in their forensic application, the first metric we evaluate is the true positive rate at $\gamma = 10^{-3}$ false positive rate in the ROC curve. We denote this metric for the *c*-th camera by $\mathcal{P}_c$. As a second metric to provide a more complete picture, we evaluate the false positive (or negative) rate at the point where the false positive and false negative rates are equal in the ROC curve, which is known as the equal error rate (EER). We use the notation $\mathcal{R}_c$ to represent this performance measure for the *c*-th camera.

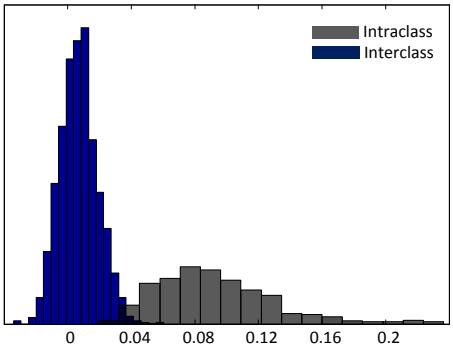

Fig 2. The distributions of the intraclass and interclass correlation coefficients of the PRNU estimates for an example camera.

B. Dataset and Experimental Setup

Forty-five cameras, listed in Table I and II below, from our database and the 'Dresden image database' [17], [18] are used to benchmark all the studied techniques. (We note that some images from Dresden image database do not have the original resolution of their cameras because of digital zooming, and they must be excluded in PRNU based applications.) For each camera, we create 144 sets of non-overlapping sub-images of size 64×64 cropped from the original images. Such cropping would allow us to expand our database, and it would increase the challenge on the PRNU-based discrimination in order to noticeably differentiate between the performances of the studied methods. The sub-images of each set of each camera are grouped into $L = 50$ sub-images, and they are de-noised and combined to yield a PRNU fingerprint estimate (for a camera in our database for example, this would yield 10 PRNU estimates within each of the 144 sets and hence 1440 estimates per camera). The PRNU estimates within each set are compared with each other as intraclass experiments. On the other hand, the similarities between PRNU estimates of different cameras but of the same model are measured as interclass experiments (for one camera in our database, this would lead to 6480 intraclass experiments and over 2×10^6 interclass experiments). For unique cameras in Table I which have no other cameras of the same model, since sub-images cropped from non-overlapping parts of the original images are meant to contain different PRNU signals, the similarities of those PRNU estimates are measured as interclass experiments. We can observe that the number of experiments per camera is much larger than $1/\gamma$, which would lead to a much reliable measure of the delicate metric $\mathcal{P}_c$, i.e. the true positive rate at $\gamma = 10^{-3}$ false positive rate in the ROC curve.

All experiments are performed in MATLAB on 45 computers in parallel with an Intel Core Duo i7-4770 @ 3.40GHz processor and 16 GB of memory. Before we close this section, we note that in colour images, the PRNUs can be estimated from the three colour channels separately and then combined using the standard RGB to grey scale conversion. Or, PRNU extraction can be performed once on the combination of the three colour channels, i.e. luminance channel. An alternative approach is to focus solely on the green channel since it carries most of the PRNU information and the least interpolation noise (a further discussion is included in Section VIII). We adopt the last approach herein to ease our long experiments.

III. (PRNU-Focused) Camera Model

Regardless of the sensor type, the average signal generated at a sensor from $\mathbf{Y} \in \mathbb{R}^{M \times N}$ illumination is

$$\mathbf{Y} + \mathbf{KY}\,, \tag{2}$$

where $\mathbf{K} \in \mathbb{R}^{M \times N}$ represents the PRNU that follows a white Gaussian distribution**.** Another source of pattern noise is

TABLE I. THE LIST OF CAMERAS FROM DRESDEN IMAGE DATABASE USED IN OUR NUMERICAL EVALUATION.

Camera No.	Camera Name	No. of Images
1.	SONYDSC-HX200V	630
2.	Kodak_M1063_4	571
3.	Kodak_M1063_0	464
4.	Kodak_M1063_3	460
5.	Kodak_M1063_1	458
6.	Kodak_M1063_2	438
7.	SamsungL301	522
8.	Panasonic_DMC-FZ50_1	415
9.	Panasonic_DMC-FZ50_0	265
10.	Panasonic_DMC-FZ50_2	251
11.	Nikon_D200_1	380
12.	Nikon_D200_0	372
13.	Agfa_Sensor530s	373
14.	Agfa_DC-830I	363
15.	Sony_DSC-H50_0	284
16.	Agfa_DC-733S	281
17.	Samsung_L74wide_0	232
18.	Samsung_L74wide_2	231
19.	Samsung_L74wide_1	224
20.	Canon_Ixus55	224
21.	Olympus_mju_1050SW_2	218
22.	FujiFilm_J50_2	215
23.	FujiFilm_J50_1	205
24.	FujiFilm_J50_0	210
25.	Sony_DSC-H50_1	257
26.	Samsung_NV15_0	214
27.	Samsung_NV15_1	211
28.	Samsung_NV15_2	211
29.	Olympus_mju_1050SW_1	209
30.	Olympus_mju_1050SW_3	207
31.	Praktica_DCZ5.9_0	209
32.	Praktica_DCZ5.9_3	206
33.	Praktica_DCZ5.9_1	205
34.	Praktica_DCZ5.9_2	205
35.	Sony_DSC-W170_0	205

TABLE II. THE LIST OF OUR CAMERAS USED IN THE NUMERICAL EVALUATION.

Camera No.	Camera Name	No. of Images
1.	Nikon L330_0	500
2.	Nikon L330_1	500
3.	Panasonic TZ20_0	500
4.	Panasonic TZ20_1	500
5.	Fujifilm S2950_0	500
6.	Fujifilm S2950_1	500
7.	Canon IXUS_0	500
8.	Canon IXUS _1	500
9.	Samsung PL120_0	500
10.	Samsung PL120_1	500

introduced at the imaging sensors, known as dark currents. This is due to thermal energy that can generate free electrons in silicon with no illumination exposed on the sensor. There are small fluctuations in the number of generated dark electrons from pixel to another. Yet, this sensor pattern imperfection cannot be used in image forensic. This is due to its high dependence on the temperature and its direct proportionality to the exposure time setting in the camera that is not always available for the analyst. Also, the dark currents are suppressed in some cameras by subtracting a dark frame from the final image. The two sensor imperfections are known combined as sensor pattern noise (SPN). However, PRNU noise is the most dominant part of SPN, and unlike the other component it is always present in an image and cannot be subtracted in common consumer cameras. Hence, several papers in the field recognise SPN as the fingerprint of a camera sensor. With slight abuse of terms, we use the terms exchangeably to maintain the consistency of the terminology in this paper.

As we mentioned above, (2) represents the average number of collected electrons. The actual number can be more/less than or equal to the average, and its distribution about the average follows a Poisson distribution (where its variance equals its mean). It is usually referred to as shot noise or photonic noise. From above, the number of collected electrons can be expressed as

$$\mathbf{Y} + \mathbf{KY} + \mathbf{N_{DC}} + \mathbf{N_s}\,, \tag{3}$$

where $\mathbf{N_{DC}} \in \mathbb{R}^{M\times N}$ is the number of electrons due to thermal energy, and $\mathbf{N_s} \in \mathbb{R}^{M\times N}$ is the zero-mean result of the Poisson shot noise. The output amplifier that transforms the photon-induced electrons at the sensors into a measurable signal adds a zero-mean read-out noise that is independent of the value of the signal. The signal is then gamma corrected to adjust to human vision and quantised with an ADC before saving. The final image can be expressed as

$$\mathbf{I} = g^{\gamma}(\mathbf{Y} + \mathbf{KY} + \mathbf{N_s} + \mathbf{N_{DC}})^{\gamma} + \mathbf{N_q}\,, \tag{4}$$

where g is the amplifier gain, γ(=0.45 typically) is the gamma factor, and $\mathbf{N_q} \in \mathbb{R}^{M\times N}$ is the quantisation noise (the reader can refer to [19] for more details about camera noise sources and characteristics). With the Taylor expansion $(1+x)^{\alpha} = 1 + \alpha x + O(x^2)$ at $x = 0$, and by re-arranging the bracket in (4) into the former, we reach

$$\mathbf{I} = g^{\gamma}\mathbf{Y}^{\gamma}\left[1 + \gamma\mathbf{K} + \gamma\mathbf{N_s} + \gamma\mathbf{N_{DC}} + O\left(\left|\mathbf{K} + \frac{\mathbf{N_s}+\mathbf{N_{DC}}}{\mathbf{Y}}\right|^2\right)\right] + \mathbf{N_q}. \tag{5}$$

The last term in the square bracket is small and can be ignored. Let $\mathbf{I_0} \coloneqq g^{\gamma}\mathbf{Y}^{\gamma}$ and $\mathbf{N_t} \coloneqq \gamma\mathbf{N_s} + \gamma\mathbf{N_{DC}} + \mathbf{N_q}$ denotes the combination of the independent random noise components. To avoid introducing many notations, the symbols are absorbed as follows $\mathbf{K} \coloneqq \gamma\mathbf{K}$. This leads to

$$\mathbf{I} = \mathbf{I_0} + \mathbf{KI_0} + \mathbf{N_t}\,. \tag{6}$$

The model is more or less adopted in all the existing PRNU-based techniques despite the various terminologies. And, many techniques model $\mathbf{KI_0} + \mathbf{N_t}$ combined as white Gaussian process. In the literature, some authors distinguish $\mathbf{K}$ by the PRNU factor and $\mathbf{KI_0}$ by the PRNU signal. Nonetheless, $\mathbf{K}$ is the actual fingerprint of a camera, and all the techniques implicitly or explicitly seek to estimate this quantity or a scaled version of it—which we simply refer to by the PRNU.

IV. DE-NOISING OPERATIONS

Various de-noising methods have been exploited in the PRNU extraction in the image forensics research literature. In the next subsections, we discuss all the techniques developed and adopted in the estimation of this weak signal.

A. Wavelet-Based Filter

This filter was originally proposed in [20], and it operates as follows. The fourth-level wavelet decomposition of the image with the 8-tap Daubechies quadrature mirror filter is first calculated. Let the wavelet coefficients in the vertical, horizontal, and diagonal subbands be respectively denoted by

$\mathbf{h}(i,j), \mathbf{v}(i,j), \mathbf{d}(i,j),\ (i,j) \in \mathcal{T}$, where $\mathcal{T}$ is the index set of the wavelet coefficients that depends on the decomposition level. The de-noised wavelet coefficients are obtained using the Wiener filter:

$$\mathbf{h}_w(i,j) \coloneqq \mathbf{h}(i,j)\frac{\widehat{\boldsymbol{\sigma}}^2(i,j)}{\widehat{\boldsymbol{\sigma}}^2(i,j)+\sigma_0^2}, \qquad (7)$$

and similarly for $\mathbf{v}(i,j)$ and $\mathbf{d}(i,j)$. σ_0^2 is the variance of the noise that is assumed to be a white Gaussian process, and $\widehat{\boldsymbol{\sigma}}^2(i,j)$ represents the estimated local variance of the wavelet coefficients of the "original" noise-free image—these coefficients are modelled as locally stationary iid variables with zero mean. The maximum a posteriori (MAP) estimation is used to obtain the local variance:

$$\widehat{\boldsymbol{\sigma}}^2{}_q(i,j) = \max\left(0, \frac{1}{q^2}\sum_{(x,z)\in\mathcal{B}_q}\mathbf{h}^2(x,z) - \sigma_0^2\right), \qquad (8)$$

where $q \times q$ is the size of the window $\mathcal{B}_q$ around (i,j); it was proposed to set $q \in \{3,5,7,9\}$. The minimum of the four variances is used in (7), i.e.

$$\widehat{\boldsymbol{\sigma}}^2(i,j) = \min\left(\widehat{\boldsymbol{\sigma}}^2{}_3(i,j), \widehat{\boldsymbol{\sigma}}^2{}_5(i,j), \widehat{\boldsymbol{\sigma}}^2{}_7(i,j), \widehat{\boldsymbol{\sigma}}^2{}_9(i,j)\right). \qquad (9)$$

The de-noised image is then obtained by applying the inverse wavelet transform on the de-noised coefficients. It was shown in [1] that the choice of σ_0^2 has little impact on the performance of the filter in PRNU extraction. The authors, throughout their various versions of this work, suggested setting σ_0 between 2 and 5.

B. Context-Adaptive Interpolator (CAI)

In [21] the authors proposed an estimator based on an eight-neighbour context-adaptive interpolation algorithm to supress the effect of image scenes (a four-neighbour version was proposed in their conference paper [22]). It aims to identify edges and to produce a high-quality PRNU estimate. According to this method, the local regions are classified into six types: smooth, horizontally edged, vertically edged, forward-diagonally edged, backward-diagonally edged, and other. A mean filter is used to estimate the centre pixel value in the smooth regions. In edge regions the centre pixel is predicted along the edge. In other regions a median filtering is used. To put this in a concise equation, the pixels $\mathbf{I}(m,n+1)$ are designated by e, $\mathbf{I}(m+1,n+1)$ by se, $\mathbf{I}(m+1,n)$ by s, $\mathbf{I}(m+1,n-1)$ by sw, $\mathbf{I}(m,n-1)$ by w, $\mathbf{I}(m-1,n-1)$ by nw, $\mathbf{I}(m-1,n)$ by no, and $\mathbf{I}(m-1,n+1)$ by ne, and $\mathbf{A} = [e, se, s, sw, w, nw, no, ne]'$. The centre pixel value is given by:

$$\hat{\mathbf{r}}(m,n) = \begin{cases} \mathbf{I}(m,n) - \text{mean}(\mathbf{A}), & \left(\max(\mathbf{A}) - \min(\mathbf{A})\right) \le 20 \\ \mathbf{I}(m,n) - \frac{(s+no)}{2}, & (|e-w| - |no-s|) > 20 \\ \mathbf{I}(m,n) - \frac{e+w}{2}, & (|s-no|, -|e-w|) > 20 \\ \mathbf{I}(m,n) - \frac{se+nw}{2}, & (|sw-ne| - |se-nw|) > 20 \\ \mathbf{I}(m,n) - \frac{sw+ne}{2}, & (|se-nw| - |sw-ne|) > 20 \\ \mathbf{I}(m,n) - \text{median}(\mathbf{A}), & \text{otherwise}. \end{cases} \qquad (10)$$

Then spatial Wiener filtering operation is exploited to eliminate the impact of the image scenes leaked in (10), where $\hat{\mathbf{r}}$ is modelled as an additive mixture of the locally stationary iid zero-mean signal of image content and white Gaussian noise. The noise variance in this operation is set to 9 and the MAP estimation, described in Subsection IV.A above, is used to obtain the local variance of the image content with a window of size 3×3.

C. 2-Pixel Approach

Modern low-medium end cameras have high sensor pixel density, and it is very likely that two pixels in the same vicinity have close values in natural images. This method [23] capitalises on this observation in spatial domain filtering. It counts on engaging as little as one adjacent pixel at estimating the PRNU at a pixel location in order to suppress the correlation between neighbouring pixels in the estimated (supposedly white) signal. Let (x,z) designate a pixel location in the close vicinity of (m,n). Simply speaking, based on the model of $\mathbf{I}$ in (6), a noisy estimate of $\mathbf{K}(m,n)\mathbf{I_0}(m,n)$ at an individual pixel can be produced by subtracting its value $\mathbf{I}(m,n)$ by another pixel of the same (amplified and gamma corrected) illumination $\mathbf{I_0}(x,z)$ but with opposite signed $\mathbf{K}(x,z)$. Thus to obtain a rough PRNU estimate, i.e. noise residual, at a certain pixel of an image, this technique searches the 16 next-to-the-adjacent pixels to find one pixel with a comparable value $\mathbf{I}(x,z)$ and an alternative signed $\mathbf{K}(x)\mathbf{I_0}(x,z)$. We saw in Section III that the level of a considerable part of the random noise at a pixel (m,n) depends on its illumination. This search considers this pixel dependent noise in finding a neighbouring pixel with a close value $\mathbf{I_0}(x,z)$. The noise residual at pixel (m,n) of an image is estimated as

$$\hat{\mathbf{r}}(m,n) = \frac{\mathbf{I}(m,n) - \mathbf{I}(x,z)}{2},$$
$$|\mathbf{I}(m,n) - \mathbf{I}(x,z)| < \sqrt{\mathbf{I}(m,n)} \text{ and } \mathbf{K}(m,n)\mathbf{K}(x,z) < 0. \qquad (11)$$

Starting from the pixel on the (let's say) top right corner, the search runs (anti)clockwise to find the first pixel that fulfills the conditions in (11). Otherwise, $\hat{\mathbf{r}}(m,n)$ is set to no value and not considered in the combining process. The add-hoc threshold $\sqrt{\mathbf{I}(m,n)}$ in (11) reflects the standard deviation of the shot noise at the addressed pixel. The method exploits a prior rough estimate of the signs of $\mathbf{K}$, which can be obtained by any basic filtering operation such as median filtering of few images of L. In the original paper, the 8 adjacent pixels are engaged instead. It has been modified because of the high correlation in adjacent pixels resulted from the interpolation operation that has been observed in a few cameras.

D. Adaptive Spatial (AS) Filtering

In [24], it was suggested that a rather simple space variant filtering technique may be more useful in estimating the PRNU because of the relaxed requirement of de-noising every image entirely. Their work is based on two stage filtering. The first stage is the standard adaptive Wiener filter [25] that operates directly in the spatial domain:

$$\hat{\mathbf{r}}(m,n) = \mathbf{I}(m,n) - \mathbf{U}(m,n) + [\mathbf{I}(m,n) - \mathbf{U}(m,n)]\frac{\widehat{\boldsymbol{\sigma}}^2(m,n)}{\widehat{\boldsymbol{\sigma}}^2(m,n)+\sigma_0^2}, \qquad (12)$$

where σ_0^2 is the variance of the noise that is assumed to be a white process, and $\mathbf{U}(m,n)$ and $\widehat{\boldsymbol{\sigma}}^2(m,n)$ are the local mean and variance of the original image within the $q \times q$ pixel-size window $\mathcal{B}_q$ around the pixel (m,n), respectively:

$$\mathbf{U}(m,n) = \frac{1}{q^2}\sum_{(x,z)\in\mathcal{B}_q}\mathbf{I}(x,z), \quad (13)$$

$$\widehat{\boldsymbol{\sigma}}^2(m,n) = [\frac{1}{q^2}\sum_{(x,z)\in\mathcal{B}_q}\mathbf{I}^2(x,z) - \mathbf{U}^2(m,n)] - \sigma_0^2. \quad (14)$$

It was recommended in [24] to use a window of size 9×9 pixels and set σ_0^2 to 5 in extracting the noise residuals from the L images. The second stage consists of two 2×2 cascaded median filters to suppress the impulse pixels in $\hat{\mathbf{r}}$.

E. Other Image De-Noising Filters

Renowned image de-noising filters that have been explored in thousands of studies and applications have also been adopted in PRNU extraction. In this subsection, we briefly outline these seminal filters:

1) Perona-Malik diffusion (PMD) filter*:* The filter was adopted in PRNU estimation in [26]. Perona and Malik presented this filter in [27] based on the anisotropic diffusion equation:

$$\frac{\partial\mathbf{I}(m,n,t)}{\partial t} = \nabla.\left(c(|\nabla\mathbf{I}(m,n,t)|)\nabla\mathbf{I}(m,n,t)\right), \quad (15)$$

where ∇ and $\nabla.$ are the gradient and divergence operations, respectively. In their discrete form of (15), $\mathbf{I}(m,n,t)$ is a de-noised version of the image at scale t (where $\mathbf{I}(m,n,0)$ is the original image), and $\frac{\partial\mathbf{I}(m,n,t)}{\partial t}$ $=\mathbf{I}(m,n,t+1)-\mathbf{I}(m,n,t)$. $c(|\nabla\mathbf{I}(m,n,t)|)$ is the diffusion coefficient that is chosen to preserve edges and textures. One of the two Perona and Malik's diffusion coefficients was chosen in [26] owing to its better performance in PRNU extraction. That is,

$$c(\upsilon) \coloneqq -e^{-(|\upsilon|/\chi)^2}, \quad (16)$$

where χ is a gradient threshold parameter that is determined at each iteration; it represents the value below which 90% of the absolute value of the whole image gradient occur. The Perona-Malik's discrete form of the right hand side of (15) is

$$\frac{\lambda}{|\mathcal{M}|}\sum_{(x,z)\in\mathcal{M}} c(|\mathbf{I}(m,n,t)-\mathbf{I}(x,z,t)|)(\mathbf{I}(m,n,t)-\mathbf{I}(x,z,t))$$

where $\mathcal{M}=\{(m,n+1),(m,n-1),(m+1,n),(m,n)\}$, $|\mathcal{M}|=4$ (an 8-neighbouring pixel approximation can also be adopted) and $\lambda\in(0,1]$ determines the rate of diffusion. In [26], λ was fixed to 1/7, and the number of scales t was set to three, i.e. $\hat{\mathbf{r}}(m,n)=\mathbf{I}(m,n,3)-\mathbf{I}(m,n,0)$.

2) Total variation (TV) filter*:* In the classical definition of this filter that was introduced in [28], the total variation of the image is minimized subject to constraints involving the statistics of the noise. The constraints are imposed using Lagrange multipliers. The solution is obtained using the gradient-decedent method (many new other optimization techniques have appeared in the compressive sensing field). The filter preserves edges whilst smooth away noise in flat regions, even at low signal-to-noise ratios. The PRNU estimation work of [29] is based on a simplified version of the total variation filter. The authors adopted the unconstrained total variation method proposed in [30] and used the gradient-decedent optimization. With each step

$$\mathbf{I}(m,n,t) = -\nabla.\left(\frac{\nabla\mathbf{I}}{|\nabla\mathbf{I}|_\varepsilon}\right) + \mathbf{I}(m,n,t-1), \quad (17)$$

$\mathbf{I}(m,n,t)$ is a de-noised version of the image at iteration step t (where $\mathbf{I}(m,n,0)$ is the original image) and $|\nabla\mathbf{I}|_\varepsilon = \sqrt{|\nabla\mathbf{I}|^2+\varepsilon^2}$, where ε is inserted to avoid singularities (we set $\varepsilon=10^{-3}$ in our implementation). To simplify it further, the authors recommended using only one step in the gradient-decedent optimization, which they hence name it as the first step total variation (FSTV) filter. The noise residual for this case will be given by $\hat{\mathbf{r}} = -\nabla.(\nabla\mathbf{I}/|\nabla\mathbf{I}|_\varepsilon)$. Whilst the authors' primary aim is to adopt a simple, fast de-noising operation in PRNU extraction, they also seek a more accurate estimate of the PRNU compared to other filters.

3) Block-matching and 3D (BM3D) algorithm: The filter, that was introduced in [31], has been explored in PRNU estimation in [32] and [33]. The filter combines sliding-window transform processing with *block-matching*, where a pixel of the true image is estimated from regions which are found similar to the region centered at the estimated pixel. The filter operates in the following steps. Image blocks are processed in a sliding manner to search for blocks that exhibit similarity to the currently-processed one. The matched blocks are stacked together to form a 3D array. A 3D transformation of the array is applied to produce a sparse representation of the true signal in 3D transform domain. Then efficient noise attenuation is achieved by applying a shrinkage operator (e.g. hardthresholding or Wiener filtering) on the transform coefficients. Inverse 3D transform of the filtered coefficients yields the local estimates of the blocks. This results in an improved de-noising performance and preserves the finest details in the local estimates of the matched blocks. After processing all blocks, the final estimate is the weighted average of all overlapping local block-estimates. To maintain the clarity and precession in this limited-space paper, the reader is referred to the original paper and its web page for the details of the implementation of this innovative filter [31].

F. Results Analysis

Because of the different sensor types and images, each camera produces a different ROC curve and hence various values of $\mathcal{P}_c$ and $\mathcal{R}_c$ for each method. To reach a conclusion under an evaluated method, we average the two metrics $\mathcal{P}_c$ and $\mathcal{R}_c$ across all the cameras, respectively, yielding $\bar{\mathcal{P}}$ and $\bar{\mathcal{R}}$. The results for the aforementioned de-noising operations are listed in Table III. The filters are listed and arranged in order according to their performances. At the top, we can observe the excelling performance of the block-matching and 3D (BM3D) filter with significant improvement that can be seen at both metrics (it was also observed that the excelling performance is consistent across all the 45 cameras unlike any other method studied in this paper). It is respectively followed by the popular wavelet-based (WB) filter. The computationally undemanding total variation (TV) filter, at two iteration steps, falls slightly behind the WB filter. We tested the TV filter with less/more steps; and the results were similar or inferior. The basic operation of adaptive spatial (AS) does not fall much behind the WB and TV filters. The 2-pixel (2P), the context-adaptive interpolator (CAI) and the Perona-Malik diffusion (PMD) come sequentially. We remark that the last three filters have been shown to provide superior performances in the forensic applications in the original papers. We mainly attribute that to the imperfection of the filters at suppressing specific types of noise that can be beneficial in some experimental setups, whereas in our study

we are evaluating the estimation quality of the supposedly unique PRNU fingerprint.

It was pointed out in various works that the filters operate different amount of de-nosing and hence the number of noise residuals L used to estimate the reference PRNU would affect the filters' relative performances. Therefore, we evaluate the filters for $L = 100$ and $L = 150$. We note that only cameras with more than 300 images in Table I and II have been used for those experiments. Also, since the PRNU estimates would give very high accuracy with such large L, we crop the images to size 32×32 . This would allow us to expand the numerical experiments and manifest the relative performances of the filters. The results are listed in Table IV. We can observe that the ranking of the filters persist. But, the performances of some of the filters seem to climb up.

Although our main focus is on the quality of the PRNU estimation, the computational time of the filtering operation is a worth considering aspect—especially with the large number of images usually involved in forensic applications. The CPU time of the implementation of a filter in MATLAB could serve as a good measure of this aspect. The CPU times of the filters are listed in Table V—all the computations are performed on the same machine of Intel Core Duo i7-4770 @ 3.40GHz processor and 16 GB of memory on images of size 512×512. It is highlighted that the strong performance of BM3D comes at the cost of relatively demanding computations. Whereas, the FSTV requires minimal computations. Before we close this section, we remark few notes about our implementations of the filters. The 8-neighbouring pixel approximation of PMD was adopted here. And, we implemented the central difference approximation, as described in the appendix of [26], which is a touch different from the original work [27]. In the TV filters, the forward finite difference approximation of the gradient was used in our implementation.

TABLE III. THE PERFORMANCE RESULTS (IN %) OF THE LISTED FILTERS USING L=50 IMAGES

Method	$\bar{\mathcal{P}}$	$\bar{\mathcal{R}}$
BM3D	82.4	5.1
WB	66.7	7.6
(2S)TV	54.9	8.8
AS	54.9	8.9
2P	40.2	15.6
CAI	24.3	19.7
PMD	16.3	18.1
FSTV	13	24.6

TABLE IV. THE PERFORMANCE (IN %) OF THE LISTED FILTERS USING L=100 AND L=150 FOR AN ALTERED EXPERIMENTAL SETUP.

Method	$\bar{\mathcal{P}}$		$\bar{\mathcal{R}}$	
	$L = 100$	$L = 150$	$L = 100$	$L = 150$
BM3D	86.2	86.2	3.6	3.2
WB	69.1	79	6.5	4.8
(2S)TV	60.1	74	7	5.2
AS	49.4	59.4	9.7	7.8
2P	42.8	58.6	14.0	10.3
CAI	34.0	50.2	16.2	8.8
PMD	20.9	50.5	15.8	10.0
FSTV	11	34.0	25.1	16.8

TABLE V. CPU TIMES OF THE LISTED FILTERING OPERATIONS.

Method	CPU Time (ms)
CAI	4344
BM3D	3155
WB	851
PMD	298
2P	72
AS	44
FSTV	12

V. COMBINING PROCESS

As we mentioned earlier the noise residual of the filtering process (1) contains a considerable amount of random noises that cannot be used in image forensics, as well as partial scene details of the image itself caused by the imperfections of the filtering process—which are referred to by image contamination. Thus, to provide a reliable estimate of the PRNU, the noise residuals of L images taken by the same camera are combined. The underlying model of the noise residual that is implicitly adopted in a lot of work in the literature is given by $\hat{\mathbf{r}} = \mathbf{K}\mathbf{I_0} + \boldsymbol{\theta}$, where $\boldsymbol{\theta} \in \mathbb{R}^{M \times N}$ is a combination of the random noises and image contamination that is independent of $\mathbf{KI_0}$ and has constant mean and variance. Since $\mathbf{I_0}(m,n)$ and $\mathbf{K}(m,n)$ are independent at a pixel location, a pixel-wise average: $\mathbf{R} \coloneqq \sum_{l=1}^{L} \hat{\mathbf{r}}_l / L$, where $\hat{\mathbf{r}}_l$, $l = 1, \dots, L$ are the noise residuals extracted from the images $\mathbf{I}_l$, $l = 1, \dots, L$, respectively, converges to the (scaled and DC-shifted) PRNU with increasing L. Alternative combining approaches have been adopted in the field, which we discuss below.

A. Maximum Likelihood Estimator (MLE)

The work of Chen *et al*. [34] models the extracted noise residual as $\hat{\mathbf{r}} = \mathbf{K}\mathbf{I} + \boldsymbol{\theta}$. The authors accept that the random noises across *all* the L images at a certain pixel location, i.e. $\boldsymbol{\theta}_l(m,n)$, $l = 1, \dots, L$, are (zero-mean, fixed variance) white Gaussian process. Correspondingly, a maximum likelihood estimator can be simply adopted to estimate $\mathbf{K}$:

$$\mathbf{R}_{MLE} \coloneqq \frac{\sum_{l=1}^{L} \hat{\mathbf{r}}_l \mathbf{I}_l}{\sum_{l=1}^{L} (\mathbf{I}_l)^2} \ . \tag{18}$$

The assumption of fixed (random) noise variance per pixel in the L images can be met in uniformly illuminated images taken under controlled conditions.

B. Weighted Averaging (WA)

The variance of the random noise is not constant in all natural uncontrolled images taken by a camera, even for fixed ISO sensitivity. This is due to several reasons among which is the variation of the camera settings such as integration time, shutter speed and focal length at the times of taking the pictures. Relying on this fact, it was proposed in [35] to capitalise on a weighted averaging operation to reduce the estimation error:

$$\mathbf{R}_w \coloneqq \sum_{l=1}^{L} \mathrm{w}_l \hat{\mathbf{r}}_l \ , \tag{19}$$

where w_l is the weight for the l-th image, and it is given by

$$w_l = \frac{1}{\sigma_l^2}\left(\sum_{s=1}^{L}\frac{1}{\sigma_s^2}\right)^{-1}. \quad (20)$$

where σ_l^2 is the variance of the undesirable noise in $\hat{\mathbf{r}}_l$. In [35], the noise variances are estimated using the *difference signal estimation* approach proposed in [36]. Based on the assumption that the PRNU is deterministic and invariant for an individual pixel of a camera from image to another, the random noise component can be obtained by subtracting the noise residual by the $\mathbf{KI_0}$. But since the latter is not known, the estimated PRNU $\mathbf{R}$ is used:

$$\hat{\mathbf{n}}_l \coloneqq \hat{\mathbf{r}}_l - \mathbf{R} \quad (21)$$

Then, the variance estimate $\hat{\sigma}_l^2$ is simply calculated using the following.

$$\hat{\sigma}_l^2 \coloneqq \frac{\sum_{(m,n)}(\hat{\mathbf{n}}_l(m,n) - \bar{n}_l)^2}{MN}, \quad (22)$$

where $\bar{n}_l$ is the mean of the random noise component in the $\hat{\mathbf{r}}_l$:

$$\bar{n}_l \coloneqq \frac{\sum_{(m,n)}\hat{\mathbf{n}}_l(m,n)}{MN}. \quad (23)$$

It was recommended in [35] to divide each image into a number of sub-images, where the noise is rather stationary, and run the above procedure on each sub-image yielding different weights.

C. Results Analysis

The performance results of the maximum likelihood estimator (MLE) and weighted averaging (WA), over the basic averaging approach are considered here. As the wavelet-based filter was originally implemented by the first PRNU-based work [1] and it is still the most popular filter in this forensic field, we use this filter in all the noise residuals combining approaches. For easy interpretation of the results, we benchmark the studied methods against the basic averaging approach. That is, the performance results are calculated through $\dot{\mathcal{P}} = \sum_{c=1}^{c=45}(\mathcal{P}_c - \bar{\mathcal{P}}_{WB})/45$ and $\dot{\mathcal{R}} = \sum_{c=1}^{c=45}(\bar{\mathcal{R}}_{WB} - \mathcal{R}_c)/45$, where $\bar{\mathcal{P}}_{WB}$ and $\bar{\mathcal{R}}_{WB}$ are the overall performance results of the basic averaging with the wavelet-based filter which are equal to 66% and 7.6%, respectively. In fact, for the rest of the paper, we will use the basic averaging with the wavelet-based filter to evaluate the studied techniques and benchmark them accordingly. We can see in Table VI that the MLE and the WA approaches provide clear improvements on the PRNU estimation. But, they seem to deliver similar results to each other.

VI. PRNU Enhancement Techniques

The estimated PRNU can still contain considerable amount of contamination even after combining the noise residuals of a large number of images. Various additional enhancement techniques have been adopted in the literature to improve the purity of the estimated signal, which are described in the next subsections.

TABLE VI. THE PERFORMANCE IMPROVEMENTS (IN %) OF THE NOISE RESIDUALS COMBINING APPROACHES OVER BASIC AVERAGING.

Method	$\dot{\mathcal{P}}$	$\dot{\mathcal{R}}$
MLE	+3.9	+1.4
WA	+4.8	+0.9

A. Removing the Sharing Components (RSC)

The estimated PRNU contains all the components that are systematically present in every image of an individual camera. These components include the sought PRNU and other artifacts that are not unique for a camera, not even for a model or make. These usually appear because of cameras employing the same processing algorithms in their pipelines. Hence these artifacts cannot serve as a reliable forensic tool and must be removed from the estimated PRNU to improve its quality. There are various types of artifacts; although the following two operations were not originally developed in [34] to tackle all the different types, they seem to effectively suppress them [37].

The first step is the 'zero-mean' operation, denoted by $\mathcal{Z}(.)$, where the column average is subtracted from each pixel in the column and then the row average is taken from every pixel in the row. It targets the artifacts induced due to colour interpolation and the row-wise/column-wise operations of processing circuits and sensors.

The second operation is Wiener filtering the PRNU estimate in the Fourier domain. It operates by filtering the magnitude of the Fourier transform, keeping the only noise components. This would result into a flatter frequency spectrum. These can be summarised as

$$\text{real}\left[\mathcal{F}^{-1}\left(\frac{\mathcal{F}(\mathcal{Z}(\mathbf{R}))}{|\mathcal{F}(\mathcal{Z}(\mathbf{R}))|}\left[|\mathcal{F}(\mathcal{Z}(\mathbf{R}))| - \omega(|\mathcal{F}(\mathcal{Z}(\mathbf{R}))|)\right]\right)\right] \quad (24)$$

where $\mathcal{F}(.)$ and $\omega(.)$ are the Fourier transform and the Wiener filtering, respectively. The noise variance in the latter is set as the sample variance of the magnitude of the Fourier transform $|\mathcal{F}(\mathcal{Z}(\mathbf{R}))|$. And, the assumption is that the non-unique artifacts in $|\mathcal{F}(\mathcal{Z}(\mathbf{R}))|$ are locally stationary iid variables with zero mean. Indeed, PRNU estimates constructed from any of the combining techniques can be plugged in above analogously.

B. Phase-Only Operation

Similar to the approaches above, the authors in [38], [39] proposed a method to clear the noise residue in the frequency domain from image contents and non-unique artifacts of JPEG compression, on-sensor signal transfer, sensor design, and colour interpolation. The method counts on the established assumption that the sensor pattern noise is a white noise, and hence it has a flat frequency spectrum. To this end, the noise residuals are whitened first through:

$$\mathbf{Ph}_l \coloneqq \frac{\mathcal{F}(\hat{\mathbf{r}}_l)}{|\mathcal{F}(\hat{\mathbf{r}}_l)|}, \qquad l = 1, \dots, L \quad (25)$$

where $\mathcal{F}(.)$ denotes the Fourier transform as seen above, and hence $\mathbf{Ph}_l$ represents the phase component of the noise residual $\hat{\mathbf{r}}_l$ of the *l*-th image. The phase components are then combined before taking the inverse Fourier transform to yield the PRNU estimate:

$$\mathbf{R}_p \coloneqq \text{real}\left[\mathcal{F}^{-1}\left(\frac{\sum_{l=1}^{L}\mathbf{Ph}_l}{L}\right)\right]. \quad (26)$$

C. Sensor Pattern Noise Enhancer Models

In [40], the author proposed an enhancing technique based on the hypothesis that the stronger a signal component in noise residual is, the more likely that it is associated with strong scene details, and hence the less trustworthy the component should be. Working in conjunction with the wavelet-based de-noising operation, the hypothesis suggests that an improved PRNU can be attained by assigning less weighting factors on strong components of the noise residual in the digital wavelet domain in order to supress the contamination of scene details. To this end, the author proposed five models to be applied. Let the wavelet coefficients of the noise residual be denoted by $\hat{\mathbf{r}}_{\boldsymbol{w}}(i,j), (i,j) \in \mathcal{T}$, where $\mathcal{T}$ is the index set of the wavelet coefficients that depends on the decomposition level. The two models (Model 3 and Model 5 in the original work) that were shown there to deliver the best results are, respectively:

$$\hat{\mathbf{r}}_{M1}(i,j) \coloneqq \begin{cases} 1 - e^{-\hat{\mathbf{r}}_{\boldsymbol{w}}(i,j)}, & 0 \le \hat{\mathbf{r}}_{\boldsymbol{w}}(i,j) \le \alpha \\ (1 - e^{-\alpha}).\, e^{\alpha - \hat{\mathbf{r}}_{\boldsymbol{w}}(i,j)}, & \hat{\mathbf{r}}_{\boldsymbol{w}}(i,j) > \alpha \\ -1 + e^{\hat{\mathbf{r}}_{\boldsymbol{w}}(i,j)}, & -\alpha \le \hat{\mathbf{r}}_{\boldsymbol{w}}(i,j) < 0 \\ (-1 + e^{-\alpha}).\, e^{\alpha + \hat{\mathbf{r}}(i,j)}, & \hat{\mathbf{r}}_{\boldsymbol{w}}(i,j) < -\alpha, \end{cases} \tag{27}$$

and

$$\hat{\mathbf{r}}_{M2}(i,j) \coloneqq \begin{cases} e^{-0.5\hat{\mathbf{r}}_{\boldsymbol{w}}{}^2(i,j)/\alpha^2}, & \hat{\mathbf{r}}_{\boldsymbol{w}}(i,j) \ge 0 \\ -e^{-0.5\hat{\mathbf{r}}_{\boldsymbol{w}}{}^2(i,j)/\alpha^2}, & \hat{\mathbf{r}}_{\boldsymbol{w}}(i,j) < 0, \end{cases} \tag{28}$$

where α is a threshold to be decided by the user. The enhanced noise residuals are then obtained by applying the inverse wavelet transform on the coefficients of (27) and (28). In the original work, these models were developed to suppress the scene contamination in the noise residual of a single uncontrolled test image in camera identification/verification, and they were not applied to the noise residuals used to estimate the reference PRNU (the assumption there is that the camera is available to the analyst, and hence uniformly illuminated images can be taken which contain no scene details to suppress). It was implied in the original work that those models are applied in the pixel domain. This led few researchers to adopt the models directly in the spatial domain in their implementation.

D. Results Analysis

The described enhancement operations are implemented here over the wavelet-based filter with basic averaging, and they are benchmarked against it as we described in Section V. The relative performances are listed in Table VII. Our findings highlight the effectiveness of removing the sharing components (RSC). In our experiments, we observed a certain amount of false correlation in the estimated PRNUs not only between cameras of the same model but also between other camera models. And, the RSC operations seem to efficiently suppress such adversary effect. In the original work of RSC, the Wiener filter is described as 3×3, however, in the authors' implementation they use the variance estimation procedure, described in Subsection IV.A, to obtain the minimum local variance within windows of sizes 3×3, 5×5, 7×7, and 9×9. We adopt the same procedure in our implementation. The other Fourier-based operation, i.e. phase-only operation, also seems to deliver sound performance. However, the improvement is not as significant (or consistent through different cameras) as the RSC results. This is despite that the phase-only operation is applied at each of the L images, whereas the RSC operations are only applied once on the reference PRNU estimate. As the two methods operate in the same Fourier domain, combining the two operations does not improve the results further. Finally, the two enhancing models seem to have rather harmful impact on the performance of the basic approach (the user parameters α in the numerical experiments are set to the optimal values for the addressed image size, as proposed in the original work). An explanation could be seen in (27) and (28); as they might suppress the significant contamination they also magnify the small components which highlight the effect of the shared non-unique artifacts in the estimated PRNUs.

TABLE VII. THE PERFORMANCE IMPROVEMENTS (IN %) OF THE LISTED METHODS OVER THE BASIC APPROACH.

Method	$\hat{\mathcal{P}}$	$\hat{\mathcal{R}}$
RSC	+6.59	+0.9
Phase-Only	+2.5	+0.01
Model1	-9.0	-3.0
Model2	-13.9	-6.6

VII. COMPACT PRNU-BASED FINGERPRINT

In this section, we discuss methods that aim to exploit the information in the estimated PRNU by constructing a modified PRNU-based fingerprint. These methods aim to enhance the accuracy of the forensic application and lessen the computational and storage requirement. The latter is achieved by tapering the size of the PRNU-based fingerprint that is required to be stored and engaged in the forensic calculations. These PRNU enhancing methods are exclusive for camera origin identification applications, and cannot be incorporated in image forgery detections.

A. Significant Components (SC) Only Technique

The authors in [41] proposed to *only* use the large components in the estimated PRNU to cut down the overall random noise. In theory, the large components carry more of the signal of interest in comparison to small components that are mainly random noise. Based on the magnitude, they sort the components (i.e. pixels) of the estimated PRNU in a descending order. Then, the first d largest components are used while masking the rest to yield a new reduced-size PRNU representation $\mathbf{R}_{sc} \in \mathbb{R}^{d \times 1}$. Along with the new PRNU representation, the locations of those significant components in the original PRNU estimate are saved to apply on the other PRNU estimates engaged in the forensic application.

B. Clustering Technique

In [42] and [43], a new system was proposed to suppress the random noises in the reference PRNU estimation by clustering PRNU pixels of comparable values. The method starts by re-arranging the estimated PRNU pixels according to their values in a descending/ascending order into a vector $\mathbf{H} \in \mathbb{R}^{Z \times 1}$ where $Z = MN$ is the size of the PRNU signal. Then, every C pixels are simply averaged to give a PRNU

representation vector $\mathbf{R}_{CL} \in \mathbb{R}^{d\times 1}$, where $d = Z/C$. Along with $\mathbf{R}_{CL}$, a map of the locations of the clustered pixels in the original PRNU is saved and used in the forensic application. In theory, this procedure produces a suppressed-noise reduced-size PRNU representation, which could serve as a more robust fingerprint than its original full-size counterpart.

C. Principal Component Analysis (PCA) based Approach

In [44], the Principal Component Analysis (PCA) has been used to reduce the dimensionality of the PRNU noise and attenuate image content contamination and other undesired noise components. The approach operates on the PRNU estimation of S cameras. After collecting and reshaping the noise residuals of every camera into column vectors of size $Z = MN$, the technique forms a covariance matrix from the SL noise residuals. Then, the PCA is performed by obtaining the eigenvectors of the mean-centered covariance matrix to convert the Z-dimension noise residuals space into a smaller orthogonal space. The underlying idea is that the energy of the noise residuals characterising the reference PRNU is concentrated in a small subspace of the attained orthogonal space, while the (image-dependent) noise energy that represents undesirable components is spread over the whole space. Therefore, by preserving only the most important subspace (characterized by the d eigenvectors which are associated to the most significant eigenvalues that correspond to 99% of the variance explained by the eigenvectors) and projecting the re-arranged noise residuals of a camera into the objective subspace, we obtain enhanced noise residual representations $\mathbf{P}_l \in \mathbb{R}^{d\times 1}$, $l = 1, \dots, L$, where $d \ll Z$. The reference PRNU representation is then obtained by component-wise averaging the $\mathbf{P}_l$, $l = 1, \dots, L$.

D. Fingerprint Compression

Unlike the other studied methods in this paper that seek to improve the accuracy of the PRNU estimation, the aim of the addressed methods herein is to ease the potentially burdening aspects of storage and computations of the PRNU signal in its applications. The PRNU compression techniques are visited here for their relation to the studied methods. Indeed, the PRNU signal cannot be compressed using standard methods such as JPEG because of the signal's lack of redundancy. In [45], the authors proposed to represent the PRNU signal in a binary-quantization form, i.e. 1-bit representation per pixel. And, it was analytically shown that the reduction in the accuracy of the PRNU matching is insignificant. A more thorough study in fingerprint compression can be found in [46] based on random projection. The idea is to project the PRNU estimate, reshaped into the column vector $\mathbf{H} \in \mathbb{R}^{Z\times 1}$ where $Z = MN$, using a random matrix $\mathbf{\Phi} \in \mathbb{R}^{d\times Z}$ where $d < Z$, to yield the PRNU representation:

$$\mathbf{R}_{CM} \coloneqq \mathbf{\Phi} * \mathbf{H} \tag{29}$$

of reduced size, i.e. $\mathbf{R}_{CM} \in \mathbb{R}^{d\times 1}$. Herein, $*$ designates matrix multiplication. The same random matrix is used to project the other PRNU estimates in the application. The considered random matrices in [46] are the most-studied Gaussian random matrices, which are practically addressed using circulant matrices (the requirements on the suitable $\mathbf{\Phi}$ are thoroughly studied in the field of compressive sensing [47]). The key idea is based on from Johnson-Lindenstrauss: if points in a vector space are projected on a suitable lower dimensional space then the distances are approximately preserved [48]. And, since PRNU fingerprints from different cameras are highly uncorrelated and thus the angles (equivalent to the distance herein) between them are wide, the angles between the compressed PRNUs are preserved to be wide. Inspired by [45] and the 1-bit compressive sensing [49], the authors also considered binarizing the compressed PRNU, which could be seen as a generalized case of [45] with identity projecting matrix. Theoretical results concerning the compressed PRNU matching accuracy show insignificant reduction.

TABLE VIII. THE PERFORMANCE IMPROVEMENTS (IN %) OF THE LISTED METHODS OVER THE BASIC APPROACH.

Method	$\hat{\mathcal{P}}$	$\hat{\mathcal{R}}$
SC Only	-6.7	-1.4
Clustering	+0.1	+0.3
PCA	+6.9	+2.1
1-Bit	-5.5	-2.6
CD	+6.4	-0.3
DA	-40.3	-15.7

E. Results Analysis

It is evident in Table VIII that the technique of keeping the significant components (SC) only does not benefit the PRNU classification (in this implementation, the 20% largest components in magnitude are kept). This indicates that there is information in the small components of the estimated PRNU that would be adversary to flush. The clustering approach is evaluated on the RSC-PRNU as proposed and highlighted in the original work. That is, the sharing components are removed before applying the clustering technique. Benchmarked against the RSC results, the clustering technique seems to provide no improving effect on the classification of the PRNU estimates. We recall and attribute these findings to the fact that the images used here in the estimation of the reference PRNU are of random nature as opposed to the fixed illumination images used in the numerical experiments of the original work. Nonetheless, considering that the clusters sizes are set to 64 pixels here, the clustering technique would constitute an excellent PRNU compression tool. In contrast, we can see that the principle component analysis (PCA) technique can bring considerable improvements. We note that PCA technique is trained against the interclass and intraclass images prior to estimating their final PRNU signals. This facility is not always available in forensic applications. Finally, the results of the 1-Bit representation of the PRNU show the expected slight reduction in the performance.

VIII. MODIFIED PRNU ESTIMATION PROCEDURES

There are other research developments in PRNU estimation that cannot be categorised in one of the sections above. They operate a modified strategy to the standard procedure in Fig. 1.

A. Colour-Decoupling (CD) Approach

The work [50] takes into the account the characteristics of the colour filter array (CFA) structure. That is, the lenses of

most cameras let through rays of the three colour components, but for every pixel only the rays of one colour are passed through the CFA and subsequently captured by the sensor pixel. Then, a colour interpolation algorithm generates the other two colour components of every pixel. The artificial colours obtained through the colour interpolation process (known is de-mosaicking) are not physically acquired from the scene by the sensor. Therefore, it is assumed that the PRNU extracted from the physical components should be more reliable. The almost universal CFA in cameras is the Bayer filter where pixels in odd/even rows alternate between green and red, and pixels in even/odd rows alternate between blue and green. Based on this assumption, the authors proposed a new strategy that first decomposes each image into 4 sub-images (interlaced along the two dimensions) and then extracts the PRNU from each sub-image. The PRNU noises of the sub-images are then assembled to obtain the final one. This method aims to prevent the interpolation noise from propagating into the PRNU estimation of the physically captured pixels. According to our numerical findings shown in Table VIII, the promising idea enhances the PRNU results notably.

B. Direct Average (DA) Technique

Given the deterministic nature of the PRNU, as opposed to the other random noise components, the PRNU can be estimated by simply averaging a very large number of images without any de-noising step. The work in [51] counts on this concept to estimate the PRNU of an available camera by capturing a very large number of uniform random noise images displayed on a high-resolution monitor. Using the model of $\mathbf{I}$ in (6), the pixel-wise mean of L images is given by:

$$\frac{1}{L}\sum_{l=1}^{L}\mathbf{I}_l = \frac{(1+\mathbf{K})}{L}\sum_{l=1}^{L}\mathbf{I}_{0,l} + \frac{1}{L}\sum_{l=1}^{L}\mathbf{N}_{t,l}. \tag{30}$$

When L tends to infinity, the last term will be a negligible constant. The channel gains $\{g_l\}_{l=1}^{L}$ and the image illuminations $\{\mathbf{Y}_l(m,n)\}_{l=1}^{L}$ are supposed to be mutually independent in $\{\mathbf{I}_{0,l}(m,n)\}_{l=1}^{L}$. Hence, the expected value:

$$E\left[\frac{1}{L}\sum_{l=1}^{L}\mathbf{I}_{0,l}(m,n)\right] = E[g^{\gamma}]E[\mathbf{Y}^{\gamma}(m,n)]. \tag{31}$$

Since g is a global variable for an image that is independent of the pixel location, and the random images are displayed on the monitor with constant mean, then the expectation of (31) is constant across all the pixels. As $\mathbf{K}$ is zero mean, the PRNU can be simply extracted by removing the DC component from (30) when L tends to infinity. In practice, we deal with a limited number of images and the conditions are not ideal. Hence, (31) does not hold strictly. The work in [51] takes the logarithm of the mean of a very large number of images L:

$$\ln\left(\frac{1}{L}\sum_{l=1}^{L}\mathbf{I}_l\right) \cong -\ln L + \ln(1+\mathbf{K}) + \ln\left(\sum_{l=1}^{L}\mathbf{I}_{0,l}\right). \tag{32}$$

Carrying out MacLaurin expansion

$$\ln\left(\frac{1}{L}\sum_{l=1}^{L}\mathbf{I}_l\right) \cong -\ln L + \mathbf{K} + O(\mathbf{K^2}) + \ln\left(\sum_{l=1}^{L}\mathbf{I}_{0,l}\right), \tag{33}$$

and since the values of $\mathbf{K}$ are very small, the higher order term $O(\mathbf{K^2})$ is of an insignificant value and can be ignored. Taking in consideration the linear manipulations in the digital camera pipeline, the authors model $\ln\left(\sum_{l=1}^{L}\mathbf{I}_{0,l}\right)$ as the auto-regressive and moving average (ARMA), where $\mathbf{K}$ is the additive white Gaussian noise, and estimate $\mathbf{K}$ using

$$\mathbf{R}_{DA} \coloneqq \ln\left(\sum_{l=1}^{L}\mathbf{I}_l\right) - \omega\left(\ln\left(\sum_{l=1}^{L}\mathbf{I}_l\right)\right) \tag{34}$$

where $\omega(.)$ is the standard 3 x 3 Wiener filter. Thus, the filtering operation is only applied once to estimate the reference PRNU, as opposed to the standard procedure. This approach has been tested in our evaluation system, and the results are shown in Table VIII. Our experiments indicate a considerable inferiority of the direct averaging (DA) approach as opposed to the standard procedure when natural uncontrolled images are used.

IX. Similarity Measures

As we mentioned in previous sections, PRNU-based forensic applications generally rely on measuring the similarity between the estimated PRNU signals in a binary hypothesis test for decision-making. In this section, we outline the various similarity measures used in PRNU-based forensics because of its close relation to the studied methods. Since most of the developments in the similarity measures are in camera identification/verification application (where the reference PRNU $\mathbf{R}$ and the noise residual $\hat{\mathbf{r}}_q$ of the query image $\mathbf{I}_q$ are compared), we present them in terms of this application. Let $\mathbf{X}$ represent $\mathbf{R}$, $\mathbf{R}_w$, $\mathbf{R}_p$, or the product $\mathbf{R}_{MLE}\mathbf{I}_q$. The basic measure is the normalized cross-correlation (after mean centering the two signals):

$$\rho \coloneqq \frac{\mathbf{X}\odot\hat{\mathbf{r}}_q}{\|\mathbf{X}\|\|\hat{\mathbf{r}}_q\|}, \tag{35}$$

where $\odot$ and $\|.\|$ are the dot product and norm operations, respectively. A development was proposed in [52] where the aim is to eliminate the effect of contamination of the two compared signals with the same periodic noise that could adversely increase their correlation. It is referred to by the peak-to-correlation energy (PCE), and based on the circular cross-correlation:

$$\mathcal{C}(x,z) \coloneqq \frac{1}{MN}\sum_{(m,n)}\mathbf{X}(m,n)\,\hat{\mathbf{r}}_q(m+x \bmod M, n + z \bmod N),$$
$$x = 0,\dots,M-1,\ \ z = 0,\dots,N-1, \tag{36}$$

it is given in

$$\rho_e \coloneqq \mathrm{sign}[\mathcal{C}(0,0)]\frac{\mathcal{C}^2(0,0)}{\frac{1}{MN-|\mathcal{A}|}\sum_{(x,z),\ (x,z)\notin\mathcal{A}}\mathcal{C}^2(x,z)}, \tag{37}$$

where $\mathcal{A}$ is a small area around (0,0) and $|\mathcal{A}|$ is its cardinality. The sign in (37) was not included in the first introduction of PCE; it was inserted in their later work to eliminate the false alarm of the negative correlations. The same idea appeared in [39] by considering $\mathcal{C}(0,0)$ (and a square root of the dominator) rather than its squared value to retain its sign; the authors referred to it by the correlation over circular cross-correlation norm (CCN).

A more optimal and complex similarity measure was pursued by Chen *et al.* in [34]. It begins with a new model for

$\hat{\mathbf{r}}_q$. That is, a pixel-wise multiplicative shaping factor $\mathbf{T} \in \mathbb{R}^{M\times N}$ is introduced to capture the de-noising process imperfection and other operations on the PRNU signal. And, the random noises and image contamination are modeled as coloured Gaussian noise:

$$\hat{\mathbf{r}}_q = \mathbf{TX} + \boldsymbol{\theta}, \tag{38}$$

where $\boldsymbol{\theta} \in \mathbb{R}^{M\times N}$ is a matrix of independent Gaussian variables with unequal variances. The work divides the noise residual signal into B non-overlapping, equal sized blocks. The pixels (m,n) within the b-th block , $b = 1, \dots, B$ are allocated a fixed $\mathbf{T}(m,n)$ denoted by T_b **,** and their noise $\boldsymbol{\theta}(m,n)$ are assumed to have a fixed variance σ_b^2. The similarity measure is the generalized matched filter that is given in:

$$\rho_m \coloneqq \frac{\sum_b \hat{T}_b/\hat{\sigma}_b^2 \,(\mathbf{X}_{\boldsymbol{b}} \odot \hat{\mathbf{r}}_{\boldsymbol{b}})}{\sqrt{\sum_b \left\|\hat{T}_b \mathbf{X}_{\boldsymbol{b}}\right\|^2/\hat{\sigma}_b^2}\sqrt{\sum_b \|\hat{\mathbf{r}}_{\boldsymbol{b}}\|^2/\hat{\sigma}_b^2}}, \tag{39}$$

where $\hat{\mathbf{r}}_{\boldsymbol{b}}$ and $\mathbf{X}_{\boldsymbol{b}}$ are the noise residual from the tested image and the PRNU term $\mathbf{X}$ within the b-th block, respectively. $\hat{T}_b$ and $\hat{\sigma}_b^2$ represent estimates of T_b and σ_b^2 respectively, which are obtained from the normalized cross-correlation within a block under the positive hypothesis:

$$\begin{aligned}\rho_b &\coloneqq \frac{\mathbf{X}_{\boldsymbol{b}} \odot \hat{\mathbf{r}}_{\boldsymbol{b}}}{\|\mathbf{X}_{\boldsymbol{b}}\|\|\hat{\mathbf{r}}_{\boldsymbol{b}}\|} \\ &= \frac{T_b\|\mathbf{X}_{\boldsymbol{b}}\|^2 + \mathbf{X}_{\boldsymbol{b}} \odot \boldsymbol{\theta}_{\boldsymbol{b}}}{\|\mathbf{X}_{\boldsymbol{b}}\|\sqrt{{T_b}^2\|\mathbf{X}_{\boldsymbol{b}}\|^2 + \|\boldsymbol{\theta}_{\boldsymbol{b}}\|^2 + 2T_b\mathbf{X}_{\boldsymbol{b}} \odot \boldsymbol{\theta}_{\boldsymbol{b}}}},\end{aligned} \tag{40}$$

with $\boldsymbol{\theta}_{\boldsymbol{b}}$ being white process that is independent of $\mathbf{X}_{\boldsymbol{b}}$, the term $\mathbf{X}_{\boldsymbol{b}} \odot \boldsymbol{\theta}_{\boldsymbol{b}}$ will be small and can be ignored:

$$\rho_b \approx \frac{1}{\sqrt{1 + C\sigma_b^2/{T_b}^2\|\mathbf{X}_{\boldsymbol{b}}\|^2}}, \tag{41}$$

where C is the number of pixels in each block. And, by using a predictor of ρ_b to give $\hat{\rho}_b$, we have estimates of those parameters:

$$\hat{T}_b = |\hat{\rho}_b|\|\hat{\mathbf{r}}_{\boldsymbol{b}}\|/\|\mathbf{X}_{\boldsymbol{b}}\| , \tag{42}$$

$$\hat{\sigma}_b^2 = \left(1 - {\hat{\rho}_b}^2\right)\|\hat{\mathbf{r}}_{\boldsymbol{b}}\|/C . \tag{43}$$

Chen *et al*. developed a simple predictor of ρ_b based on features derived from blocks of a few diverse images. They noted that other standard predictors and features provided the same performance. A more flexible, pixel-wise weighting approach based on similar features was proposed in [53]. In [54], it was proposed to only use the significant blocks of the query noise residual. The significance of a block is measured by its signal to noise ratio (SNR). Indeed, the signal in SNR is the PRNU noise part that we seek, whilst the noise refers to all the other noise components and image contamination. The SNR of the b-th block is approximated using

$$\mathrm{SNR}_b \coloneqq \frac{\left\|\hat{T}_{\boldsymbol{b}}\mathbf{X}_{\boldsymbol{b}}\right\|^2}{C\hat{\sigma}_b^2}, \tag{44}$$

Then, the SNR values of all the blocks are sorted and only the blocks with the largest SNR values are used:

$$\rho_s \coloneqq \frac{\sum_{b,b\in\mathcal{H}} \hat{T}_b/\hat{\sigma}_b^2 \,(\mathbf{X}_{\boldsymbol{b}} \odot \hat{\mathbf{r}}_{\boldsymbol{b}})}{\sqrt{\sum_{b,b\in\mathcal{H}}\left\|\hat{T}_b\mathbf{X}_{\boldsymbol{b}}\right\|^2/\hat{\sigma}_b^2}\sqrt{\sum_{b,b\in\mathcal{H}}\|\hat{\mathbf{r}}_{\boldsymbol{b}}\|^2/\hat{\sigma}_b^2}} , \tag{45}$$

where $\mathcal{H}$ is the set of the indices of the most significant blocks.

Despite that (41) represents the optimal detector, PCE is the most favorite detection statistics for the two facts. First, the assumption on the models to derive the optimal detector may not be satisfied. Second, PCE can facilitate selecting the decision threshold to achieve the sought probability of false detection.

V. CONCLUSION

In this paper, we introduced a systematic comparative analysis of all the techniques concerned with the estimation of PRNU noise. In order to conduct a profound study, we categorised the techniques based on their roles in the PRNU estimation procedure and analysed each category correspondingly. We created a large database of 45 cameras with effectively over 2.2 million test images for our numerical evaluation; the relatively large experiments were necessary given the variant performance of the techniques across cameras and images. The carefully selected performance metrics were adequate to benchmark the techniques and provide a conclusive study. Our findings provided some concrete conclusions whilst others can be extrapolated. We hope that the presented results can support the research community in digital forensics in general and PRNU-based image forensics in particular. With some practical aspects considered here along with our insight, we hope that this paper would benefit forensic practitioners with sharp implementation decisions.

REFERENCES

[1] J. Lukas, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," *IEEE Trans. on Information Forensics and Security*, vol. 1, no. 2, pp. 205- 214, Jun. 2006.

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

[3] T. H. Thai, R. Cogranne, F. Retraint "Camera Model Identification Based on the Heteroscedastic Noise Model, " *IEEE Trans. on Image Processing*, vol. 23, pp. 250-263, 2014.

[4] A. Swaminathan, M. Wu, and K. J. R. Liu, "Nonintrusive component forensics of visual sensors using output images," *IEEE Trans. on Information Forensics and Security*, vol. 2, pp. 91-106, 2007.

[5] M. Kharrazi, H. T. Sencar, and N. Memon, "Blind source camera identification," *in Proc. of IEEE Int. Conf. on Image Processing (ICIP)*, Singapore, 2004, Vol. 1, pp. 709 – 712.

[6] Z. J. Geradts, J. Bijhold, M. Kieft, K. Kurosawa, K. Kuroki, and N. Saitoh, "Methods for identification of images acquired with digital cameras," *in Proc. of Enabling Technologies for Law Enforcement and Security*, Feb. 2001, vol. 4232, pp. 505–512.

[7] K. San Choi, E. Y. Lam and K. K. Y. Wong, "Source camera identification using footprints from lens aberration," *in Proc. of SPIE Conf. in Digital Photography II*, vol. 6069, pp. 172–179, 2006.

[8] Y. Sutcu, S. Bayram, H. T. Sencar, and N. Memon, "Improvements on sensor noise based source camera identification," *in Proc. of IEEE Conf. on Multimedia and Expo*., July 2007, pp. 24–27

[9] T. Gloe, E. Franz, and A. Winkler, "Forensics for flatbed scanners" *in Proc. of SPIE conf. in Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX* , San Jose, CA, Jan. 29–Feb. 1, 2007, vol. 6505, pp. 1I–1J.

[10] N. Khanna, A. K. Mikkilineni, G. T. C. Chiu, J. P. Allebach, and E. J. Delp, III, "Scanner identification using sensor pattern noise," in *Proc.*

of SPIE Conf. in Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX, San Jose, CA, 2007, vol. 6505, pp. 1K–1L.

[11] A. C. Popescu and H. Farid, "Exposing digital forgeries by detecting traces of resampling", *IEEE Trans. on Signal Processing*, vol. 53, pp. 758–767, 2005.

[12] A. C. Popescu and H. Farid, "Exposing digital forgeries in color filter array interpolated images", *IEEE Trans. on Signal Processing*, vol. 53, pp. 3948–3959, 2005.

[13] M. K. Johnson and H. Farid, "Exposing digital forgeries by detecting inconsistencies in lighting", *in Proc. of ACM Workshop in Multimedia and Security*, New York, 2015, pp. 1–9.

[14] Z. Lin, R. Wang, X. Tang, and H.-Y. Shum, "Detecting doctored images using camera response normality and consistency," *in Proc. of IEEE Conf. on Computer Vision Pattern Recognition*, 2005, vol.1, pp. 1087–1092.

[15] W. Chen and Y. Shi, "Image splicing detection using 2D phase congruency and statistical moments of characteristic function," *in Proc. of SPIE Conf. on Electronic Imaging, Security, Steganography, Watermarking of Multimedia Contents IX*, San Jose, CA, 2007, vol. 6505, pp. 0R–0S.

[16] M. Goljan, J. Fridrich, and T. Filler, "Large scale test of sensor fingerprint camera identification," *in Proc. of SPIE conf. in Media Forensics and Security*, San Jose, CA, Jan. 2009, vol. 7254, pp. 1–12.

[17] T. Gloe and R. Böhme, "The 'Dresden image database' for benchmarking digital image forensics," *in Proc. of ACM Symposium Applied Computing*, Mar. 2010, vol. 2, pp. 1585–1591.

[18] Dresden Image Database, (Technische Unversitaet Dresden, Dresden). http://forensics.inf.tu-dresden.de/ddimgdb. Accessed 1 May 2015.

[19] G.E. Healey and R. Kondepudy, "Radiometric CCD camera calibration and noise estimation*," IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 16, no. 3, pp. 267-276, 1994.

[20] M. Mihcak, I. Kozintsev, and K. Ramchandran, "Spatially adaptive statistical modeling of wavelet image coefficients and its application to denoising," *in Proc. of IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, 1999 , pp. 3253–3256.

[21] X. Kang, J. Chen, K. Lin, and A. Peng, "A context-adaptive SPN predictor for trustworthy source camera identification," *EURASIP Journal on Image and Video Processing*, no. 1, pp. 19–30, 2014.

[22] G. Wu, X. Kang, and K. J. R. Liu, "A context adaptive predictor of sensor pattern noise for camera source identification," *IEEE Int. Conf. on Image Processing (ICIP)*, 2012, pp.237-240.

[23] M. Al-Ani, F. Khelifi, A. Lawgaly, and A. Bouridane, "A novel image filtering approach for sensor fingerprint estimation in source camera identification," *in Proc. of IEEE Conf. on Advanced Video and Signal Based Surveillance (AVSS)*, 2015, pp.1-5.

[24] J. Cooper, "Improved photo response non-uniformity (PRNU) based source camera identification," *Forensic Science International*, pp. 132–141, 2013.

[25] J. S. Lee, "Digital image enhancement and noise filtering by use of local statistics," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 2, pp. 165–168, 1980.

[26] W. Van Houten and Z. Geradts, "Using anisotropic diffusion for efficient extraction of sensor noise in camera identification," *Journal of Forensic Sciences*, vol. 57, no. 2, pp. 521–527, 2012.

[27] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12, no. 9, pp. 629-639, 1990.

[28] L. Rudin, S. Osher and E. Fatemi, "Nonlinear total variation based noise removal algorithms, " *Physica D*, vol. 60, pp. 259–68, 1992.

[29] F. Gisolf, A. Malgoezar, T. Baar and Z. Geradts, "Improving source camera identification using a simplified total variation based noise removal algorithm," *Digital Investigation*, vol. 10, pp. 207–214, 2013.

[30] L.I. Rudin and S. Osher, "Total variation based image restoration with free local constraints," *in Proc. of IEEE conf. in image processing*, 1994, vol. 1, pp. 31–35.

[31] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering," *IEEE Trans. on Image Processing*, vol.16, no.8, pp. 2080-2095, 2007.

[32] A. Cortiana, V Conotter, G Boato, and FGB de Natale, "Performance comparison of denoising filters for source camera identification, *" in Proc. of SPIE conf. on Media Watermarking, Security, and Forensics III*, San Jose, USA, Jan. 2011, vol. 7880, pp. 778007/1 – 778007/6.

[33] G. Chierchia, S. Parrilli, G. Poggi, C. Sansone, and L. Verdoliva, "On the influence of denoising in PRNU based forgery detection," *in Proc. of ACM workshop on Multimedia Forensics, Security and Intelligence*, 2010, pp. 117–122.

[34] M. Chen, J. Fridrich, M. Goljan, and J. Lukáš, "Determining image origin and integrity using sensor noise," *IEEE Trans. on Information Forensics and Security*, vol. 3, no. 1, pp. 74–90, 2008.

[35] A. Lawgaly, F. Khelifi, and A. Bouridane, "Weighted averaging-based sensor pattern noise estimation for source camera identification," *in Proc. of IEEE Conf. on Image Processing (ICIP)*, Paris, 2014, pp. 5357–5361.

[36] E. Laciar and R. Jane, "An improved weighted signal averaging method for high-resolution ECG signals," *in Proc. of IEEE Conf. on Computers in Cardiology, Rotterdam*, vol. 1, 2001, pp. 69-72, 2001.

[37] T. Gloe, S. Pfennig, and M. Kirchner, "Unexpected artefacts in PRNU-based camera identification: a 'Dresden image database' case-study," *in Proc. of ACM Conf. on Multimedia and Security*, 2012, pp. 109–114.

[38] X. Kang, Y. Li, Z. Qu, and J. Huang, "Enhancing source camera identification performance with a camera reference phase sensor pattern noise," *IEEE Trans. on Information Forensics and Security*, vol. 7, no. 2, 2012.

[39] X. Kang, Y. Li, Z. Qu, and J. Huang, "Enhancing ROC performance of trustworthy camera source identification," *in Proc. of SPIE Conf. on Media Watermarking, Security, Forensics XIII*, San Francisco, CA, 2011, vol. 7880, pp. 778009/1 – 778009/11.

[40] C.-T. Li, "Source camera identification using enhanced sensor pattern noise," *IEEE Trans. on Information Forensics and Security*, vol. 5, no. 2, pp. 280–287, 2010.

[41] Y. Hu, B. Yu, and C. Jian, "Source camera identification using large components of sensor pattern noise," *in Proc. of IEEE Conf. Computer Science Applications*, Jeju, South Korea, 2009, pp. 1–5.

[42] Y. Tomioka, Y. Ito, and H. Kitazawa, "Robust digital camera identification based on pairwise magnitude relations of clustered sensor pattern noise," *IEEE Trans. on Information Forensics and Security*, vol. 8, no. 12, pp. 1986-1995, 2013.

[43] Y. Tomioka and H. Kitazawa, "Digital camera identification based on the clustered pattern noise of image sensors," *in Proc. of IEEE Conf. on Multimedia and Expo.*, Jul. 2011, pp. 1–4.

[44] R. Li, Y. Guan, and C-T Li, "PCA-based denoising of sensor pattern noise for source camera identification," *in Proc. of IEEE conf. in Signal and Information Processing*, China, 2014, pp. 436-440.

[45] S. Bayram, H. Sencar, and N. Memon, "Efficient sensor fingerprint matching through fingerprint binarization," *IEEE Trans. on Information Forensics and Security*, vol. 7, pp. 1404–1413, 2012.

[46] D. Valsesia, G. Coluccia, T. Bianchi, and E. Magli, "Compressed fingerprint matching and camera identification via random projections," *IEEE Trans. on Information Forensics Security*, vol. 10, no. 7, pp. 1472–1485, 2015.

[47] D. L. Donoho, "Compressed sensing," *IEEE Trans. on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[48] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, 2008.

[49] L. Jacques, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk, "Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors," *IEEE Trans. on Information Theory*, vol. 59, no. 4, pp. 2082–2102, 2013.

[50] C.-T. Li and Y. Li, "Color-Decoupled Photo Response Non-Uniformity for Digital Image Forensics," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 22, no. 2, pp. 260-271, 2012.

[51] Z. Qu; X. Kang; J. Huang; Y. Li, "Forensic sensor pattern noise extraction from large image data set," *in Proc. of IEEE Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2013, pp. 3023-3027.

[52] M. Goljan and J. Fridrich, "Digital camera identification from images—Estimating false acceptance probability," *in Proc. of Int. Workshop Digital Watermarking*, Busan, Korea, 2008, pp 454-468.

[53] L.-H. Chan, N.-F. Law, and W.-C. Siu, "A confidence map and pixel-based weighted correlation for PRNU-based camera identification," *Digital Investigation*, vol. 10, no. 3, pp. 215–225, 2013.

[54] B.-B. Liu, Y. Hu, and H.-K. Lee, "Source camera identification from significant noise residual regions," *in Proc. of IEEE Conf. on Image Processing (ICIP)*, Hong Kong, 2010, pp. 1749–1752.