



Source camera attribution via PRNU emphasis: Towards a generalized multiplicative model

Samuel Fernández-Menduiña, Fernando Pérez-González*, Miguel Masciopinto

atlanTTic Research Center, Signal Processing in Communications Group, 36310, Vigo, Spain

ARTICLE INFO

Keywords:

PRNU
Camera response function
Sensor fingerprint
Sensor pattern noise
Source camera attribution
Image forensics

ABSTRACT

The photoresponse non-uniformity (PRNU) is a camera-specific pattern, which acts as unique fingerprint of any imaging sensor and thus is widely adopted to solve multimedia forensics problems such as device identification or forgery detection. Customarily, the theoretical analysis of this fingerprint relies on a multiplicative model for the denoising residuals. This setup assumes that the nonlinear mapping from scene irradiance to preprocessed luminance, that is, the composition of the Camera Response Function (CRF) with the digital preprocessing pipeline, is a gamma correction. However, this assumption seldom holds in practice. In this paper, we improve the multiplicative model by including the influence of this nonlinear mapping, termed PRNU emphasis, on the denoising residuals. On the theoretical side, we conduct first an exploratory analysis to show that the response of typical cameras deviates from a gamma correction. We also propose a regularized least squares estimator to measure this effect. On the practical side, we argue that the PRNU emphasis is especially beneficial for a source camera attribution problem with cropped images. We back our argument with an extensive empirical evaluation using different denoisers and both compressed and uncompressed images. This new model will pave the way to future PRNU estimators and detectors.

1. Introduction

The photoresponse non-uniformity (PRNU) is a noise-like pattern intrinsically embedded into digital images due to manufacturing imperfections in camera sensors. It qualifies as a unique device fingerprint for several multimedia forensics problems, including forgery detection and source device attribution [1]. Since the pattern is buried under the scene content, image processing plays a key role in extracting and handling the PRNU [2]. In this sense, although data-driven algorithms are promising [3,4], model-driven strategies still prevail.

In the research literature, the multiplicative model of the denoising residuals [5], that is, the difference between the original and the denoised images, is well-established. Given a large number of pixels, empirical results [6] sustain that it leads to almost optimal PRNU estimators and detectors [7]. In fact, it has remained essentially unchallenged since its inception; as argued by Al-Ani and Khelifi [8], “the model is more or less adopted in all the existing PRNU-based techniques [...]”. Nevertheless, there is margin for improvement when information is scarce. In this vein, the authors of [9] showed with synthetic signals that some of the predictions of the multiplicative model do not conform with the empirical results. Motivated by these findings, we put forward a new model based on the theoretical analysis of the *Transfer Curve*,

a term we use to denote the composition of the Camera Response Function (CRF) with the digital preprocessing pipeline.

The CRF is the mapping from image irradiance to image brightness. Evidence suggests that the response function “can vary significantly from a gamma curve” [10]. Moreover, in practice, the digital preprocessing pipeline, including the ISO and demosaicing, affects the captured image. Yet, as we will show, the multiplicative model is only valid when this transfer curve is a pure gamma correction.

In this paper, we show how to derive a generalized model accounting for this composition of functions. This generalization includes the multiplicative model presented in [5] as a special case. The resulting expression shows that an emphasis function weights the pixels of the PRNU according to the brightness level of the corresponding pixel in the denoised image, but the dependence with this level may be nonlinear. Hence, we call this effect *PRNU emphasis*; we also propose a method to estimate it from natural images.

We note that a related idea was proposed by Li under the name “SPN enhancers” [11]. However, we depart from a model for the captured image, whereas SPN enhancers were derived from a hypothesis about the relation between the denoising residual and the scene content. The result is also different: SPN enhancers act on the denoising residual under test, while PRNU emphasis may be applied to the denoised

* Corresponding author.

E-mail addresses: sfmenduiña@gts.uvigo.es (S. Fernández-Menduiña), fperez@gts.uvigo.es (F. Pérez-González), mmasciopinto@gts.uvigo.es (M. Masciopinto).

images in both training and testing splits. In any case, our model is backwards compatible with all the techniques that rely on the denoised image, so it can be used in conjunction with SPN enhancers.

On the practical side, we consider a source camera attribution (SCA) problem, in which we try to determine whether a given test image was taken with a specific camera device. In this work, we focus on SCA with cropped images because using the wrong model is particularly prejudicial, as we will show. To address the SCA problem, an estimate of the PRNU is necessary. This statistic results from the combination of L residuals, obtained after denoising a set of images, all of them known to be taken with the same camera. The denoising filter removes part of the image content, which benefits the subsequent operations. Regarding the combination step, one of the cornerstone discoveries was the maximum likelihood estimator (MLE) of the PRNU, introduced by Chen et al. [5]. After computing this quantity, Goljan and Fridrich [12] proposed a two step process to decide whether a test image (through its corresponding denoising residual) was taken with a given camera. In their method, the test residual is first aligned with the fingerprint estimate using a correlation based approach, namely by taking the maximum of the normalized correlation coefficient. Then, the peak to correlation energy (PCE) [12] is used as the detection statistic. In this paper, we show that the PRNU emphasis affects both the PRNU estimation and the detection processes.

To validate our proposal, we estimate the PRNU emphasis from real images considering different setups, such as JPEG compression and different denoising algorithms. In all our experiments the result differs from the linear function assumed by the multiplicative model. Then, we address an SCA problem with cropped images using the estimated PRNU emphasis. We also propose and evaluate a fixed emphasis function, which requires less computations. The results evidence the benefits of including the PRNU emphasis in the source camera attribution process.

1.1. Notation and structure

Uppercase bold letters, such as $\mathbf{A}$ or Φ , denote matrices. Lowercase bold letters, such as $\mathbf{a}$, denote vectors. We denote the n th component of the vector $\mathbf{a}$ by $[\mathbf{a}]_n$. We use $\mathbf{1}$ and $\mathbf{0}$ to denote the vector of all-ones and all-zeros, respectively. Regular letters, either Latin or Greek, denote scalar values.

The rest of this paper is organized as follows. In Section 2, we propose our generalized multiplicative model and present the PRNU emphasis. We also present the PRNU estimator and the detector to address the SCA problem with cropped images, detailing how to take advantage of the emphasis function. Different methods to estimate the PRNU emphasis are described in Section 3. We provide our experimental results in Section 4, including a discussion of the peculiarities of the SCA problem with cropped images in Section 4.2. Finally, we draw some conclusions and future research lines in Section 5.

2. Generalizing the PRNU model

In this section, we show that the transfer curve induces an emphasis function in the PRNU. In this context, we first review the derivation of the multiplicative model, which is based on the Taylor expansion of the gamma-correction. After formalizing the concept of transfer curve, we apply a similar Taylor expansion to justify the presence of the PRNU emphasis. We also derive a PRNU estimator that accounts for this emphasis function. Finally, we modify the process of SCA with cropped images to include the emphasis function.

2.1. Multiplicative model revisited

Assume that we have access to a dataset of L images from the same device. In their work, Chen et al. [1,5] argued that the output of the

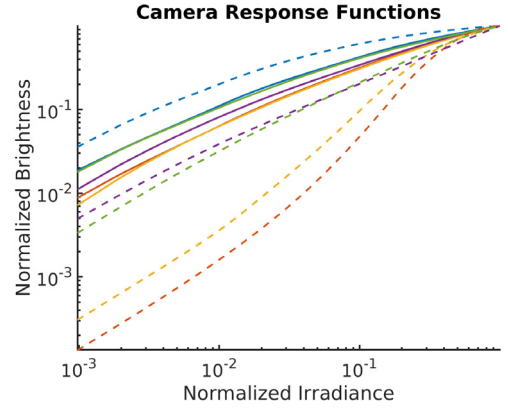


Fig. 1. Ten CRFs of real cameras, from [14]. Colors and line-styles have no meaning beyond identifying each function.

sensor, for each color channel, admits the model

$$\mathbf{y}_l = c^\gamma (\mathbf{z}_l + (\sigma_k \mathbf{k} \circ \mathbf{z}_l + \mathbf{n}_{p,l})^\gamma + \mathbf{n}_{q,l}), \quad l = 0, 1, \dots, L-1, \quad (1)$$

where the exponentiation is element-wise and the operator $\circ$ represents the Hadamard product. Focusing on the l th vector, $\mathbf{y}_l$ denotes the output image, $\mathbf{z}_l$ denotes the noiseless input to the sensor, $\mathbf{k}$ denotes the PRNU, $\mathbf{n}_{p,l}$ summarizes different noise sources, and $\mathbf{n}_{q,l}$ denotes the quantization noise. All these quantities are N dimensional, obtained after vectorizing images of size $N_1 \times N_2$. The scalars c and γ denote the color channel gain and the gamma correction factor, respectively. We assume that $\mathbf{k}$ follows a zero-mean white Gaussian process; for convenience, we include explicitly its variance, which is denoted by σ_k^2 .

The authors of [5] apply the Taylor expansion $(1+x)^\gamma = 1 + \gamma x + \mathcal{O}(x^2)$ about $x = 0$, reaching the multiplicative model:

$$\mathbf{y}_l = \mathbf{x}_l \circ (\mathbf{1} + \gamma (\sigma_k \mathbf{k})) + \mathbf{n}_{t,l}, \quad l = 0, 1, \dots, L-1, \quad (2)$$

where $\mathbf{x}_l = c (\mathbf{z}_l)^\gamma$. Applying the denoising filter, we obtain the residual

$$\mathbf{w}_l = \gamma \mathbf{x}_l \circ (\sigma_k \mathbf{k}) + \mathbf{n}_l, \quad l = 0, 1, \dots, L-1. \quad (3)$$

Although the denoising process is often non-ideal, we assume that the vectors $\mathbf{n}_l$ are mutually uncorrelated processes, which are also uncorrelated with $\mathbf{k}$.

2.2. Transfer curve

In deriving the multiplicative model, Chen et al. relied on the Taylor expansion of the gamma correction. As discussed by Grossberg and Nayar [10], the gamma correction is just a possible instance of Camera Response Function (CRF), which describes the mapping from image brightness to image irradiance. Often, this mapping is modeled as a monotonically increasing function between 0 and 1. Throughout their work, the authors of [10] argue that “camera’s response function may vary significantly from an analytic form like a gamma curve” [13]. In Fig. 1, we depict in logarithmic scale ten examples of CRFs obtained from [14]. In this plot, gamma functions would appear as straight lines, which is not the case; thus, we conclude that these CRFs differ from a gamma curve.

Moreover, in practice, we need to account for the digital preprocessing operations. We call this composition of functions, which includes the CRF, *transfer curve*; in the following, we denote it by h . In this work, we assume that it is a strictly monotonically increasing function. This assumption imposes some conditions on the preprocessing operations:

1. Pixelwise operations, like the ISO, are well approximated by strictly monotonic functions.

2. Operations that mix different pixels, such as demosaicing [15] or gray-scale conversion, are well approximated by a positive linear combination of these pixels.

In Section 3.7, we will test empirically whether the transfer curve is strictly monotonically increasing. We also assume that h is differentiable in its domain, denoting its derivative by $h'(x)$. For convenience, we also set $h(0) = 0$ and $h(1) = 1$.

2.3. Generalized model

Following the definition of transfer curve, we can write the output of the sensor as

$$\mathbf{y}_l = h(\mathbf{z}_l + (\sigma_k \mathbf{k}) \circ \mathbf{z}_l + \mathbf{n}_{c,l}) + \mathbf{n}_{q,l}, \quad l = 0, 1, \dots, L-1. \quad (4)$$

If $h(x) = (cx)^\gamma$, we recover the expression in (1). The function h acts element-wise, i.e., $[h(\mathbf{a})]_n = h([a]_n)$, for $n = 0, 1, \dots, N-1$. Our goal is to show that the transfer curve affects the denoising residuals. For simplicity, we ignore the term $\mathbf{n}_{c,l}$: this approximation is reasonable when either $\mathbf{n}_{q,l}$ or the errors introduced by the denoising process prevail, which is often the case in practice. Applying a first order Taylor expansion about $(\sigma_k \mathbf{k}) = \mathbf{0}$, we obtain, for $l = 0, 1, \dots, L-1$, the following expression:

$$\mathbf{y}_l = h(\mathbf{z}_l) + h'(\mathbf{z}_l) \circ \mathbf{z}_l \circ (\sigma_k \mathbf{k}) + \mathcal{O}((\sigma_k \mathbf{k})^2) + \mathbf{n}_{q,l}. \quad (5)$$

Summarizing the additive noise terms by $\mathbf{n}_l$ we obtain, for the l th image,

$$\mathbf{y}_l = h(\mathbf{z}_l) + h'(\mathbf{z}_l) \circ \mathbf{z}_l \circ (\sigma_k \mathbf{k}) + \mathbf{n}_l. \quad (6)$$

Let us define another element-wise function, $\tilde{g}(z) \doteq h'(z)z$. Also, set $\mathbf{x}_l \doteq h(\mathbf{z}_l)$. Since the random variables have zero mean, we can assume that the output of the denoising filter is $\mathbf{x}_l$. Thus, the residual obtained after denoising $\mathbf{y}_l$ accepts the model

$$\mathbf{w}_l = g(\mathbf{x}_l) \circ (\sigma_k \mathbf{k}) + \mathbf{n}_{d,l}, \quad l = 0, 1, \dots, L-1, \quad (7)$$

where, with a slight abuse of notation, we used the shorthand $g(\mathbf{x}_l)$ to denote $\tilde{g}(h^{-1}(\mathbf{x}_l)) = h^{-1}(\mathbf{x}_l) \circ h'(h^{-1}(\mathbf{x}_l))$ and $\mathbf{n}_{d,l}$ collects the noise terms. We call function g *PRNU emphasis*; in Section 3 we analyze how to estimate it. This model matches (3) if and only if h is a gamma curve, as we show in Appendix.

2.4. Emphasized estimator of the PRNU

We can introduce the emphasis on the estimator of the PRNU. As in [7], we work with the n th pixel of the residuals. We assume independence among pixels; although images themselves are locally correlated [16] and the estimate from the denoiser is usually a linear combination of the surrounding pixels [17]. Our goal is to reach a balance between rigor and tractability; in this sense, assuming independence allows us to work over the model while providing reasonable results. This assumption simplifies the result in vector form to stacking the individual estimates for each pixel.

Suppose the vector $\mathbf{n}_{d,l}$ in (7) is a spatially white Gaussian process with variance $\sigma_{d,l}^2$, that is, when focusing on a given residual, its pixels are modeled as independent and identically distributed Gaussian random variables. Aiming to be as general as possible, we assume that this variance may not be the same for every residual, as argued in [18]. Hence, when focusing on a particular pixel, the residuals are mutually independent, but not white. Then, given the L denoised images $\{\mathbf{x}_l\}_{l=0}^{L-1}$ and the PRNU $\mathbf{k}$, we can write the joint log-likelihood of the L denoising residuals $\{\mathbf{w}_l\}_{l=0}^{L-1}$ for the n th pixel as

$$l_n = -\frac{1}{2} \sum_{l=0}^{L-1} \log(2\pi\sigma_{d,l}^2) - \sum_{l=0}^{L-1} \frac{([\mathbf{w}_l]_n - \sigma_k [\mathbf{k}]_n g([\mathbf{x}_l]_n))^2}{2\sigma_{d,l}^2}. \quad (8)$$

Taking the derivative with respect to $[\mathbf{k}]_n$ and equating to zero to maximize the log-likelihood, we obtain

$$\sum_{l=0}^{L-1} \frac{g([\mathbf{x}_l]_n)}{\sigma_{d,l}^2} ([\mathbf{w}_l]_n - \sigma_k [\mathbf{k}]_n g([\mathbf{x}_l]_n)) = 0. \quad (9)$$

Solving the equation and writing the result in vector form, we reach the emphasized estimator of the PRNU:

$$\sigma_k \hat{\mathbf{k}} = \sum_{l=0}^{L-1} \frac{g(\mathbf{x}_l)}{\sigma_{d,l}^2} \circ \mathbf{w}_l / \sum_{l=0}^{L-1} \frac{g(\mathbf{x}_l) \circ g(\mathbf{x}_l)}{\sigma_{d,l}^2}, \quad (10)$$

where the division is entry-wise. This estimator includes the method proposed in [7] as a special case, because the expression

$$\sigma_k \hat{\mathbf{k}} = \sum_{l=0}^{L-1} \mathbf{x}_l \circ \mathbf{w}_l / \sum_{l=0}^{L-1} \mathbf{x}_l \circ \mathbf{x}_l. \quad (11)$$

corresponds to our estimator when $g(x) = x$ and the variance of each residual is assumed constant.

Our result also resembles [18, Eq. (10)], but it includes weights in the denominator. When the weights are constant, the estimator can be seen as an ordinary least squares (OLS) regressor, which is optimal when the variances $\sigma_{d,l}^2$ are the same for all l . To account for different variances, weighted least squares is a popular approach. In this sense, Aitken [19] proved that (1) the optimal weights are given by the inverse of the variance of the residuals, and (2) these weights must be applied in both the denominator and the numerator. Our proposal above aligns with this result: with uncorrelated Gaussian random variables, the weighted OLS estimator and the MLE match.

In practice, to estimate the values of $\{\sigma_{d,l}^2\}_{l=0}^{L-1}$, we follow a modified version of the method in [18]. There, the authors first aim to remove the component due to the PRNU in the denoising residuals. Then, they exploit the “pixel-dimension” to compute the variance of each residual. We propose to estimate first the PRNU using the OLS estimator, i.e., assuming constant variance. Let this estimate be $\tilde{\mathbf{k}}$; we use it to compute the errors,

$$\mathbf{e}_l = \mathbf{w}_l - g(\mathbf{x}_l) \circ \tilde{\mathbf{k}}, \quad l = 0, \dots, L-1, \quad (12)$$

(cf. [18, Eq. (5)], where they subtracted the average of the residuals). Finally, we compute the variance using the errors for different pixels, as in [18, Eq. (6)]:

$$\hat{\sigma}_{d,l}^2 = \frac{1}{N-1} \sum_{n=0}^{N-1} \left([\mathbf{e}_l]_n - \sum_{m=0}^{N-1} [\mathbf{e}_l]_m \right)^2, \quad l = 0, \dots, L-1. \quad (13)$$

Under our model, this modification should improve the results.

We can provide a geometrical interpretation of the strategy above. The operation $\mathbf{w}_l - g(\mathbf{x}_l) \circ \tilde{\mathbf{k}}$ projects the denoising residuals into the null space of the emphasized denoised images, giving access to a modified version of the error terms. We approximate the variance of each $\mathbf{n}_{d,l}$ in (7), for $l = 0, 1, \dots, L-1$, by the variance of these “modified errors”.¹ Finally, we exploit that close pixels share almost the same statistical properties. The estimate of the variance provided in [18] follows the same idea, but assuming that the covariates are constant. Our interpretation may allow the introduction of different strategies to estimate the variances, as MINQUE [20]. We plan to explore this option in a future work.

In our derivation, we assume that the noise is spatially white. But the denoising filter is far from ideal, and we expect it to perform better in less textured regions of the image. Although Lawgaly et al. [18] proposed a block-wise approach to account for this effect, both the problem of estimating the variance of the noise through a given residual and the problem of incorporating this quantity into the PRNU estimator are still open. Yet, as we will argue in Section 4.2, the model above is particularly beneficial in a detection problem with cropped images. We plan to consider the behavior of the noise terms in our future work.

¹ This procedure underestimates the variance, because it ignores the component of the error lying in the “emphasized-denoised-image subspace”.

2.5. Source camera attribution with cropped images

In this section, we modify the mainstream method to conduct SCA with cropped and scaled images [12] in order to include the emphasis function. Our procedure remains very similar, but it includes some modifications. For the sake of simplicity, in this work we focus on cropped images; the extension to scaled images is straightforward. Although the authors of [12] use the multiplicative model, note that here we include the PRNU emphasis to explain the method. We remark, in any case, that we do not follow any optimality criterion to derive this detector.

The algorithm proposed in [12] relies on a two-channel hypothesis testing problem. To account for the two-dimensional nature of images, in this subsection we conduct our analysis with matrices of size $N \times M$, i.e., we do not apply vectorization. To clarify that we are working with matrices, we follow the notation in Section 2.3, but using the corresponding capital letters. Let us restate (4) in matrix form:

$$\mathbf{Y} = h(\mathbf{Z} \circ (\mathbf{1} + (\sigma_{k,1} \mathbf{K}_1)) + \mathbf{N}_p) + \mathbf{N}_q, \quad (14)$$

where we dropped the subindex l because we only consider the test image. In the equation above, $\mathbf{Y}$ is the original image, $\mathbf{Z}$ is the input to the sensor, h is the transfer curve, and $(\sigma_{k,1} \mathbf{K}_1)$ is the PRNU. The noise sources are $\mathbf{N}_c$ and $\mathbf{N}_q$. Then, the denoising residual obeys the following relationship:

$$\mathbf{W} = g(\mathbf{X}) \circ (\sigma_{k,1} \mathbf{K}_1) + \mathbf{N}_d, \quad (15)$$

where g is the PRNU emphasis, $\mathbf{N}_d$ is the additive noise term, and $\mathbf{X} = h(\mathbf{Z})$. Now, let us define the operator T_s , which crops its argument at the position given by $s \doteq (s_1, s_2)$. We have access to the residual of the cropped test image, $T_s(\mathbf{W}) \in \mathbb{R}^{P \times Q}$. Assume, also, that we know an estimated fingerprint $(\sigma_k \hat{\mathbf{K}})$ for a source device with PRNU $(\sigma_k \mathbf{K})$. Then, the authors of [12] aim to test whether the two PRNUs are the same, that is,

$$\begin{aligned} H_0 : (\sigma_k \mathbf{K}) &\neq (\sigma_{k,1} \mathbf{K}_1), \\ H_1 : (\sigma_k \mathbf{K}) &= (\sigma_{k,1} \mathbf{K}_1). \end{aligned} \quad (16)$$

After some mathematical derivations, they arrive at the following procedure. First, they propose a brute-force search to estimate the cropping point using a correlation-based method. In particular, they rely on the NCC [21]; for $0 \leq s_1 < N$ and $0 \leq s_2 < M$, we have:

$$[\text{NCC}]_{s_1, s_2} = \sum_n \sum_m \frac{((\sigma_k \hat{\mathbf{K}})_{n,m} - \mu_k)([\mathbf{W}_g]_{(n+s_1)_N, (m+s_2)_M} - \mu_w)}{\|(\sigma_k \hat{\mathbf{K}}) - \mu_k\|_F \|\mathbf{W}_g - \mu_w\|_F}, \quad (17)$$

where $\mathbf{W}_g \in \mathbb{R}^{N \times M}$ is the result of padding with zeros the cropped test residual after multiplying it by the emphasized denoised image, $T_s(g(\mathbf{X}) \circ \mathbf{W}) \in \mathbb{R}^{P \times Q}$, and the expressions $(n+s_1)_N$ and $(m+s_2)_M$ denote the shifts modulo N and M , respectively. The scalars μ_k and μ_w are the averages of all values in $(\sigma_k \hat{\mathbf{K}})$ and $T_s(g(\mathbf{X}) \circ \mathbf{W})$, respectively. We denote the Frobenius norm of the matrix $\mathbf{A}$ as $\|\mathbf{A}\|_F \doteq (\sum_n \sum_m (|\mathbf{A}|_{n,m})^2)^{1/2}$.

Due to the statistical whiteness of the PRNU, the correlation with its shifted versions has zero mean, save for the true alignment, where we expect to find a large positive value. Thus, the position of the maximum of the NCC as a function of the shifts is a good estimator of the position of the cropping point:

$$s_{\text{peak}} \doteq \arg \max_{(s_1, s_2) \in B} [\text{NCC}]_{s_1, s_2}, \quad (18)$$

where $B \doteq \{0, 1, \dots, N-P-1\} \times \{0, 1, \dots, M-Q-1\}$. Note that we only search for the maximum in a reduced range, which corresponds to the difference between the sizes of the original and the cropped images (any value of (s_1, s_2) outside this range would be unfeasible).

After solving the estimation problem above, the authors address the detection stage by focusing on the position of s_{peak} . The statistic they consider is the PCE [12], which can be computed using the NCC:

$$\text{PCE} = \frac{([\text{NCC}]_{s_{\text{peak}}})^2}{\frac{1}{NM-|A|} \sum_{s \notin A} ([\text{NCC}]_s)^2}, \quad (19)$$

where A is a small region surrounding the position given by s_{peak} and $|A|$ is the number of elements within this region.

We consider the same pipeline with two modifications. First, we use the signed PCE instead of the PCE [22]:

$$\text{SPCE} = \text{sgn}([\text{NCC}]_{s_{\text{peak}}}) \frac{([\text{NCC}]_{s_{\text{peak}}})^2}{\frac{1}{NM-|A|} \sum_{s \notin A} ([\text{NCC}]_s)^2}, \quad (20)$$

where $\text{sgn}(x)$ equals -1 when x is negative and 1 when x is non-negative. Second, we use $T_s(\mathbf{W})$, i.e., the cropped test residual, instead of $T_s(g(\mathbf{X}) \circ \mathbf{W})$, because the results remain very similar and the procedure is simpler; we plan to introduce the multiplication step jointly with the variance of the noise in a future work. We declare H_1 whenever the signed PCE is above a preset threshold τ . During our analysis in Section 4.2, we will also use the CCN [22]:

$$\text{CCN} = \text{sgn}([\text{NCC}]_{s_{\text{peak}}}) \sqrt{\text{PCE}}, \quad (21)$$

which is an equivalent statistic to the signed PCE in terms of detection results. In Fig. 2, we provide a block diagram with the different steps of the process, including the different estimators of the PRNU emphasis that we will present in Section 3 and the post-processing operations (zero-meaning and Wiener filtering in the DFT domain [1]) that we will mention in both Sections 3.5 and 4.1.

3. Estimating the PRNU emphasis

After introducing the PRNU emphasis, in this section we present some methods to estimate it. Our strategy relies on the covariance of the residuals. First, we show that the covariance is linked to the PRNU emphasis through a spectral decomposition. Then, we provide a simplified method, which is based on an estimate of the PRNU. We evaluate both techniques to obtain a preliminary idea of the shape of the PRNU emphasis.

Based on the results, we propose a regularized least squares estimation method, which is useful when the amount of information is scarce. Finally, we show a method to approximate the transfer curve from the PRNU emphasis.

3.1. Linking the covariance with the PRNU emphasis

Following the model in (7), given two denoising residuals obtained from images taken with the same camera, which we will denote by $\mathbf{w}_i$ and $\mathbf{w}_j$, we can write the covariance of the n th entry as:

$$\mathbb{E}([\mathbf{w}_i]_n [\mathbf{w}_j]_n) = \sigma_k^2 [g(\mathbf{x}_i)]_n [g(\mathbf{x}_j)]_n, \quad n = 0, \dots, N-1, \quad (22)$$

where $\mathbb{E}$ takes the expectation of its argument. Given the vectors $\mathbf{x}_i$ and $\mathbf{x}_j$, we can define the following bivariate function:

$$\phi : (z_1, z_2) \mapsto \sigma_k^2 g(z_1) g(z_2). \quad (23)$$

The first step to recover the PRNU emphasis is to estimate ϕ , which is the function binding the pair of pixels of the denoised images $([\mathbf{x}_i]_n, [\mathbf{x}_j]_n)$ with the covariance between the entries of the denoising residuals $[\mathbf{w}_i]_n$ and $[\mathbf{w}_j]_n$.

3.2. Regressogram

From the definition of transfer curve, we can assume that the function to estimate is continuous. We note that our aim in this section is exploratory [23]: we want to obtain a curve that describes the PRNU emphasis minimizing the number of assumptions. Then, one option to retrieve ϕ is the regressogram [24]; although more sophisticated methods may exist, this approach is simple and fast.

First, under our model, the product of the residuals is an unbiased estimator of the covariance. Thus, we can access an estimate of the right hand side of (23).

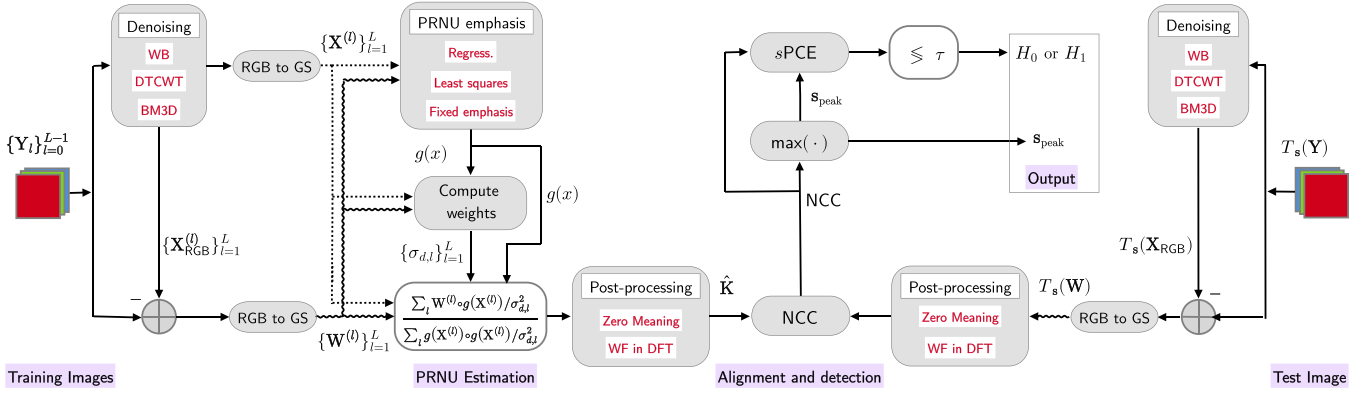


Fig. 2. Block diagram showing the different steps in the source camera attribution problem with cropped images. “RGB to GS” stands for “RGB to grayscale conversion”; “WF in DFT” is a shorthand for “Wiener filtering in the DFT domain”. We use $\{Y_l\}_{l=0}^{L-1}$ to denote the training images in matrix form.

On the other hand, we have two covariates, which have the same span. Thus, we can divide the 2D plane in N_b^2 squares. Let $I_{p,q}$ be the square corresponding to the p th division in one axis and the q th division in the other axis, for $p, q = 0, 1, \dots, N_b - 1$. Then, the regressogram is

$$\hat{\phi}(z_1, z_2) = \frac{\sum_{i,j \neq i} \sum_n [\mathbf{w}_i \circ \mathbf{w}_j]_n \mathbb{1}_{p,q}(([\mathbf{x}_i]_n, [\mathbf{x}_j]_n))}{\sum_{i,j \neq i} \sum_n \mathbb{1}_{p,q}(([\mathbf{x}_i]_n, [\mathbf{x}_j]_n))}, \quad (24)$$

if $(z_1, z_2) \in I_{p,q}$, where $\mathbb{1}_{p,q}$ is the indicator function of $I_{p,q}$.

3.3. From the regressogram to the PRNU emphasis

From the 2D function with the covariances, we can obtain the PRNU emphasis by solving a matrix-nearness problem. Let us define the matrix $\hat{\Phi} \in \mathbb{R}^{N_b \times N_b}$, which contains the values of the regressogram for each possible p and q . To reach $g(x)$, we use two properties:

1. Ideally, the matrix $\hat{\Phi}$ is symmetric, because $\phi(z_1, z_2) = \phi(z_2, z_1)$.
2. In the noiseless case, the matrix $\hat{\Phi}$ is rank one, because it corresponds to the outer product of a function by itself sampled at the same points in both dimensions.

Then, we need to approximate our matrix by a symmetric rank-one matrix [25]. We opt for a two-step approach.² First, we symmetrize our matrix:

$$\hat{\Phi}_s = (\hat{\Phi} + \hat{\Phi}^T)/2. \quad (25)$$

Now, to obtain the rank-one approximation, we apply the following theorem.

Theorem 1 (Eckhart-Young, [26]). Let $\mathbf{A}$ be a rank r matrix and $\mathbf{USV}^T$ be its singular value decomposition, with singular values s_i sorted in descending order. Then, if $\mathbf{D}$ is a rank $k < r$ matrix,

$$\min_{\mathbf{D}} \|\mathbf{A} - \mathbf{D}\|_2 = \left\| \mathbf{A} - \sum_{i=0}^{k-1} s_i \mathbf{u}_i \mathbf{v}_i^T \right\|_2 = s_{k+1}, \quad (26)$$

where $\mathbf{u}_i$ and $\mathbf{v}_i$ are the i th columns of $\mathbf{U}$ and $\mathbf{V}$, respectively.

Let the closest rank one matrix to $\hat{\Phi}_s$ be $\hat{\Phi}_1$. To obtain the estimate of $g(x)$ from the rank one approximation, we first observe that, due to symmetry, $\hat{\Phi}_1 = s_1 \mathbf{u}_1 \mathbf{u}_1^T$. From the definition of ϕ , we can conclude that, in the ideal case, each of the N_b entries of the vector $\mathbf{u}_1$ is proportional to a sample of $g(x)$.

² In this paper, we leave the sign of the eigenvalues unconstrained; we plan to impose positive-definiteness in our future work.

3.4. Simplified regressogram

In the method above, we are computing more points than necessary. In particular, we estimate N_b^2 points to obtain N_b samples of the emphasis. To simplify it, we can rely on an estimator of the PRNU. Following our model of the PRNU in (7), but computing the estimator presented in (11), we obtain

$$\hat{\mathbf{k}} = (\sigma_k \mathbf{k}) \circ \frac{\sum_l \mathbf{x}_l \circ g(\mathbf{x}_l)}{\sum_l \mathbf{x}_l \circ \mathbf{x}_l} + \frac{\sum_l \mathbf{x}_l \circ \mathbf{n}_{d,l}}{\sum_l \mathbf{x}_l \circ \mathbf{x}_l}. \quad (27)$$

Then, if L is sufficiently large, we can assume that the term $\beta \doteq \sum_l \mathbf{x}_l \circ g(\mathbf{x}_l) / \sum_l \mathbf{x}_l \circ \mathbf{x}_l$ is close to a constant multiple of the all-ones vector. Then, we can estimate $\sigma_k g(x)$ via the product between the residuals and the estimated PRNU, applying a simplified version of the regressogram with only one covariate. This strategy reduces to N_b the number of points to estimate.

In practice, given a set of L residuals in the training stage, we follow a leave-one-out strategy: we use $L - 1$ residuals to evaluate (11), applying zero-meaning and Wiener filtering in the DFT domain to this estimate [1]. Then, we feed the product between this quantity and the remaining residual to the 1D regressogram. This method allows us to gather L products to estimate the function $g(x)$.

We can refine this estimator using an iterative strategy. Let g_1 be the output of the regressogram in the method above. The subindex indicates the iteration number. Then, for $m = 1, 2, \dots$, we can evaluate

$$\hat{\mathbf{k}}_{m+1} \doteq \sum_l g_m(\mathbf{x}_l) \circ \mathbf{w}_l / \sum_l g_m(\mathbf{x}_l) \circ g_m(\mathbf{x}_l), \quad (28)$$

repeating our leave-one-out strategy to estimate g_{m+1} . We expect the term $\sum_l g_m(\mathbf{x}_l) \circ g(\mathbf{x}_l) / \sum_l g_m(\mathbf{x}_l) \circ g_m(\mathbf{x}_l)$ to be progressively closer to a constant multiple of the all-ones vector when m grows.

3.5. Assessing the shape of the emphasis

In the following, we estimate the PRNU emphasis from real images using different versions of the regressogram. As we highlighted at the beginning of this section, our aim is exploratory; in particular, we have three goals:

1. confirm that the PRNU emphasis is non-linear,
2. corroborate that the regressogram and the simplified regressogram provide similar results,
3. analyze the typical shape of the emphasis, aiming to propose parametric estimators.

We consider a database of 16 cameras, mixed from public [27,28] and private sources, whose details appear in Table 1. For each camera, we gather 40 images, cropping them to squares of size 2000×2000 to make the experiments independent of the native resolution of the

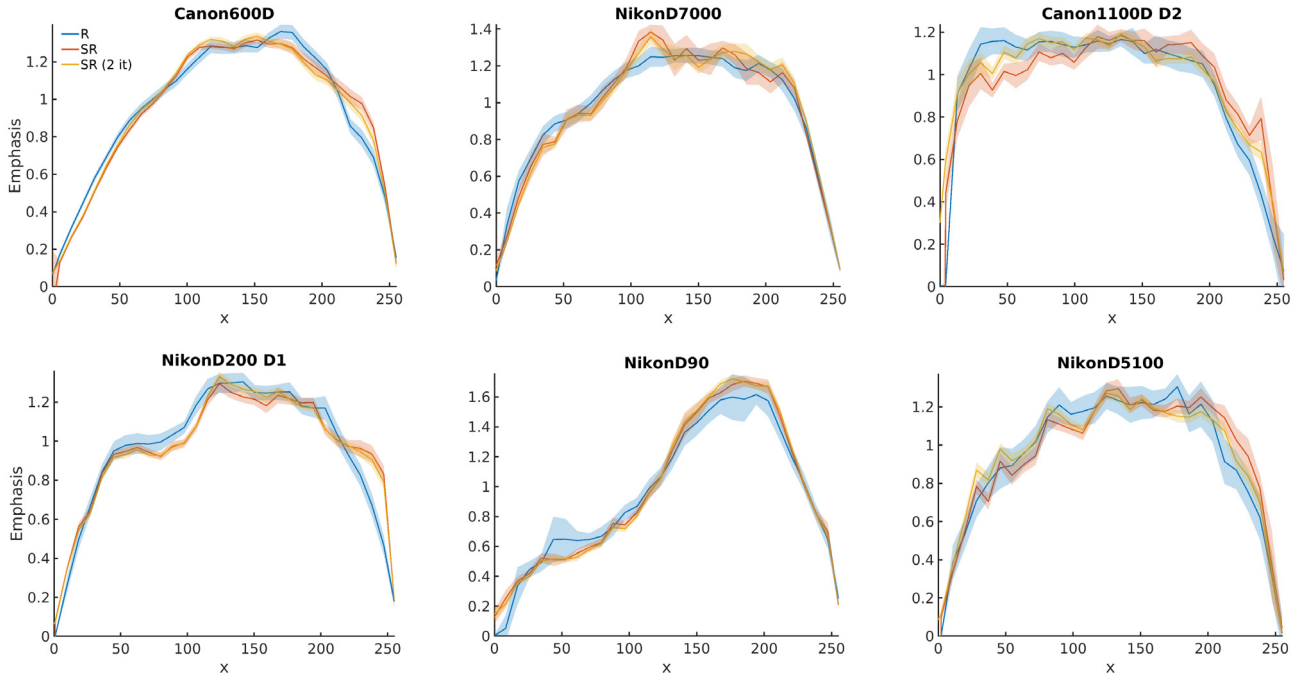


Fig. 3. Estimated PRNU emphasis. Blue: regressogram (R); red: simplified regressogram (SR); yellow: simplified regressogram, two iterations (S. R. 2 it). The semi-transparent bands are the 95% Gaussian confidence intervals. Note that, in some cases, the red curve overlaps the yellow line, rendering the latter barely visible.

Table 1

Cameras used in our experiments. Both Dresden [27] and RAISE [28] are public databases.

Camera model	Sensor	Devices	Num. images	Database
Canon600D	22.3 × 14.9 mm CMOS	1	241	Own
Canon1100D	23.2 × 14.7 mm CMOS	3	316/122/216	Own
NikonD60	23.6 × 15.8 mm CCD	1	197	Own
NikonD70	23.7 × 15.6 mm CCD	2	43/43	Dresden
NikonD70S	23.7 × 15.6 mm CCD	2	43/47	Dresden
NikonD90	23.6 × 15.8 mm CMOS	1	250	RAISE
NikonD200	23.6 × 15.8 mm CCD	2	48/43	Dresden
NikonD3000	23.6 × 15.8 mm CCD	1	230	Own
NikonD3200	23.2 × 15.4 mm CMOS	1	250	Own
NikonD5100	23.6 × 15.6 mm CMOS	1	250	Own
NikonD7000	23.6 × 15.6 mm CMOS	1	250	RAISE

camera. The cropping point is random. All our images are uncompressed (TIFF), and we use the wavelet-based (WB) denoising algorithm proposed in [29].

We denoise each color channel independently, converting to grayscale the result. Although special versions to deal jointly with the three color channels exist, as the CBM3D [30], we note that non-positive or non-linear combinations of the pixels of the color channels violate one of our assumptions about the transfer curve in Section 2.2. Thus, these methods may provide a different estimated emphasis. We also apply zero-meaning and Wiener filtering in the DFT domain to the fingerprint and the remaining residuals, as in [1].

We estimate the PRNU emphasis using the regressogram and the simplified regressogram with $m = 1$ and 2 iterations, all of them with 32 bins per dimension. To do so, we sample 15 residuals randomly from our set of 40 residuals. To compute confidence bands, we repeat the sampling process 50 times.

The results show that the PRNU emphasis significantly differs from the linear function assumed by the linear model in (3) (Fig. 3). Also, the difference between the regressogram and the simplified regressogram after two iterations is negligible in most cases.

In practice, images are often compressed. Moreover, the wavelet-based denoising filter is fast, but several studies [31,32] reported that other denoising algorithms are better at handling the PRNU. Thus, we also estimate the PRNU emphasis under three different scenarios,

namely (1) JPEG compression with quality factor 90, (2) BM3D denoising [31,33], because different studies report its good performance [8, 32], and 3) using the dual-three complex wavelet transform (DTCWT) denoising filter [34]: this method accounts for directionality and shift-invariance [35] and imposes symmetric boundary conditions. We show the results in Fig. 4; note that we used three cameras, each of them different from the cameras in Fig. 3. We can see that these operations preserve the PRNU emphasis. However, the estimates when considering JPEG compression are noisier than in the uncompressed scenario, as one would expect due to the extra quantization noise.

We conclude that the regressogram is suitable when the size of the training images is large. Yet, the performance degrades when the number of pixels decreases. Next, we exploit the structure of the PRNU emphasis to circumvent this limitation.

Remark. To estimate the PRNU emphasis, we need enough samples for each possible bin on the covariate axis. In other words, the denoised images must span the whole range of pixel intensities. Although natural images often fulfill this requirement, flat-field images do not; this is troublesome because the lack of content is known to benefit the estimation of the PRNU [2]. In practice, however, we may avoid the problem by ensuring that the dataset contains a sufficient number of natural images.

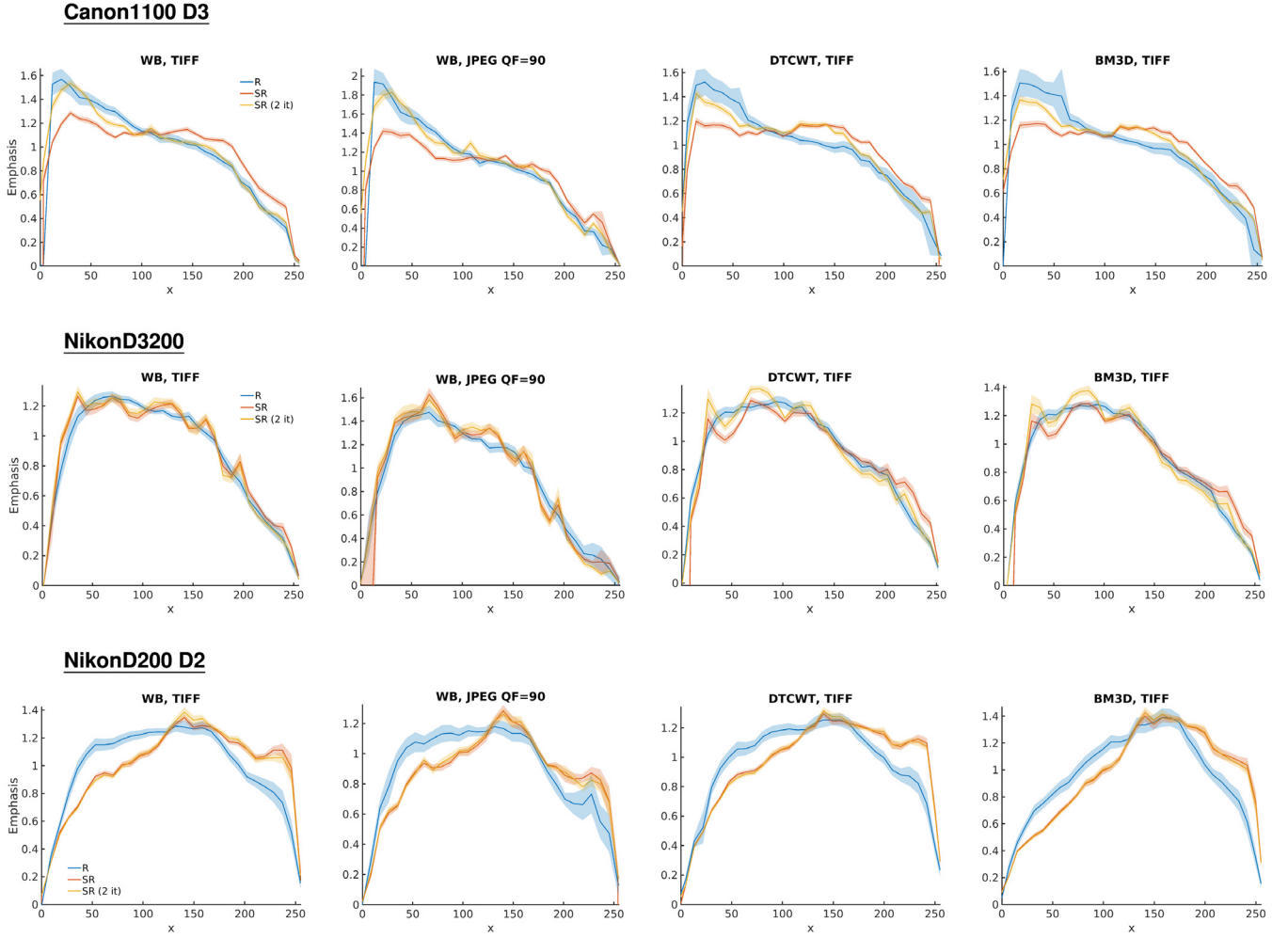


Fig. 4. Estimated PRNU emphasis, considering JPEG compression with quality factor 90 (center) and either BM3D or DTCWT denoising (right). Blue: regressogram (R); red: simplified regressogram (SR); yellow: simplified regressogram, two iterations (SR 2 it). The semi-transparent bands are the 95% Gaussian confidence intervals.

Remark. JPEG compression shapes the power spectrum of the noise, providing it with a certain structure and introducing artifacts. This phenomenon contradicts our assumption in (1) about the quantization noise being negligible. As we will show in Section 4, the phase-only method by Kang et al. [22] performs comparatively better under moderate JPEG compression. We conjecture that one of the reasons is because it introduces a whitening step that mitigates compression artifacts. Also, the larger the compression, the less white the noise, increasing the gap between the method by Kang et al. and the emphasized estimators.

3.6. Least squares estimation

In some applications we have to deal with small images. Hence, the estimate of the emphasis may show a large variance. In the following, we present a method to reduce these fluctuations by introducing some bias in the estimates.

First, from our exploratory results, we observe that fitting the emphasis to a dictionary of skewed concave curves is reasonable. Thus, we can exploit the covariance between the denoising residual and the estimated PRNU to use a least squares method.

To construct the column vector with dependent variables, denoted by $\mathbf{r}$, we stack the estimates of the covariance $\mathbf{w}_i \circ \mathbf{k}$; as we did with the simplified regressogram in Section 3.4, we use the products between the estimate of the PRNU using the leave-one-out approach and the

remaining residual. We repeat the same process with the corresponding covariates, obtaining $\mathbf{x}$. Then, we aim to find the matrix of coefficients $\mathbf{b}$ to minimize

$$\arg \min_{\mathbf{b}} \|\mathbf{r} - \Psi(\mathbf{x})\mathbf{b}\|_2 \Rightarrow \mathbf{b}_{opt} = (\Psi(\mathbf{x})^T \Psi(\mathbf{x}))^{-1} \Psi(\mathbf{x})^T \mathbf{r}, \quad (29)$$

where $\Psi(\mathbf{x})$ is a matrix containing skewed concave functions of the entries of $\mathbf{x}$. In this work, we use a Gammatone parametric dictionary [36] without frequency modulation:

$$\psi_{\beta}(x) = a x \exp(-2\pi\beta x), \quad \Psi(\mathbf{x}) = [\psi_{\beta_1}(\mathbf{x}), \dots, \psi_{\beta_m}(\mathbf{x})], \quad (30)$$

where ψ_{β} acts element-wise and a is a normalization parameter. We can choose β to control the skewness (the larger β , the more skewed the atom). To add an additional degree of flexibility, we include the inverse parabola, $-x^2 + 255x$. We depict an example of this basis in Fig. 5. In this work, we use only $\beta = 0.003$, its reversed version (along the x -axis), and the inverse parabola. These three functions provide sufficient flexibility while limiting the number of parameters.

From the figure, we can conclude that the atoms of this dictionary are correlated. As a result, (1) the singular values of $\Psi(\mathbf{x})$ decay gradually to zero, and (2) the ratio between the largest and the smallest singular values is large. These two properties induce instabilities in the estimator [37]. To mitigate them, we use Tikhonov regularization using a multiple of the identity as Tikhonov matrix (in other words, we use ridge regression). That is, we compute

$$\mathbf{b}_{opt} = (\Psi^T(\mathbf{x})\Psi(\mathbf{x}) + \lambda \mathbf{I})^{-1} \Psi^T(\mathbf{x}) \mathbf{r}, \quad (31)$$

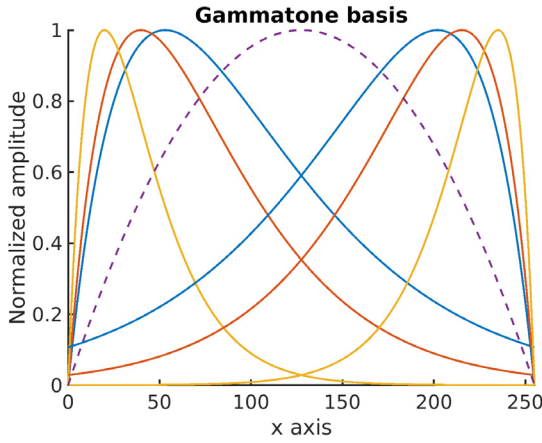


Fig. 5. Gammatone basis, obtained evaluating the expression in (30) using $\beta = 0.003, 0.004, 0.008$, and their reversed versions. The larger β , the more skewed the atom. We also included the function $-x^2 + 255x$ after normalizing the amplitude (dashed).

where $\mathbf{I}$ is the identity matrix of the proper size. To select λ , we use Generalized Cross-Validation (GCV) [38]; we consider 200 possible regularization parameters for the GCV, chosen accordingly to [39].

We can follow the iterative procedure in Section 3.4 to enhance the estimates. In this case, we consider the inverse parabola, $-x^2 + 255x$, in the first iteration (instead of the linear function), which is closer to the typical value of the emphasis. Also, to estimate the PRNU, we can apply the weighting procedure described in Section 2.4. In this section, we divide each image into 16 blocks to compute and apply the weights [18].³

3.6.1. Experiments with the LS estimator

We depict in Fig. 6 some examples of the estimated PRNU emphasis using the least squares method as described above on images of size 512×512 pixels. To obtain them, we gather a set of 40 images from the database in Table 1. Then, we sample uniformly at random 15 images from the set, repeating this process 50 times to compute the confidence bands. To ease the comparison, we include the results of the simplified regressogram (two iterations) using 2000×2000 pixels. We use the Gammatone basis with $\beta = 0.003$, its reversed version, and the inverse parabola (see Fig. 5). Although the number of data points was severely reduced, the estimate preserves the overall shape of the original PRNU emphasis while stabilizing the variance.

3.7. Estimating the transfer curve

The last step is to show that a valid transfer curve induces the PRNU emphasis effect. Although obtaining this function is not necessary to apply our generalized model in practice, we provide a method to approximate it in this section, aiming to confirm our previous assumptions about the shape of the curve.

Note, however, that we make no attempt to provide a robust algorithm to retrieve the transfer curve; rather, we aim to be illustrative. In practice, computational instabilities may hinder the application of this technique.

We start from the analytical definition of PRNU emphasis:

$$\tilde{g}(z) = zh'(z). \quad (32)$$

³ We can apply these ideas to the regressograms too, but in the large sample size regime the improvement is unnoticeable, so we skip them to reduce the computational burden.

The output of the denoising filter is $x = h(z)$; thus, to write the output of the regressogram in terms of quantities we know, the proper expression is $\tilde{g}(h^{-1}(x))$. Then, we have

$$\tilde{g}(h^{-1}(x)) = h^{-1}(x) \frac{dh(y)}{dy} \Big|_{y=h^{-1}(x)}. \quad (33)$$

First, note that by the chain rule, $(f^{-1})'(x) = 1/f'(f^{-1}(x))$. Therefore, we can rewrite the expression above as

$$\tilde{g}(h^{-1}(x)) = \frac{h^{-1}(x)}{(h^{-1})'(x)} \implies \frac{d \log h^{-1}(x)}{dx} = \frac{1}{\tilde{g}(h^{-1}(x))}. \quad (34)$$

Integrating both sides, we obtain

$$h^{-1}(x) = \exp \left(\int_0^x \frac{1}{\tilde{g}(h^{-1}(y))} dy + C \right). \quad (35)$$

We can evaluate the integral numerically using quadrature methods [41]. In practice, we add a small quantity to $\tilde{g}(h^{-1}(x))$ to avoid the division by zero.

To account for the boundary conditions, we rely on our definition of transfer curve. First, since we assumed that the transfer curve was strictly monotonic, the inverse function h^{-1} exists. Moreover, we assumed $h(0) = 0$ and $h(1) = 1$; from the definition of inverse, we observe that $h^{-1}(0) = 0$ and $h^{-1}(1) = 1$. However, enforcing the first condition leads to an initial value problem without solution because $\log(0)$ is undefined. To overcome this limitation, we restrict our attention to the interval $[\epsilon, 1]$; we also set $h^{-1}(\epsilon) = \epsilon$, with ϵ equal to a small value. Therefore, solving the equation, we obtain $C = \log \epsilon$.

To impose the condition $h^{-1}(1) = 1$, we normalize directly the result of the differential equation. Thus, the final expression is

$$h^{-1}(x) \approx \frac{\exp \left(\int_{\epsilon}^x \frac{1}{\tilde{g}(h^{-1}(y))} dy + \log \epsilon \right)}{\exp \left(\int_{\epsilon}^1 \frac{1}{\tilde{g}(h^{-1}(y))} dy + \log \epsilon \right)}. \quad (36)$$

Plotting x against $h^{-1}(x)$ provides the graph of $h(x)$.

3.7.1. Experimental evaluation

In this section, we aim to verify the method to recover the transfer curve. In all cases, we compute the integral of $1/(\tilde{g}(h^{-1}(x)) + 10^{-3})$ to avoid instabilities.

To validate the feasibility and assess the accuracy of the method, we compute the PRNU emphasis for the transfer curves shown in Fig. 1. Then, we apply the recovery algorithm in Section 3.7. As we depict in Fig. 7 (left), the results are reasonable. We believe that the source of the discrepancy is the small quantity we are adding in the denominator to avoid computational instabilities. Note that the axes are in natural units as opposed to the curves shown in Fig. 1.

Now, we compute the transfer curves for some of our PRNU emphases in Figs. 3–4 following the same procedure (Fig. 7, right). To show these results, we are setting manually the first value of the curves to zero because the value of ϵ was relatively large. We also interpolate the PRNU emphasis to five times the sampling density. From the results, we can conclude that they are strictly monotonically increasing functions, thus confirming the assumption in Section 2.2.

4. Empirical results

In this section, we show the performance of the emphasis in a SCA scenario with cropped images, following the method presented in Section 2.5. In particular, given a test image, we repeat our detection process with each camera from the database. For practical purposes, the test image is always taken with one of these cameras; that is, there is always a correct match.

First, we overview the experimental setup. We build upon this section to provide a detailed explanation about the effects of model misspecification. Finally, we conduct some experiments and discuss the results. We also include an evaluation of the computational complexity of our method.

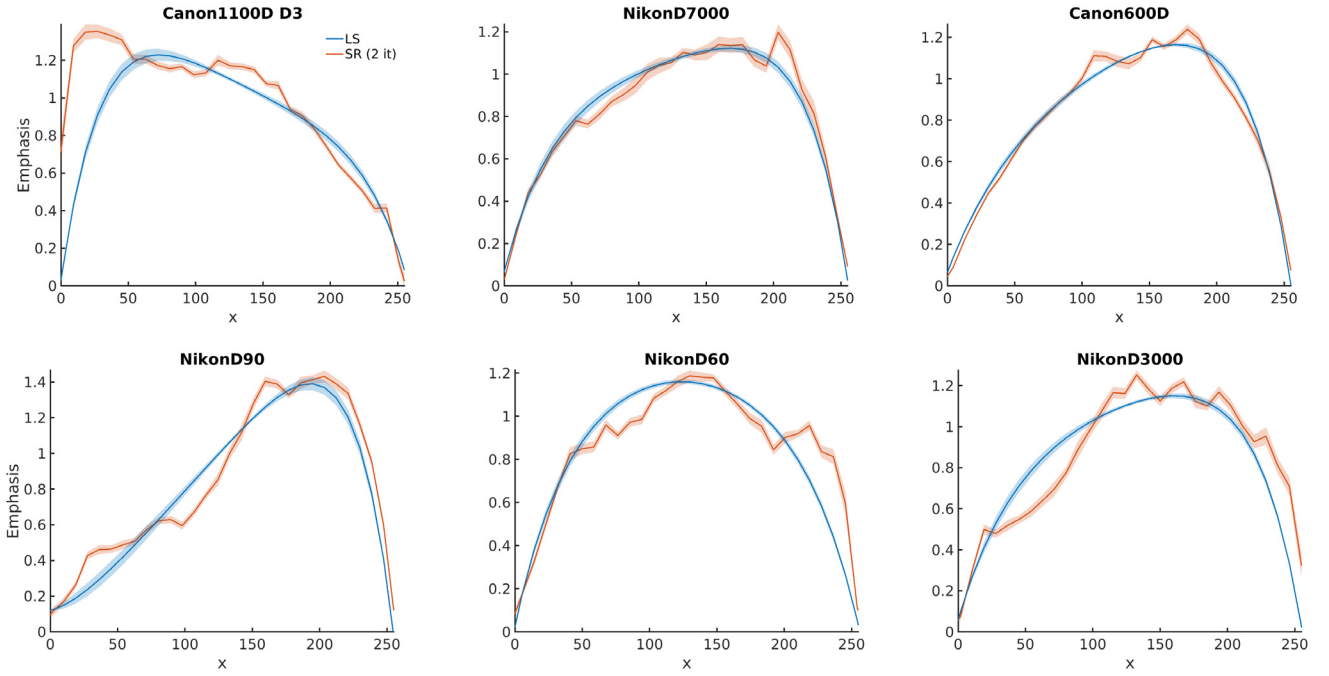


Fig. 6. Estimated PRNU emphasis. Red: simplified regressogram (SR, 2 it), using images of size 2000×2000 pixels; blue: least squares estimate (LS), using images of size 512×512 pixels. The semi-transparent bands are the 95% Gaussian confidence intervals.

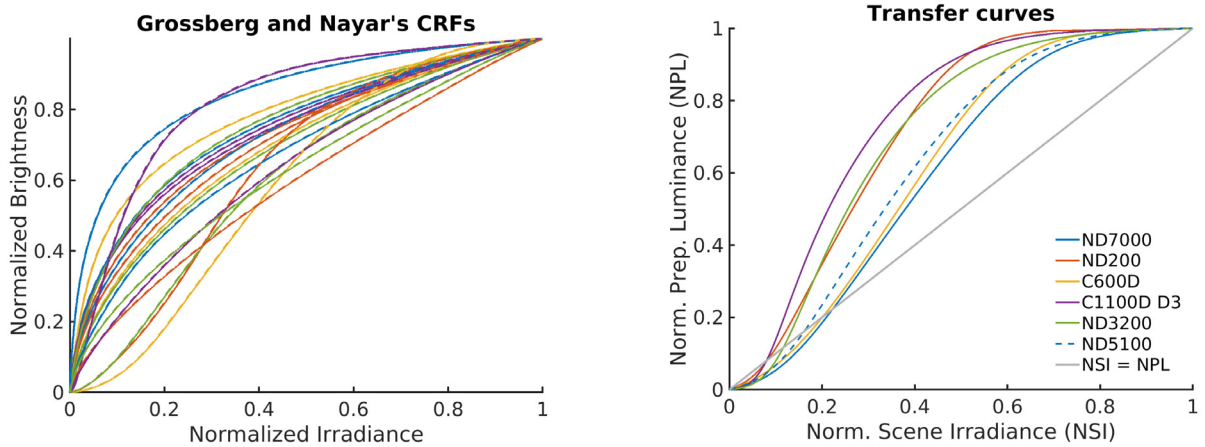


Fig. 7. Results of the algorithm presented in Section 3.7. (Left) CRFs from Grossberg and Nayar [40] (solid lines), and their recovered version (dashed lines). (Right) Estimated transfer curves from some of our PRNU emphases. “Prep.” stands for “preprocessed” and “Norm.” for “normalized”.

4.1. Experimental setup

Regarding the experimental setup, we use the same database as in Section 3.5, but excluding four cameras, the NikonD70/S, because we observed spurious correlations between them in the cropped case. We believe this effect is related to the discrimination problems discussed in [27].

As we did in Section 3.5, for each camera, we gather 40 images, cropping them to squares of size 2000×2000 to make the experiments independent of the native resolution of the camera. This cropping point is random, but constant for all the images from the same camera. We denoise each channel independently, converting to grayscale the result. We use 15 images of each camera in the training split to estimate both the emphasis and the PRNU.

To estimate the PRNU, that is, to obtain $\hat{\mathbf{K}}$, we consider:

1. The estimator in (10), but assuming that the multiplicative factor is x , that is, the weighted version of the MLE in [7], which works as a baseline. We label this method as “BL”. Although the

method presented in [5] uses the original image instead of the denoised image, the performance is almost identical.

2. The emphasized estimator presented in (10) using the estimated emphasis, which is labeled as “EE”. Most simulations rely on the simplified regressogram (with two iterations, i.e., $m = 2$) and $N_b = 32$ bins, but in Section 4.4 we use the least squares method.
3. Our emphasized estimator in (10), but using the inverse parabola, $g_f(x) = -x^2 + 255x$, as a “fixed emphasis function”. This method is less time-consuming and, given the similarity of the emphasis with an inverted parabola, it is expected to outperform the baseline. We label this method as “FE”.
4. The weighted average of the residuals (WA), as proposed in [18]. As we will show in Section 4.2, it should perform better than the classical MLE (11), because no scaling factor distorts the fingerprint spatially. Yet, the emphasized methods should provide better results.

In all cases, we apply the weighting procedure described in Section 2.4 using 16 blocks for each image [18]. We observed that some

blocks, mainly those belonging to saturated regions, have small variance, but they carry almost no information about the PRNU. Thus, we trim to 0 any normalized weight greater than 0.5; we also re-normalize again the result after this operation. When considering estimators whose weights are constant, for instance the method in (11), we will label them as unweighted (UW).

We apply zero-meaning and Wiener filtering in the DFT domain to the estimated PRNU from each method, as in [1]. We use the remaining 25 images in the testing split, cropping them to squares of size 512×512 pixels (in this case, the cropping point varies among different images). We randomly choose 50 cropping points for each test image. Following the mainstream algorithm in Section 2.5, we need to estimate these cropping points before addressing the detection process.

As a result, we have 1250 samples under H_1 and 13750 samples under H_0 for each camera. For the 12 cameras in our database, we have 15000 samples under H_1 and 165000 under H_0 .

As we argued in Section 2.5, our test statistic is the signed PCE [12, 42], because it remains the most popular similarity metric [8]. We set the length of the side of the small square region surrounding the position of the peak, A , to 11; thus, $|A|$ in (20) is 11×11 . We label a sample as “positive” (that is, as H_1) whenever its signed PCE is above a given threshold.

As a performance measure, we provide modified receiver operating characteristic (ROC) curves. These curves depict the True Positive Rate (TPR) averaged over the 12 cameras for a fixed False Positive Rate (FPR); yet, they are upper bounded by the probability of finding the correct cropping point [43,44]. To obtain these curves, we label as H_0 the samples in which the position of the peak of the NCC differs from the cropping point. To show the performance for low FPRs, we set the x -axis in logarithmic scale.

4.2. Model misspecification

Before detailing the empirical results, we briefly discuss the effects of model misspecification in the experimental setup above. For simplicity, we focus on uncompressed images and wavelet-based denoising. In a nutshell, we argue that using the baseline (or its unweighted equivalent in [5]) makes the spatial distribution of the variance of the estimated PRNU more heterogeneous. This effect modifies the variance of the cross-correlation under each shift, which hinders both the estimation of the cropping point and the computation of the signed PCE in (20).

To formalize this argument, we use the local variance [45] of the estimated PRNU. Given the fingerprint in matrix form, $\hat{\mathbf{K}} \in \mathbb{R}^{N \times M}$, we can estimate the local variance in a neighborhood of radius N_g as

$$[\mathbf{S}]_{m,n}^2 = \frac{1}{(2N_g + 1)^2 - 1} \sum_{i=-N_g}^{N_g} \sum_{j=-N_g}^{N_g} ([\hat{\mathbf{K}}]_{m-i,n-j} - [\hat{\mathbf{M}}]_{m,n})^2, \quad (37)$$

where

$$[\hat{\mathbf{M}}]_{m,n} \doteq \frac{1}{(2N_g + 1)^2} \sum_{i=-N_g}^{N_g} \sum_{j=-N_g}^{N_g} [\hat{\mathbf{K}}]_{m-i,n-j}, \quad (38)$$

is the local average of the estimated PRNU for the entry in position (m, n) . On the boundaries, we apply the symmetric extension.

As already pointed out in [1], assuming that the model is correct, the variance of the fingerprint should be inversely proportional to $\sum_i f(\mathbf{x}_i) \circ f(\mathbf{x}_i)$, where $f(\mathbf{x})$ may be the estimated emphasis, the fixed emphasis, or the linear function in the classical multiplicative model.

To conduct a deeper analysis, we compute the local variance of some of the estimates of the PRNU in the setup of Section 4.1. We show the results in Fig. 8, jointly with their histograms, after normalizing each fingerprint by its global standard deviation and multiplying by the size of the local window minus one (in our case, 80, because $N_g = 4$). For simplicity, we considered the unweighted scenario. Ideally, the variance of the PRNU should be as flat as possible; thus, we want

these histograms to resemble a χ^2 distribution with as many degrees of freedom as the size of the local window minus one (that is, 80).

From the results, we can conclude that the variance of the PRNU is far from being spatially flat. Moreover, for the baseline the tails are heavier than for the emphasized estimators. We argue that the cause is an artifact of model misspecification.

To provide an example, let us assume that among the 15 images used to extract the PRNU, some of them present highly illuminated regions. In these regions, the pixels of the denoised image take large values, but the denoising residual is almost zero. Thus, when applying the estimator in (11), we are adding a very large term in the denominator, which has no corresponding term in the numerator. Therefore, the variance decreases in this region of the fingerprint. That is, this phenomenon creates areas of the estimated PRNU where the local variance is smaller than in the remaining parts of the fingerprint, increasing the weight in the left tail.

To see how this effect translates into the estimation of the cropping point, we analyze the spatial distribution of the peak of the NCC. Under H_1 , this variable should match the cropping point, which in our experiments was uniformly distributed. To avoid clutter, we focus on the errors, that is, the case when the position of the maximum of the NCC differs from the cropping point. We depict in Fig. 9 the position of these peaks of the NCC under H_1 for two cameras, the Canon1100D D3 and the NikonD7000, using the baseline method to estimate the PRNU. Contrary to our expectation, these points do not follow a uniform distribution.

We also depict the local variance of the estimated PRNU (smoothed with a Gaussian kernel to reduce the variance of the estimates and highlight the trends of the surface), with $N_g = 4$, to show the correspondence between both variables: when the local variance of the image increases, the chance of finding an incorrect peak of the NCC also increases.

We interpret these results as follows. In the cropping scenario, we need to find the maximum of the NCC over all the possible alignments. In this vein, regions with larger local variance are more likely to induce larger correlation values, even larger than the correlation for the correct alignment. As a result, these regions are more prone to generate errors.

Regarding the distribution of the signed PCE, we focus on the case when the test image was taken with a different device than the estimated fingerprint, i.e., we consider H_0 . Also, we work with the CCN presented in (21) [22], because the resulting distributions are simpler and the conversion between this quantity and the signed PCE is straightforward. First, note that our argument in the previous paragraph still holds under H_0 : most of the estimated cropping points lie in regions of the estimated PRNU with a large variance. Thus, in most cases, the denominator of the CCN is smaller than the true standard deviation of the maximum of the NCC. As a result, the random variable is not the maximum of a set of standardized Gaussians; instead, it is shifted to the right, which increases the probability of false alarm.

We provide some examples from the H_0 samples in our experimental setup with weighted estimators in Fig. 10. We also provide the maximum of a set of standardized Gaussians: in our setup, we can compute the probability density function (pdf) using

$$p_\theta(x) = \theta \phi(x) \Phi(x)^{\theta-1}, \quad (39)$$

where $\phi(x)$ is the Gaussian pdf, $\Phi(x)$ is the Gaussian cumulative distribution function (cdf), and θ is the number of standardized Gaussians competing for the maximum. In our case, $\theta = (2000 - 512 + 1)^2$. We observe that this effect is more evident for the fingerprint estimated using the baseline method, as expected from our argument above.

Then, the PRNUs estimated using any of the emphases should provide better results than the baseline in the SCA problem. Note that these effects are deeply connected with the leakage problem we studied and reported in our previous work [46,47]. In the following sections, we test this conjecture experimentally.

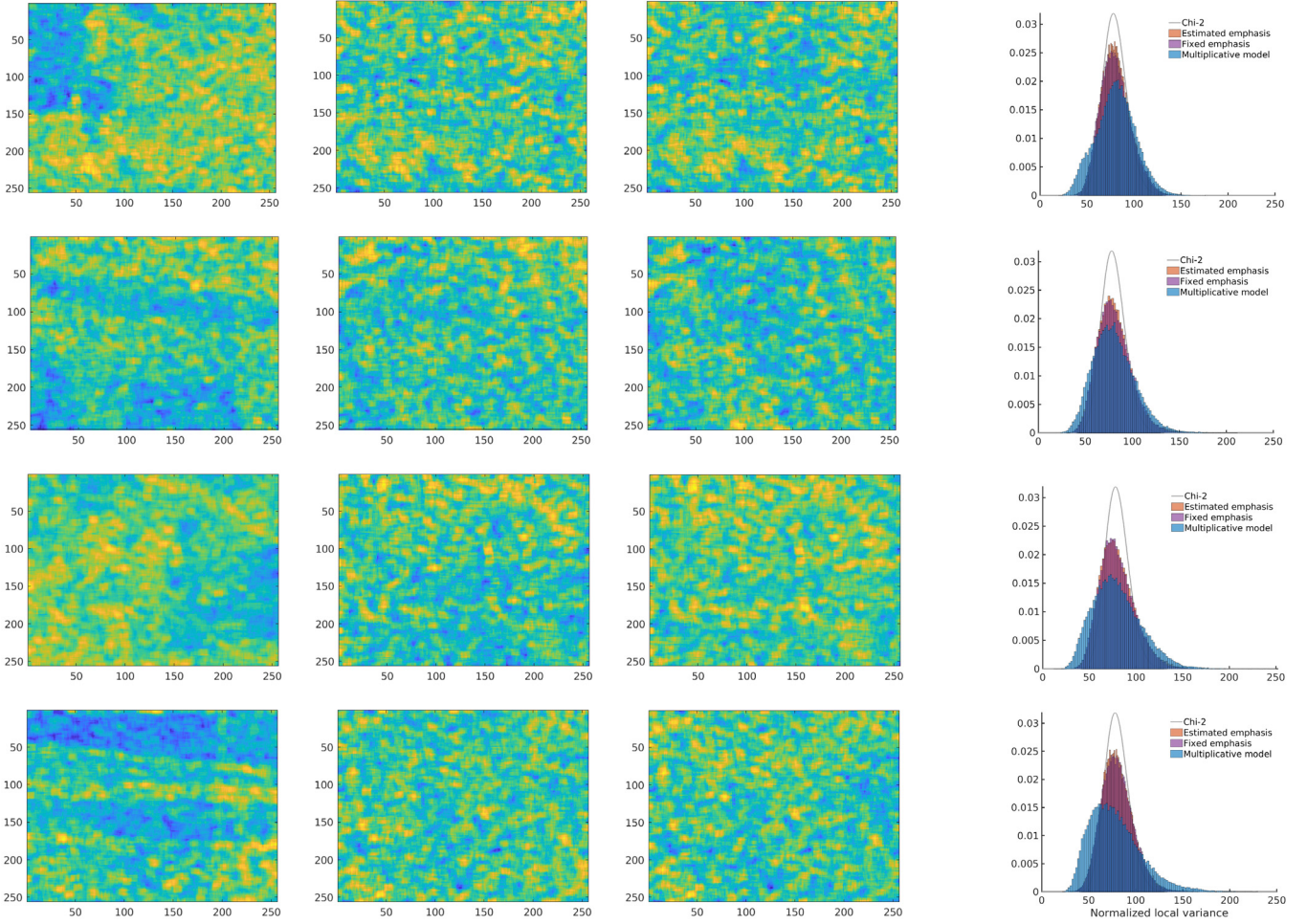


Fig. 8. Local variance of the estimated PRNU using (11) (left), and using (10) with the estimated emphasis (center-left), and the fixed emphasis (center-right). In the rightmost place, we show the histograms, including the probability density function (pdf) of a χ^2 with 80 degrees of freedom. The rows correspond, from top to bottom, to the NikonD60, the NikonD7000, the Canon1100D D3, and the NikonD200 D1.

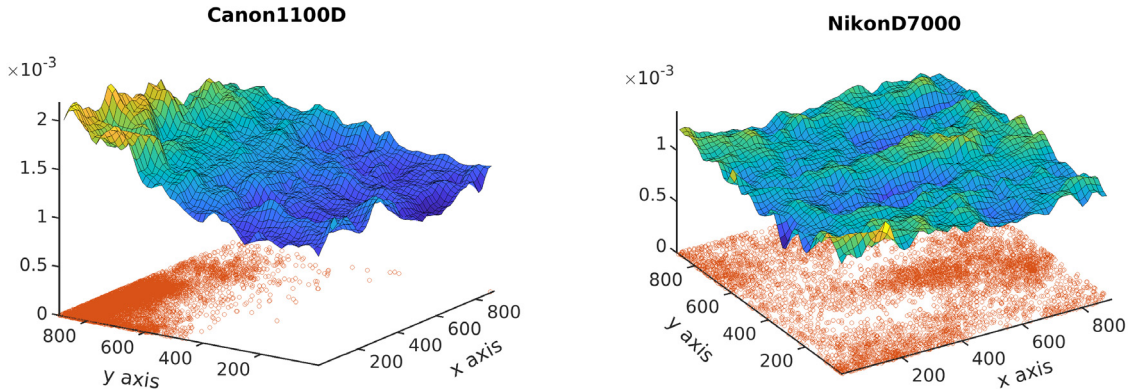


Fig. 9. Position of the peak of the cross-correlation (red circles) when it differs from the true cropping point, jointly with a smoothed surface representing the local variance of $\hat{k}$ using the estimator in (11).

4.3. Uncompressed images and wavelet-based denoising filter

We show the results for uncompressed images and the wavelet-based denoising filter, using the estimators we presented in Section 4.1. We also consider the unweighted versions of the baseline and the emphasized methods to verify our argument in Section 2.4 (Fig. 11, left).

As we expected, the emphasized estimators outperform the baseline method. The WA also performs better than the baseline. Moreover,

the estimated emphasis provides slightly better results than the fixed emphasis. We note that applying the weighting procedure improves the results using the emphasized estimators, but decreases the TPR of the baseline, which is in line with our discussion in Section 2.4.

We provide the individual TPRs for a FPR of 10^{-3} in Table 2 for the WA, the baseline, and the EE methods. The estimated emphasis provides the best results for almost every camera in our database. In some cases, using the WA leads to better results. For the NikonD5100 and the NikonD3200, we argue that they are extreme cases with

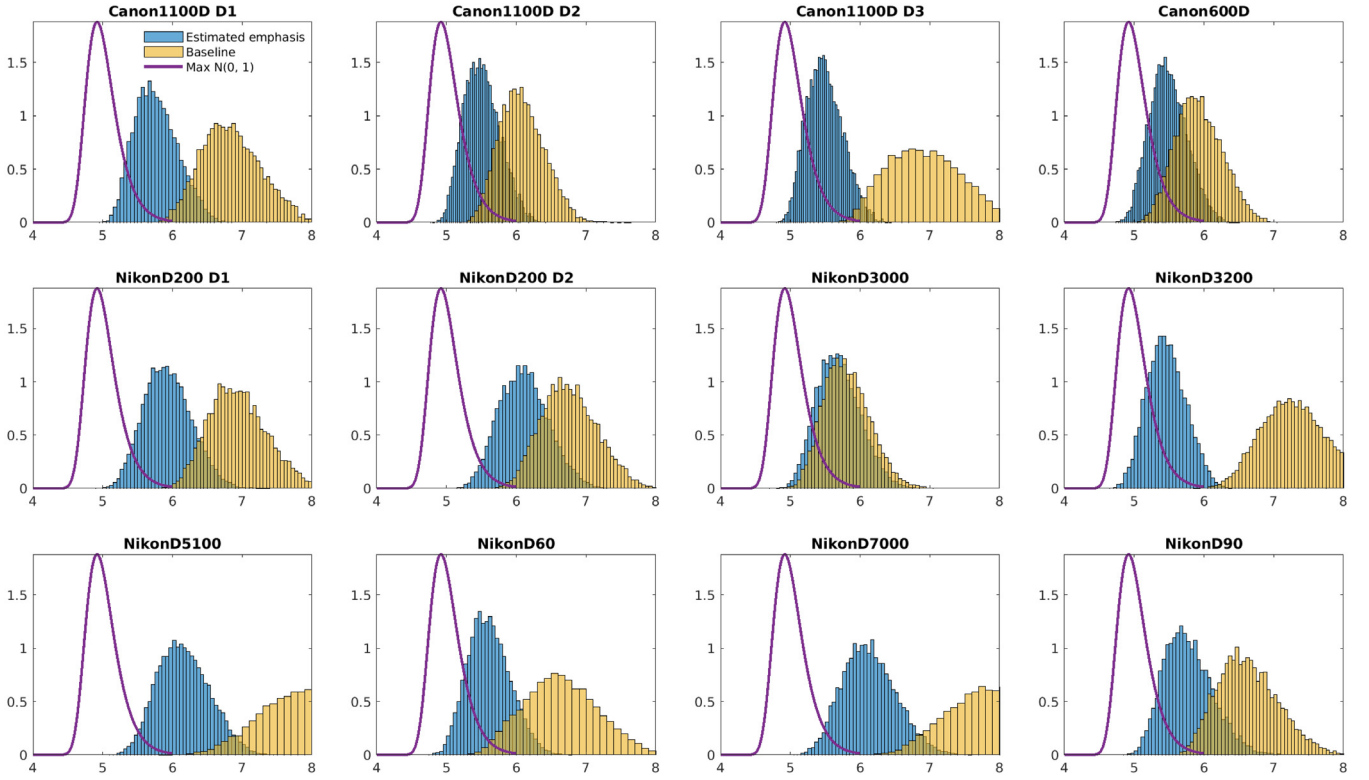


Fig. 10. Distribution of the CCN under H_0 for 12 of the cameras in our dataset, jointly with the theoretical distribution of the maximum of a set of Gaussians. The shift effect is more apparent using the baseline method (multiplicative model) than using the emphasized estimator.

Table 2

TPR for a fixed FPR of 10^{-3} using the wavelet-based denoising filter and uncompressed images. “EE” stands for estimated emphasis. The highest TPR for each camera is written in bold type.

Camera model	WA	EE	Baseline
Canon600D	1.000	1.000	1.000
Canon1100D D1	0.682	0.718	0.307
Canon1100D D2	0.710	0.714	0.440
Canon1100D D3	0.894	0.896	0.444
NikonD60	0.663	0.698	0.458
NikonD90	0.682	0.685	0.598
NikonD200 D1	0.915	0.906	0.837
NikonD200 D2	0.963	0.968	0.941
NikonD3000	0.735	0.738	0.731
NikonD3200	1.000	0.998	0.871
NikonD5100	0.320	0.316	0.098
NikonD7000	0.620	0.758	0.434

very high/very low SNR, respectively. The NikonD200 D1 deserves, however, special consideration. For this camera, most errors appear in particular training sets, where images are textured and the estimated fingerprint is very noisy. Since we rely on the estimated fingerprint to estimate the emphasis, the result is noisy too, which increases the number of missed detections.

4.4. Estimated emphasis via least squares

To check the accuracy of the least squares method presented in Section 3.6, we repeat the simulations above with this algorithm. As described in Section 3.6, we consider two iterations, using the fixed emphasis in the first iteration instead of the linear function. Also, when estimating the PRNU, we apply the weighting procedure described in Section 2.4. We divide each image into 16 blocks to compute and apply the weights. For the basis, we consider again $\beta = 0.003$, its reversed version (along the x -axis), and the inverse parabola.

To test the method in a challenging setup, we estimate the emphasis using only a cropped region of the training images of size 512×512 . The remaining parameters are unaltered. The results show that the least squares method is superior to the fixed emphasis (Fig. 11, right), performing almost as well as our simplified regressogram despite using far less samples.

4.5. Experiments with different denoisers

We also present the results for uncompressed images using the BM3D and the DTCWT denoising filters (Fig. 12). We focus on the weighted scenario. We observe the same trend: the emphasized estimators outperform both the WA and the baseline. In any case, the TPR increases for all methods, which is in line with other studies [8,32].

4.6. Li’s enhancers and phase only method

To contextualize the PRNU emphasis within the PRNU literature, we compare it with two different techniques. First, we show that it improves an already existing method: Li’s enhancers [11], in particular the “Model 3”, which has been exhaustively studied. Second, we test the results using the PRNU emphasis against the “phase only” method by Kang et al. [22], which is incompatible with our proposal but provides competitive results in the aligned case [8].

Regarding Li’s enhancers, the different models are compatible with the PRNU emphasis. Then, we can apply them together (Fig. 13). Introducing the PRNU emphasis improves the performance of the enhancer. However, in line with previous studies [8,22], using the Model 3 degrades the performance of the detector.

On the other hand, the phase-only method [22] is agnostic to the denoised image and thus it is incompatible with the PRNU emphasis. Then, we compare the performance of this method with the results from our emphasis (Fig. 13, right). Our emphasized estimators outperform the phase-only approach for uncompressed images.

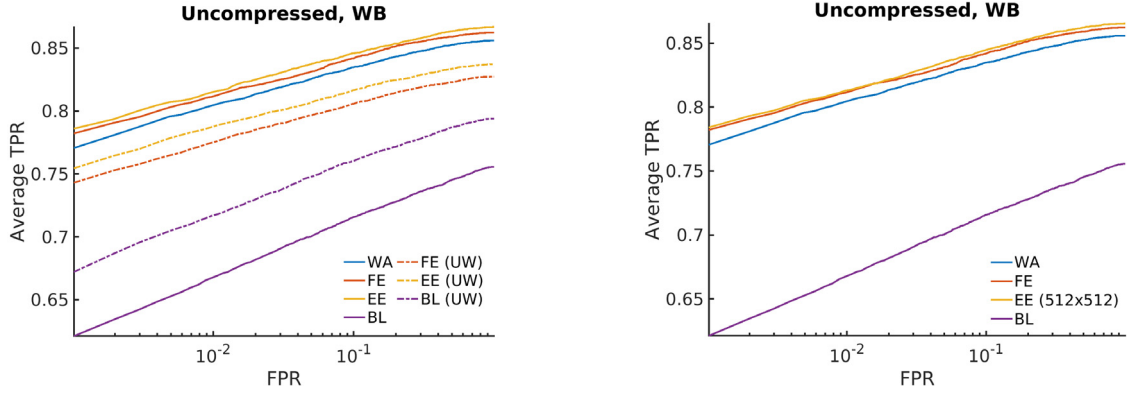


Fig. 11. Modified ROCs considering uncompressed images and wavelet-based denoising. “BL” and “UW” stand for baseline and unweighted, respectively. (Left) Results using the simplified regressogram with two iterations. (Right) Results using the least squares method with images of size 512×512 pixels.

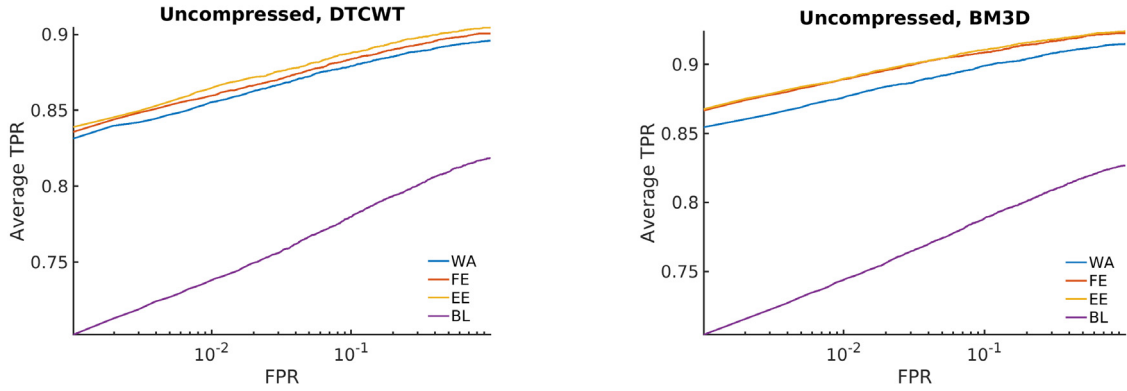


Fig. 12. Modified ROCs considering uncompressed images. “BL” stands for “baseline”. (Left) DTCWT denoising. (Right) BM3D denoising.

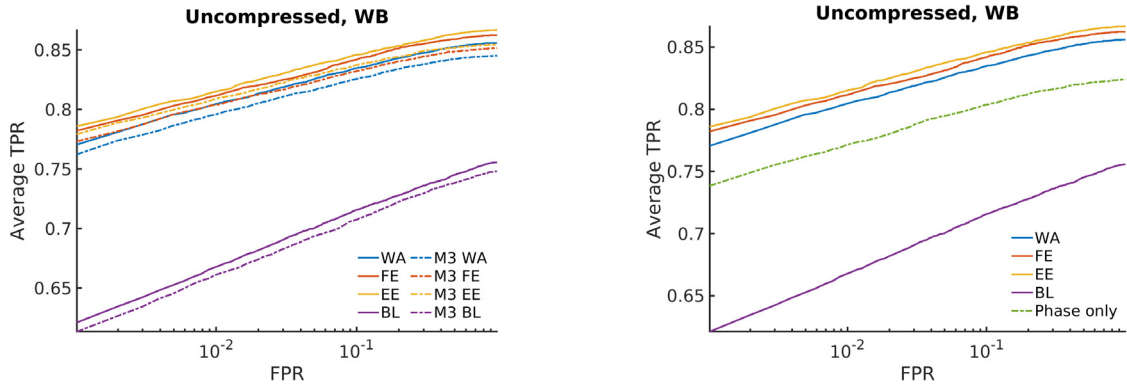


Fig. 13. Modified ROCs considering uncompressed images and wavelet-based denoising. “FE” stands for “fixed emphasis”, while “EE” stands for “estimated emphasis”. (Left) Considering the Model 3 by Li [11], labeled as “M3”. (Right) Including the phase only method proposed by Kang et al. [22].

Nevertheless, the advantage of using only the phase is more evident when considering JPEG compression [22]. In the next section, we test that scenario.

4.7. JPEG compression

We repeat the experiments above considering WB denoising and JPEG-compressed images with Quality Factor (QF) 90. The results show (Fig. 14, left) that using the emphasized estimators improves the average TPR with respect to the baseline method. Given the low TPR, results may be deemed unrepresentative. Thus, we also repeat the experiments with 55 images in the training set. Since we have less than 50 images for the two NikonD200s, we skip these devices. In any case, the same pattern remains, as depicted in Fig. 15. The results using a

quality factor of 98 and 15 images in the training split (Fig. 14, right) show a similar trend.

Although with quality factor 98 the emphasized estimators outperform the phase-only approach, the situation is reversed with quality factor 90. We believe that the method by Kang et al. mitigates the blocking artifacts [48], which is beneficial because they tend to alter the local variance [49]. As we argued in Section 4.2, there is a connection between the local variance, the process to find the cropping point, and the computation of the signed PCE. Thus, when the quality factor decreases, the phase-only method seems to outperform the emphasized methods. When the quality factor increases, the severity of this effect decreases. Then, in the high quality factor scenario, our estimators outperform the method by Kang et al.; one of the reasons may be that retaining only the phase removes some PRNU traces. In any case, we

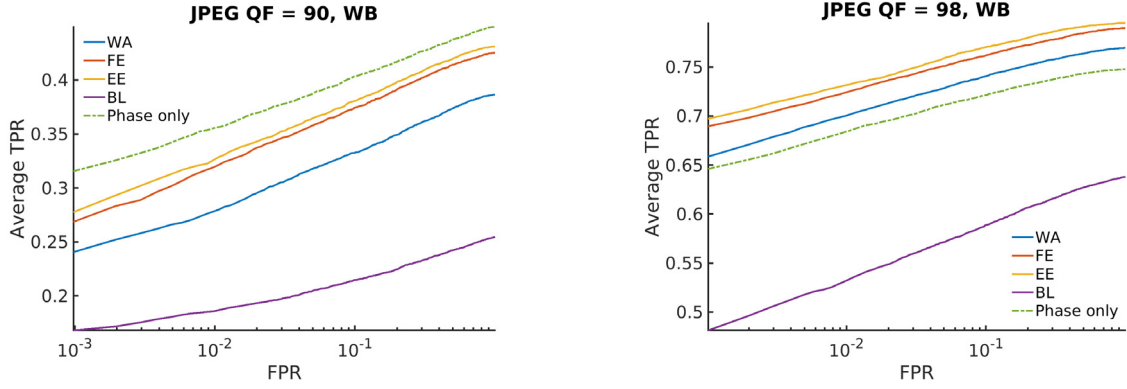


Fig. 14. Modified ROCs considering JPEG compressed images and the wavelet-based denoising filter. We include the method by Kang et al. [22], labeled as “Phase only”. (Left) Quality factor 90. (Right) Quality factor 98.

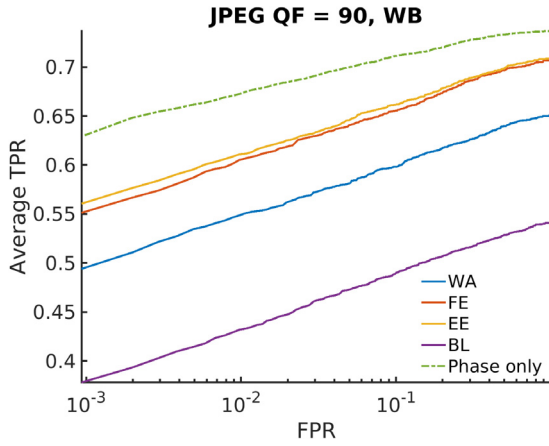


Fig. 15. Modified ROCs considering JPEG compressed images with quality factor 90 and the wavelet based denoising filter. To raise the TPR, we used 55 images in the training split. We include the method by Kang et al. [22], labeled as “Phase only”.

note that the techniques presented in this paper provide a remarkable improvement over the multiplicative model.

Considering either the spatial variance of the noise to estimate the PRNU or different spectral equalization strategies [42] might improve our results.

4.8. Computational complexity

In this subsection, we assess the computational complexity of our method.

Assume we have L images of size $N \times M$. The computational complexity of estimating the PRNU emphasis using the vanilla simplified regressogram (without improvements like computing the local variance) is:

- Computing the L fingerprints using $L-1$ residuals: $2(L-1)NLM + 1$ products, therefore $\mathcal{O}(L^2NM)$.
- Computing the L pairs of fingerprint-residual products: LNM products, therefore $\mathcal{O}(LNM)$.
- Computing the L regressograms: $L-1$ linear time sortings, therefore $\mathcal{O}(LMN)$.

Evaluating the PRNU emphasis is equivalent to evaluating a look-up table; therefore, it has unit complexity.

5. Conclusion

Throughout this paper, we have shown that the transfer curve, that is, the composition of the camera response function with the digital

preprocessing pipeline, induces an emphasis function in the denoising residuals. This emphasis function leads to a new method to estimate the PRNU, which also accounts for the variance of each residual. We have also described different estimators to implement the emphasis in practice. Our results indicate that this curve considerably deviates from the linear function assumed by the multiplicative model. Aiming to be illustrative, we also introduced a method to approximate the transfer curve from the PRNU emphasis, which shows that the result is more complex than a gamma curve.

To test our model, we addressed a SCA problem, with different denoising filters and different levels of JPEG compression. The results show that including the PRNU emphasis provides the highest average TPR for the cameras and the methods we tested, except when dealing with JPEG compression (unless the quality factor is very high), where the method by Kang et al. [22] provides better results. We also showed that using an inverted parabola as fixed emphasis is a competitive option.

As a side note, to estimate the PRNU emphasis we need enough data for each possible pixel intensity. Flat-field images, which are often preferred in practice, may lack this property. Thus, to estimate the emphasis, it may be beneficial to acquire some natural images from the camera under study, alongside the flat-field images used to estimate the fingerprint. Another option is resorting to the fixed emphasis.

We are investigating several extensions to this work. First, we plan to provide a method to account for the local variance of the PRNU; doing so should boost the results given our discussion in Section 4.2. Furthermore, this might reduce the leakage of the estimated fingerprint [46,47] and help in the detection of local forgeries and manipulations. Second, some work is necessary in the mathematical front to incorporate formally the PRNU emphasis and the weighting approach under our new model. This line of work should improve the performance of the detector in both the aligned and the unaligned setups. Finally, we are also considering the application of the emphasis with different PRNU representations, as the “significant components only” method by Hu et al. [50], and its possible impact on the conversion from RGB to grayscale, including the interplay with CFA arrays [51].

CRediT authorship contribution statement

Samuel Fernández-Mendiña: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – original draft, Visualization. **Fernando Pérez-González:** Conceptualization, Methodology, Validation, Formal analysis, Resources, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition. **Miguel Masciopinto:** Conceptualization, Methodology, Software, Validation, Investigation, Data curation, Resources, Writing – review & editing, Visualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was funded by Agencia Estatal de Investigación (Spain) and the European Regional Development Fund (FEDER “Una manera de hacer Europa”) under project RODIN (PID2019-105717RB-C21), and by Xunta de Galicia, Spain “Grupos de Referencia Competitiva” (ED431C 2021/47).

Appendix. Multiplicative model and transfer curve

In this appendix, we provide a result relating the models in (3) and (7).

Proposition 1. *Let $x \doteq h(z)$. Then, the output of the camera response function can be written as*

$$y = x(1 + ck) + \mathcal{O}(\sigma_k^2 k^2), \quad (\text{A.1})$$

for some real constant c if and only if $h(z) = c_1 z^\gamma$ for some real positive constants c_1 and γ .

To prove the “if” part, replace $h(z) = c_1 z^\gamma$ in (7). The “only if” part follows from solving the following equation

$$h(z)(1 + ck) = h(z) + zh'(z)c_1 k \implies c_2 = zh'(z)/h(z).$$

Rewriting as $h'(z)/h(z) = c_2/z$ and integrating, we obtain $\log(h(z)) = c_2 \log(z) + c_3$; thus, $h(z) = e^{c_3} z^{c_2}$.

References

- [1] M. Chen, J. Fridrich, M. Goljan, J. Lukáš, Determining image origin and integrity using sensor noise, *IEEE Trans. Inf. Forensics Secur.* 3 (1) (2008) 74–90.
- [2] J. Lukáš, J. Fridrich, M. Goljan, Digital camera identification from sensor pattern noise, *IEEE Trans. Inf. Forensics Secur.* 1 (2) (2006) 205–214.
- [3] M. Kirchner, C. Johnson, SPN-CNN: Boosting sensor-based source camera attribution with deep learning, in: 2019 IEEE International Workshop on Information Forensics and Security, WIFS, IEEE, 2019, pp. 1–6.
- [4] D. Cozzolino, L. Verdoliva, Noiseprint: A CNN-based camera model fingerprint, *IEEE Trans. Inf. Forensics Secur.* 15 (2019) 144–159.
- [5] M. Chen, J. Fridrich, M. Goljan, J. Lukáš, Source digital camcorder identification using sensor photo response non-uniformity, in: Security, Steganography, and Watermarking of Multimedia Contents IX, Vol. 6505, International Society for Optics and Photonics, 2007, p. 65051G.
- [6] M. Goljan, J. Fridrich, T. Filler, Large scale test of sensor fingerprint camera identification, in: Media Forensics and Security, Vol. 7254, International Society for Optics and Photonics, 2009, p. 72540I.
- [7] F. Pérez-González, M. Masciopinto, I. González-Iglesias, P. Comesaña, Fast sequential forensic detection of camera fingerprint, in: 2016 IEEE International Conference on Image Processing, ICIP, 2016, pp. 3902–3906, <http://dx.doi.org/10.1109/ICIP.2016.7533091>.
- [8] M. Al-Ani, F. Khelifi, On the SPN estimation in image forensics: A systematic empirical evaluation, *IEEE Trans. Inf. Forensics Secur.* 12 (5) (2017) 1067–1081.
- [9] M. Masciopinto, F. Pérez-González, Putting the PRNU model in reverse gear: Findings with synthetic signals, in: 2018 26th European Signal Processing Conference, EUSIPCO, IEEE, 2018, pp. 1352–1356.
- [10] M.D. Grossberg, S.K. Nayar, Modeling the space of camera response functions, *IEEE Trans. Pattern Anal. Mach. Intell.* 26 (10) (2004) 1272–1282.
- [11] C.-T. Li, Source camera identification using enhanced sensor pattern noise, *IEEE Trans. Inf. Forensics Secur.* 5 (2) (2010) 280–287, <http://dx.doi.org/10.1109/TIFS.2010.2046268>.
- [12] M. Goljan, J. Fridrich, Camera identification from cropped and scaled images, in: Security, Forensics, Steganography, and Watermarking of Multimedia Contents X, Vol. 6819, International Society for Optics and Photonics, 2008, p. 68190E.
- [13] M.D. Grossberg, S.K. Nayar, What is the space of camera response functions? in: 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings, Vol. 2, 2003, pp. II–602, <http://dx.doi.org/10.1109/CVPR.2003.1211522>.
- [14] CAVE, Camera response functions: Database (DoRF) and model (EMOR), 2003, Online URL <https://www.cs.columbia.edu/CAVE/software/softlib/dorf.php>.
- [15] B.K. Gunturk, J. Glotzbach, Y. Altunbasak, R.W. Schafer, R.M. Mersereau, Demosaicking: Color filter array interpolation, *IEEE Signal Process. Mag.* 22 (1) (2005) 44–54.
- [16] K.R. Rao, P. Yip, Discrete Cosine Transform: Algorithms, Advantages, Applications, Academic Press, 2014.
- [17] P. Milanfar, A tour of modern image filtering: New insights and methods, both practical and theoretical, *IEEE Signal Process. Mag.* 30 (1) (2012) 106–128.
- [18] A. Lawgaly, F. Khelifi, A. Bouridane, Weighted averaging-based sensor pattern noise estimation for source camera identification, in: 2014 IEEE International Conference on Image Processing, ICIP, IEEE, 2014, pp. 5357–5361.
- [19] A.C. Aitken, On least squares and linear combination of observations, *Proc. Roy. Soc. Edinburgh* 55 (1936) 42–48.
- [20] C.R. Rao, Estimation of heteroscedastic variances in linear models, *J. Amer. Statist. Assoc.* 65 (329) (1970) 161–172.
- [21] M. Goljan, J. Fridrich, Sensor fingerprint digests for fast camera identification from geometrically distorted images, in: Media Watermarking, Security, and Forensics 2013, Vol. 8665, International Society for Optics and Photonics, 2013, p. 86650B.
- [22] X. Kang, Y. Li, Z. Qu, J. Huang, Enhancing source camera identification performance with a camera reference phase sensor pattern noise, *IEEE Trans. Inf. Forensics Secur.* 7 (2) (2011) 393–402.
- [23] J.W. Tukey, Exploratory Data Analysis, Vol. 2, Reading, MA, 1977.
- [24] J.W. Tukey, Curves as parameters, and touch estimation, in: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics, The Regents of the University of California, 1961, pp. 681–694.
- [25] N.J. Higham, Computing a nearest symmetric positive semidefinite matrix, *Linear Algebra Appl.* 103 (1988) 103–118.
- [26] G.H. Golub, C.F. Van Loan, Matrix Computations, Fourth ed., in: Johns Hopkins studies in the mathematical sciences, Johns Hopkins University Press, Baltimore, MD, 2013.
- [27] T. Gloe, R. Böhme, The ‘Dresden image database’ for benchmarking digital image forensics, in: Proceedings of the 2010 ACM Symposium on Applied Computing, SAC ’10, Association for Computing Machinery, New York, NY, USA, 2010, pp. 1584–1590, <http://dx.doi.org/10.1145/1774088.1774427>.
- [28] D.-T. Dang-Nguyen, C. Pasquini, V. Conotter, G. Boato, RAISE: A raw images dataset for digital image forensics, in: Proceedings of the 6th ACM Multimedia Systems Conference, 2015, pp. 219–224.
- [29] M.K. Mihçak, I. Kozintsev, K. Ramchandran, Spatially adaptive statistical modeling of wavelet image coefficients and its application to denoising, in: 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 6, 1999, pp. 3253–3256, <http://dx.doi.org/10.1109/ICASSP.1999.757535>.
- [30] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, Color image denoising via sparse 3D collaborative filtering with grouping constraint in luminance-chrominance space, in: 2007 IEEE International Conference on Image Processing, Vol. 1, IEEE, 2007, pp. I–313.
- [31] G. Chierchia, S. Parrilli, G. Poggi, C. Sansone, L. Verdoliva, On the influence of denoising in PRNU based forgery detection, in: Proceedings of the 2nd ACM Workshop on Multimedia in Forensics, Security and Intelligence, 2010, pp. 117–122.
- [32] H. Zeng, M.D.M. Hosseini, K. Deng, A. Peng, M. Goljan, A comparison study of CNN denoisers on PRNU extraction, 2021, arXiv preprint [arXiv:2112.02858](https://arxiv.org/abs/2112.02858).
- [33] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, Image denoising by sparse 3-D transform-domain collaborative filtering, *IEEE Trans. Image Process.* 16 (8) (2007) 2080–2095, <http://dx.doi.org/10.1109/TIP.2007.901238>.
- [34] H. Zeng, Y. Wan, K. Deng, A. Peng, Source camera identification with dual-tree complex wavelet transform, *IEEE Access* 8 (2020) 18874–18883, <http://dx.doi.org/10.1109/ACCESS.2020.2968855>.
- [35] I.W. Selesnick, R.G. Baraniuk, N.C. Kingsbury, The dual-tree complex wavelet transform, *IEEE Signal Process. Mag.* 22 (6) (2005) 123–151.
- [36] M. Yaghoobi, L. Daudet, M.E. Davies, Parametric dictionary design for sparse coding, *IEEE Trans. Signal Process.* 57 (12) (2009) 4800–4810.
- [37] P.C. Hansen, Regularization tools: A matlab package for analysis and solution of discrete ill-posed problems, *Numer. Algorithms* 6 (1) (1994) 1–35.
- [38] G.H. Golub, M. Heath, G. Wahba, Generalized cross-validation as a method for choosing a good ridge parameter, *Technometrics* 21 (2) (1979) 215–223.
- [39] G. Wahba, Spline Models for Observational Data, SIAM, 1990.
- [40] M.D. Grossberg, S.K. Nayar, What can be known about the radiometric response from images? in: European Conference on Computer Vision, Springer, 2002, pp. 189–205.
- [41] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes : The Art of Scientific Computing, Cambridge University Press, Cambridge, UK New York, 2007.

- [42] X. Lin, C.-T. Li, Preprocessing reference sensor pattern noise via spectrum equalization, *IEEE Trans. Inf. Forensics Secur.* 11 (1) (2015) 126–140.
- [43] S. Fernández-Menduiña, F. Pérez-González, M. Masciopinto, PRNU emphasis: A generalization of the multiplicative model, 2022, [arXiv:2202.12357](https://arxiv.org/abs/2202.12357).
- [44] W. Yaqub, M. Mohanty, N. Memon, Towards camera identification from cropped query images, in: 2018 25th IEEE International Conference on Image Processing, ICIP, IEEE, 2018, pp. 3798–3802.
- [45] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: From error visibility to structural similarity, *IEEE Trans. Image Process.* 13 (4) (2004) 600–612.
- [46] S. Fernández-Menduiña, F. Pérez-González, On the information leakage quantification of camera fingerprint estimates, *EURASIP J. Inform. Secur.* 2021 (1) (2021) 1–13.
- [47] F. Pérez-González, S. Fernández-Menduiña, PRNU-leaks: Facts and remedies, in: 2020 28th European Signal Processing Conference, EUSIPCO, IEEE, 2021, pp. 720–724.
- [48] Z. Wang, A.C. Bovik, B.L. Evan, Blind measurement of blocking artifacts in images, in: *Proceedings 2000 International Conference on Image Processing*, Vol. 3, Ieee, 2000, pp. 981–984.
- [49] G. Yammine, E. Wige, A. Kaup, A no-reference blocking artifacts visibility estimator in images, in: 2010 IEEE International Conference on Image Processing, IEEE, 2010, pp. 2497–2500.
- [50] Y. Hu, B. Yu, C. Jian, Source camera identification using large components of sensor pattern noise, in: 2009 2nd International Conference on Computer Science and Its Applications, IEEE, 2009, pp. 1–5.
- [51] C.-T. Li, Y. Li, Color-decoupled photo response non-uniformity for digital image forensics, *IEEE Trans. Circuits Syst. Video Technol.* 22 (2) (2011) 260–271.