Highlights

Radon transform of image monotonic rearrangements as feature for noise sensor signature

Vittoria Bruni*, Silvia Marconi, Giuseppina Monteverde, Domenico Vitulano

- The paper investigates some mathematical properties of the Photo Response Non Uniformity Pattern noise that can be used for the identification of the source device of an image.
- The paper proves that the 2D decreasing rearrangement of PRNU image provides specific and device-dependent geometric structures that can be properly coded by the Radon transform.
- The empirical distribution of the Radon transform of PRNU 2D decreasing rearrangement is a feature robust to image manipulations and independent of image size.

Radon transform of image monotonic rearrangements as feature for noise sensor signature

Vittoria Bruni*, Silvia Marconi, Giuseppina Monteverde, Domenico Vitulano

[a]*Department of Basic and Applied Sciences for Engineering, Sapienza University of Rome, via Antonio Scarpa 16, Rome, 00161, Italy*

Abstract

Source camera identification represents a delicate, crucial but challenging task in digital forensics, especially when an image has to be used as a proof in a court of law. This paper investigates some properties of the Photo Response Non Uniformity (PRNU) pattern noise that represents the fingerprint of any acquisition sensor. The main goal is to define specific and distinctive features for this noise source that enable the identification of the acquisition sensor by simply analysing a single image. These features are required to be independent of image size, modifications, storage mode, etc. To this aim the discrimination power of the decreasing rearrangement of a function, combined with the Radon transform, has been investigated. Preliminary tests show that a proper rearrangement of PRNU image provides specific and device-dependent geometric structures that can be properly coded through the Radon transform. In particular, the empirical distribution of the Radon Transform of rearranged Flat Field images alone is capable to correctly characterize each device with high accuracy, showing robusteness to some standard image modifications, such as quantization and blurring; in addition, it guarantees independence of image size.

Keywords: Radon Transform, function monotonic rearrangements, source camera identification, PRNU

1. Introduction

In the last decade, the authentication of visual documents for their validation in the forensic field has gained increasing importance because of the expansion of communication networks and the massive dissemination of smart

devices [36]. The intensive use of images, and more in general multimedia content, plays a dual role: on the one hand, it makes images crucial for investigation purposes; on the other hand, it makes investigations difficult from different points of view as, for example, large amount of data to process and analyse, authenticity assessment of image content, identification of the source generating the image, robustness to image manipulation and modification due to transmission, storage, posting on social networks and so on. This contrasting aspect opened the way to different research topics and problems that are far from to be solved, including source camera identification. The latter aims at assessing the origin of an image, by determining the source sensor in terms of device type, brand, model till the specific device. The aim is to retrieve a source fingerprint, as for example the Photo Response Non-Uniformity pattern noise (PRNU), that allows to establish a correspondence between a device and the images that have been acquired by it. In particular, it is expected that devices equipped with the same sensor left a similar fingerprint; on the other hand, each device has its own fingerprint that should allow to distinguish it from any other device. In fact, PRNU consists of a noise image component that is caused by the CCD imperfections [30]. As a result, specific pixels are susceptible to giving brighter intensities than others. This causes the signature of the specific device that is different even among devices of the same brand and model [13]. Unfortunately, retrieving this fingerprint is a hard task due to the complexity of the acquisition process that makes harder PRNU modeling. In addition, images can be further manipulated after their acquisition; for example, they can be stored in different formats, sent by email, shared in social networks and so on, so that their original content, as well as their PRNU component, may be severely compromised or modified. PRNU is often modeled as a multiplicative noise source [13] and defined as one of the main components of the residual image deriving from a regularization of the original image. Therefore, the residual image is used for PRNU characterization purposes. It is then evident the importance of having a feasible and reliable mathematical model for the acquired image that allows for a straightforward PRNU extraction through a non linear regularization process; in addition, it is fundamental to define projection or approximation procedures that allow us to emphasize some peculiar PRNU features, while maintaining robustness to eventual inaccuracies in the adopted model. This is the reason why the literature concerning source camera identification focuses on both PRNU extraction and PRNU classification, either separately or jointly — see [1, 3, 10, 11, 14, 27, 29, 30, 31, 47] and references therein.

A sketch of the source identification process is depicted in Fig. 1. As it can be observed, regularization plays a fundamental and dual role in PRNU extraction as it is used in the definition of both the *reference PRNU*, i.e. the one characterizing the device, and the *image PRNU*, i.e. the one extracted from the single image. Inspired by the pioneering paper in [30], the existing literature mainly concerns on the definition of a better denoiser [2, 35, 37, 26], a better noise modeling [22, 32, 39], or on a proper selection of the image domain where to extract reliable information [8, 10, 29, 40, 41]. Regarding the estimation of the *reference PRNU*, denoising is combined with enhacement operations that mainly aim at estimating the reference PRNU from several images acquired by the same device. Maximum likelihood estimators are commonly employed with the purpose of suppressing artifacts and additional noise sources in the single residuals [13].

As far it concerns PRNU classification, the literature focuses on the definition of similarity metrics, features extraction, clustering methods [1, 9, 31, 47], and, more recently, on machine and deep learning methods properly built to accomplish this task for specific use cases [11, 15, 18, 24, 33, 43, 44, 45, 46]. In fact, the best methodology cannot be obviously independent of the specific working scenario [16]. In *closed-set* scenarios the task is to establish if a given image has been captured by a device among a predefined and available set. On the contrary, in *open-set* scenarios, the goal is to establish if two images have been acquired by the same device, independently of the specific device kind. However, in both cases, the main requirement is robustness to image manipulations, such as filtering, quantization, enhancement, or specific transmission/storage process, as for example, uploading and downloading from social networks. That is why several benchmarking datasets [12, 23, 38] have been built and made available to the scientific community; each of them with the specific aim of collecting images able to replicate the real scenarios: different social networks, device type and models, more common manipulations.

Although the plethora of existing approaches, the mathematical modeling [32] of such a problem mainly refers to the pioneering paper in [13] and remains an open research field. One of the main problems that arises when comparing *image PRNU* and *reference PRNU* is the image size. In fact, on the one hand the same device allows to set the resolution of the output image; on the other hand, image manipulations, such as sampling and interpolation, modify image size. As a result, there is the need of having classification methods that are independent of image size. Projecting a function into a

proper domain allows to emphasize some of its features, making it distinguishable from other functions. Projecting in a n-dimensional feature space, where n is fixed, allows us to compare functions (in our specific case, images) of different nature and size.

Based on these considerations, in this paper we investigate some further PRNU properties and evaluate their invariance to some modifications of the original acquisition. To this aim, flat field (FF) images have been considered as PRNU is the dominant component in such images. A specific 2-dimensional rearrangement along the two spatial directions has been applied to the image [4]. This spatial rearrangement is strictly linked to the noise cumulative distribution and highlights some geometrical structures that are peculiar for each device model and can be compactly represented by the Radon transform. The n-dimensional space is then set in this paper as the empirical probability of the Radon Transform of the 2-dimensional rearrangement of the acquired image, where n indicates the number of bins. Preliminary results achieved on flat field images from Dresden database [25] show that the proposed feature vector allows for the detection of the image source device in closed-set scenarios with high accuracy by using standard distances for probability density functions. In addition, it allows us to compare PRNU images having different size, showing robustness to some common image manipulations. Finally, the proposed feature is useful whenever only few images are available for the estimation of the *reference PRNU*.

The remainder of the paper is the following. Next section presents the conventional PRNU mathematical modeling and the main steps of the source identification process. Section 3 is devoted to the description of the proposed feature-based method. Section 4 contains some preliminary results concerning recognition accuracy, comparative studies and robustness to image manipulations. Finally, the last section draws the conclusions.

2. PRNU modeling

The image formation process is quite complex and camera/manufacturers-dependent. It includes signal quantization, white balance, demosaicking (color interpolation), color correction, gamma correction, filtering, and, often, JPEG compression. Based on these considerations, the image J at the pixel spatial location $\mathbf{x} = (x, y)$ that is acquired by a given device $\tilde{d}$ is modelled as it follows [13]

$$J(\mathbf{x}) = g^{\gamma}(Y(\mathbf{x})(1 + \bar{K}_{\tilde{d}}(\mathbf{x})) + \Lambda(\mathbf{x}))^{\gamma} + \Theta_q(\mathbf{x}), \quad \forall\, \mathbf{x} \in \Omega \subset \mathbf{R}^2 \qquad (1)$$

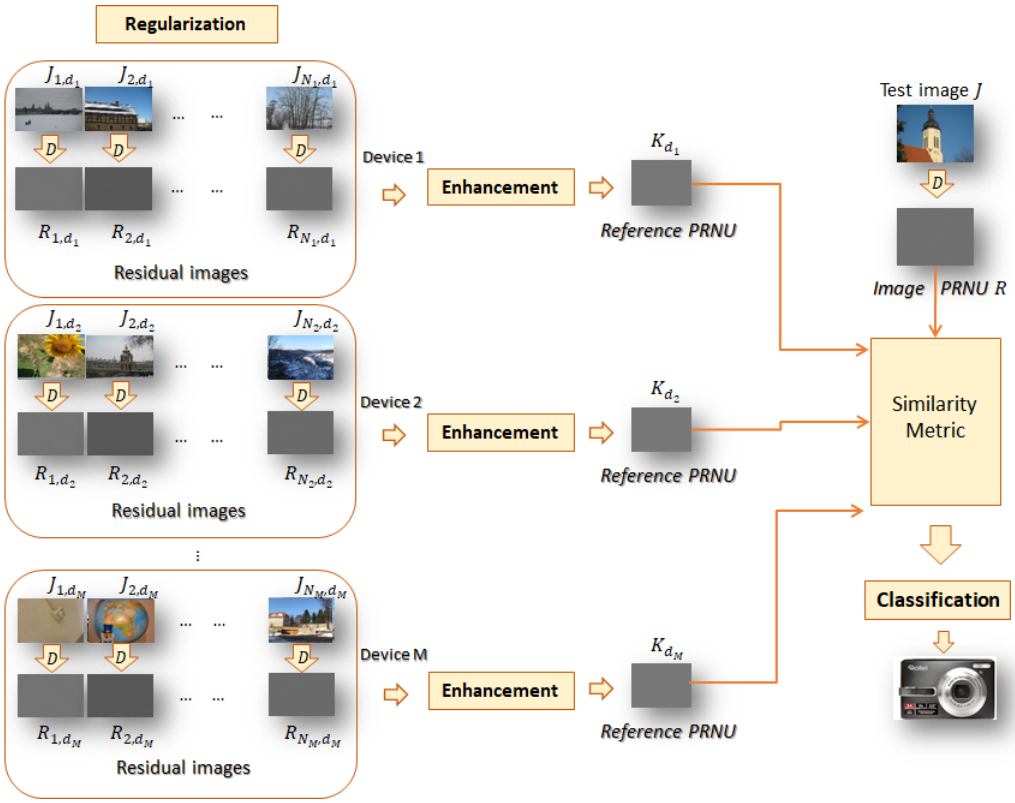

Figure 1: Block scheme of the source identification process in a closed-set scenario. The main phases are: regularization, enhancement and classification.

where $g \in \mathbf{R}^+$ denotes the color channel gain, $\gamma \in \mathbf{R}^+$ is the gamma correction exponent, Y is the scene light intensity, $\bar{K}_d$ is the device PRNU, Λ represents other noise sources including the dark current, shot noise, and read-out noise, while Θ_q is the quantization noise.

As Y is the dominant term, eq. (1) can be rewritten as it follows

$$J(\mathbf{x}) = (g\,Y(\mathbf{x}))^\gamma \left(1 + \bar{K}_{\tilde{d}}(\mathbf{x}) + \frac{\Lambda((\mathbf{x}))}{Y(\mathbf{x})}\right)^\gamma + \Theta(\mathbf{x})$$

and the first two terms in the Mc Laurin expansion of the function $(1+t)^\gamma$, with $t = K + \frac{\Lambda}{Y}$, can be kept, i.e.,

$$J(\mathbf{x}) = (g\,Y(\mathbf{x}))^\gamma \left(1 + \gamma\bar{K}_{\tilde{d}}(\mathbf{x}) + \gamma\frac{\Lambda(\mathbf{x})}{Y(\mathbf{x})}\right) + \Theta_q(\mathbf{x}).$$

By setting $I = (g\,Y(\mathbf{x}))^\gamma$, we have

$$J(\mathbf{x}) = I(\mathbf{x})(1 + K_{\tilde{d}}(\mathbf{x})) + \Theta(\mathbf{x}), \tag{2}$$

where I represents the original image content (output of the sensor without noise), $K_{\tilde{d}}$ is the device PRNU multiplied by γ, while Θ includes noise sources that are independent of $K_{\tilde{d}}$ [30]. $K_{\tilde{d}}$ is modelled as a zero-mean noise component that is independent of I and it represents the fingerprint of the device $\tilde{d}$ that took the image J.

With reference to the scheme in Fig. 1, the *image PRNU* commonly is the estimation of $K_{\tilde{d}}$ from a single image. Let $\mathcal{D}$ denote a regularization operator (denoiser), and let $\hat{I} = \mathcal{D}(J)$, with J defined as in eq. (2), be an estimate of the original image I. The *residual image* R is defined as

$$R(\mathbf{x}) = J(\mathbf{x}) - \hat{I}(\mathbf{x})$$

and a simple algebra [13] allows us to rewrite R as it follows

$$R(\mathbf{x}) = J(\mathbf{x})K_{\tilde{d}}(\mathbf{x}) + \bar{\Theta}(\mathbf{x}), \tag{3}$$

where $\bar{\Theta}$ represents all noise sources, including the one introduced by the regularization procedure. $\bar{\Theta}$ is assumed to be nearly independent of $JK_{\tilde{d}}$.

Regarding the *reference PRNU*, let $\{J_{i,d_j}\}_{i=1,\ldots,N_j}$ a set of N_j images acquired by the device d_j. For each image J_{i,d_j}, eqs. (2) and (3) hold true.

Enhancement then consists of suppressing noise contribution in the residual images

$$R_{i,d_j}(\mathbf{x}) = J_{i,d_j}(\mathbf{x})K_{d_j}(\mathbf{x}) + \bar{\Theta}_{i,d_j}(\mathbf{x}), \quad i = 1, ..., N_j. \tag{4}$$

In particular, by assuming the noise sources $\bar{\Theta}_{i,d_j}$ i.i.d., the maximum likelihood estimator [13, 30, 31] provides the following estimation for K_{d_j}

$$K_{d_j}(\mathbf{x}) = \frac{\sum_{i=1}^{N_j} R_{i,d_j}(\mathbf{x})J_{i,d_j}(\mathbf{x})}{\sum_{i=1}^{N_j} J^2_{i,d_j}(\mathbf{x})}, \quad \forall\, \mathbf{x} \in \Omega \subset \mathbf{R}^2. \tag{5}$$

Finally, to establish if the image J has been acquired by the device d_j, i.e. if $\tilde{d} = d_j$, *image PRNU* in eq. (3) and *reference PRNU* in eq. (5) have to be compared. The normalized correlation [34] is often used to address this issue. It is defined as

$$\rho(K_{d_j}, R) = \frac{\sigma_{K_{d_j}R}}{\sigma_{K_{d_j}}\sigma_R}, \tag{6}$$

where $\sigma_{K_{d_j}R}$ is the cross correlation between K_{d_j} and R, while $\sigma_{K_{d_j}}$ and σ_R are the standard deviations of K_{d_j} and R respectively. The closer to 1 ρ, the more reliable the match between d_j and $\tilde{d}$.

Finally, the task in closed-set scenarios is then accomplished by finding the device $d_{\hat{j}}$, among a set of available sensors $d_j, \quad j = 1, ..., M$, such that

$$\hat{j} = argmax_{j=1,...,M} \quad \rho(K_{d_j}, R). \tag{7}$$

As mentioned in the introduction, both denoising and enhancement play a crucial role in PRNU estimation. The former heavily influences both device and image fingerprint extraction; the latter is directly related to the *reference PRNU*, and can be somewhat dependent on the number N_j of available images. In addition, the direct comparison of *image PRNU* and *reference PRNU* is limited by the constraint on the size of the images. In order to contribute to the definition of different criteria in fingerprint estimation, the decreasing rearrangement [4, 19, 6, 7] of flat field images has been considered in this paper, as described in the following section.

3. The proposed model

A source identification method is required: *i)* to be robust to image manipulation; *ii)* to be independent of the size of the comparing images. To

accomplish the second requirement, a reasonable strategy can be the projection of the image in a feature space of fixed dimension where some significant elements of the original image are emphasized. If the features are robust to some image manipulations, the first task is accomplished too. To this aim, the 2D monotonic rearrangement of FF images is studied in this section.

Let us then consider the model in eq. (2) and let us suppose uniform light intensity, i.e. $Y(\mathbf{x}) = C, \quad \forall\, \mathbf{x} \in \Omega$, where C is constant. This corresponds to select almost constant images, commonly denoted as flat field (FF) images. With these assumptions, eq. (2) reduces to

$$J(\mathbf{x}) = C(1 + K_{\tilde{d}}(\mathbf{x})) + \Theta(\mathbf{x}), \tag{8}$$

where C denotes the intensity of the constant image background. Despite the presence of the noise component Θ, FF images better correlate with $K_{\tilde{d}}$. This is the reason why they are the preferred and recommended candidates for *reference PRNU* estimation [30]. Hence, the set of available images is supposed to satisfy the following equation

$$J_{i,d_j}(\mathbf{x}) = C_{i,d_j} + C_{i,d_j} K_{d_j}(\mathbf{x}) + \Theta_{i,d_j}(\mathbf{x}), \quad i = 1, .., N_j, \tag{9}$$

where C_{i,d_j} are constant values.

J_{i,d_j} mainly depends on noise sources having different origins and distributions; on the other hand, the decreasing rearrangment [19, 20, 21] of a function f, namely f^*, is a monotonically decreasing function and can be defined as the generalized inverse of the cumulative distribution function $M_f(y)$ of f, according to the following definition.

Definition 1 Let μ be a measure, $f(t) : \Omega_t \to \mathbb{R}$ be a μ-measurable and non negative function and

$$M_f(y) = \mu(\{t \in \Omega_t : f(t) > y\}). \tag{10}$$

The decreasing (non-increasing) rearrangement f^* of f is the function $f^*(t) : [0, \mu(\Omega_t)] \to \mathbb{R}$ such that

$$f^*(t) = M_f^{-1}(t) = \inf\{y \in f(\Omega_t) : M_f(y) \leq t\}. \tag{11}$$

f^* is then strictly related to the distribution of the function itself and satisfies the following properties:

1. $f \leq g \quad \Rightarrow \quad f^* \leq g^*$, with f and g two real-valued and non-negative functions;
2. $f_n \to f \quad \Rightarrow \quad f_n^* \to f^*$, with f_n a sequence of real-valued and non-negative functions;
3. if f is $\alpha-$ Lipschitz with Lipschitz constant L, then f^* is $\alpha-$ Lipschitz with constant $L^* \leq L$;
4. $\int_A f(t)dt = \int_{[0,\mu(A)]} f^*(t)dt$;
5. $\int_\Omega f(t)g(t)dt \leq \int_{[0,+\infty]} f^*(t)g^*(t)dt$.

It is worth observing that f^* preserves the regularity/irregularity of the original function, but the bound in the Lipschitz definition is tighter. In addition, for non negative functions, properties 4 and 5 imply $\rho(f,g) \leq \rho(f^*,g^*)$. As a result, if the monotonic rearrangements of two functions are poorly correlated, their not rearranged counterparts are likewise.

The natural extension of Definition 1 to the multidimensional case consists in cascading the 1D monotonic rearrangement along the x and y direction [5], according to the following definition.

Definition 2 Let μ be the Lebesgue measure, $f(x,y) : \Omega = \Omega_x \times \Omega_y \to \mathbb{R}_+^2$ be a μ-measurable and non negative function,

$$M_f(\sigma, y) = \mu(\{x \in \Omega_x : f(x,y) > \sigma\}), \quad \sigma \in \mathbf{R}_+$$

the distribution function of f in the first variable, and

$$M_f(x, \tau) = \mu(\{y \in \Omega_y : f(x,y) > \tau\}), \quad \tau \in \mathbf{R}_+.$$

Then

1. the decreasing rearrangement f_x^* of f in the first variable x is

$$f_x^*(s,y) = \inf\{\sigma \in \mathbf{R}_+ : M_f(\sigma, y) \leq s\}; \tag{12}$$

2. the decreasing rearrangement f_y^* of f in the second variable y is

$$f_y^*(x,t) = \inf\{\tau \in \mathbf{R}_+ : M_f(x, \tau) \leq t\}; \tag{13}$$

3. the 2D decreasing rearrangement f^* of f is

$$f^*(s,t) = [f_x^*(s,y)]_y^*(t). \tag{14}$$

Fig. 2 contains the rearranged version of three FF images that has been obtained by cascading the two 1D rearrangements, according to eq. (14). Two images have been acquired by the same device while the third one has a different source. As it can be observed, their monotonic rearrangements clearly show featuring geometrical patterns that allow source device discrimination nearly at the first glance. In fact, images acquired by the same device model (first two images in Fig. 2) exibit a similar pattern, while images captured by different device model show a different pattern (compare the first two images in Fig. 2 with the last one).

To support and prove this statement, we consider an equivalent definition of non increasing rearrangement of measurable functions that has been studied in [4] and that employs the 2D decreasing rearrangement of a subset of $\mathbf{R}^2$ and the layer-cake representation.

Definition 3 Let μ be the Lebesgue measure, $E \subset \mathbf{R}^2$

$$\varphi_E(x) = \mu\{y \in \mathbf{R} : (x, y) \in E\}, \quad x \in \mathbf{R}$$

and φ_E^* the conventional 1D decreasing rearrangement of $\varphi_E(x)$, i.e.

$$\varphi_E^*(s) = inf\{\lambda : M_\varphi(\lambda) \leq s\}, \quad s \in \mathbf{R}^+$$

with $M_\varphi(\lambda) = \mu\{x \in \mathbf{R} : \varphi_E(x) > \lambda\}$. Then, the two dimensional decreasing rearrangement of E is

$$E^* = \{(s, t) \in \mathbf{R}_+^2 : 0 < t < \varphi_E^*(s)\}, \tag{15}$$

and the two dimensional decreasing rearrangement of $f \in \mathbf{R}^2$ is

$$\hat{f^*}(s, t) = \int_0^{+\infty} \chi_{E^*}(s, t)dw, \quad (s, t) \in \mathbf{R}_+^2, \tag{16}$$

with $E = \{(x, y) \in \Omega \subset \mathbf{R}^2 : f > w\}$.

For measurable functions, the definition in eq. (16) is equivalent to the definition in eq. (14) [4].

In addition, it is worth observing that, for a fixed value w of the function f, the curve

$$\mathcal{C}_w : \{(s, t) : t = \varphi_{Ew}^*(s)\}, \tag{17}$$

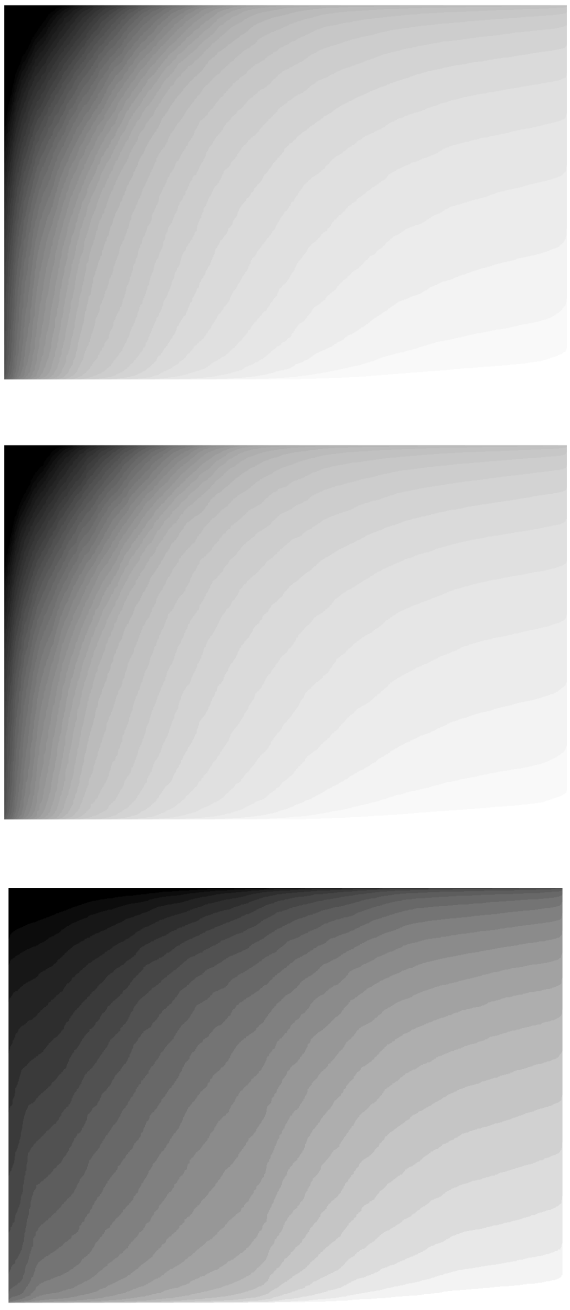

Figure 2: 2D monotonic rearranged versions J^* of FF images acquired by NikonD70 (*first two images*) and Canon Ixuss 55 (*bottommost image*).

with $E_w = \{(x,y) \in \mathbf{R}^2 : f > w\}$, is defined in the domain of f^* and separates this domain into two regions: the one where $f^* > w$ and the one $f^* \leq w$. Based on this observation, we can define the quantized version of f as it follows.

Definition 5 Let $w_k, \quad k = 1, ..., n_b$ be $n_b \in \mathbf{N}$ real and positive values such that $w_1 > w_2 > ....w_{n_b}$,

$$E_k = \{(x,y) \in \Omega \subset \mathbf{R}^2_+ : w_{k-1} \leq f_q(x,y) < \omega_k\}, \quad k = 1, ..., n_b$$

and $E_0 = \emptyset$. Let set $F_k = E_k \setminus E_{k-1}$. The quantized version of f is

$$f_q(x,y) = w_k, \quad \forall\, (x,y) \in F_k. \tag{18}$$

f_q well models an image acquired by a camera, including FF images. As a result, if f_q in eq. (18) represents the FF image J in eq. (8), then the geometrical pattern observed in J^* is composed of the curves

$$\mathcal{C}_{w_k} = \{(s,t) : t = \varphi^*_{E_k}(s)\}, \quad k = 1, ..., n_b. \tag{19}$$

.

In fact, according to Definition 5, J can be written as it follows

$$J(x,y) = \sum_{k=1}^{n_b} w_k\, \chi_{F_k}(x,y)$$

and then, using eq. (15),

$$J^*(s,t) = \sum_{k=1}^{n_b} w_k \chi_{E^*_k \setminus E^*_{k-1}}(s,t),$$

with

$$E^*_k \setminus E^*_{k-1} = \{(s,t) \in \mathbf{R}^2_+ : 0 < t < \varphi^*_{E_k}(s)\} \setminus \{(s,t) \in \mathbf{R}^2_+ : 0 < t < \varphi^*_{E_{k-1}}(s)\} =$$

$$= \{(s,t) \in \mathbf{R}^2_+ : \varphi^*_{E_{k-1}}(s) \leq t < \varphi^*_{E_k}(s)\},$$

i.e. each set $E^*_k \setminus E^*_{k-1}$ is delimited by the curves $\mathcal{C}_{w_{k-1}}$ and $\mathcal{C}_{w_k}$.

3.1. Features representation

The image feature defined in eq. (19) can be coded by applying a directional transform to J^*, as for example, the Radon Transform (RT) [17]. The latter is a non-linear transform that is widely used in tomography for the recovery of the internal structure of an object from some of its projections; however, it also is adopted in many applications as it concerns, in some sense, the geometrical structure of the object of interest.

Definition 6 Let $G(x, y) : \mathbf{R}^2 \rightarrow \mathbf{R}$ be a smooth function. The *Radon transform* of G at a point $(r, \theta) \in \mathbf{R}^2$ is defined as the integral of G along the line identified by the parameters r and θ, i.e.

$$\mathcal{R}_G(r, \theta) = \int_{\mathbf{R}} G(r\mathbf{n}_\theta + s\mathbf{n}_\theta^\perp)ds \tag{20}$$

where $\mathbf{n}_\theta = (cos(\theta), sin(\theta))$ and $\mathbf{n}_\theta^\perp = (-sin(\theta), cos(\theta))$.

Due to its nature, this transform is able to capture the main orientations of objects in the images. That is why it represents a good candidate for capturing those geometrical structures that characterize the behaviour of the monotonically rearranged version of device PRNU. Specifically, the RT values quantify how many points in the image lay on the straight line having slope equal to $tan(\theta)$ and intercept equal to r.

An equivalent definition of the transform allows us to rewrite $\mathcal{R}_G$ as it follows

$$\mathcal{R}_G(r, \theta) = \int_{\mathbf{R}} G(x, r)\delta(r - xcos(\theta) - xsin(\theta))dxdy. \tag{21}$$

It is straighforward to observe that a point in the function domain generates a sinusoid in the Radon domain, while collinear points intersect in the point whose coordinates define the straight line they lay on. As a result, a curve in the image domain is expected to distribute differently in the Radon domain, according to its main orientations. It turns out that the comparison between the RTs of the *image PRNU* and the *reference PRNU* is expected to be informative about their similarity.

However, the dimension of the discrete Radon Tranform still is dependent on the image size (see Figs. 3-5) — this would entail the need of having *image PRNU* and *reference PRNU* with the same size, limiting the use of this feature on a restricted real world cases. In order to allow for image

comparisons independently of their dimension and avoiding additional modifications, as sampling or interpolation, in this paper the empirical probability of the quantized values of the transform has been considered, whenever the number of bins b used for quantization is a priori fixed, i.e. $P_{\mathcal{R}_{J^*}}$. Even though it represents a global RT descriptor, the empirical probability of $\mathcal{R}_{J^*}$ conveys information concerning the more or less oriented image content. In addition, it is twofold advantageous. On the one hand, it allows for direct comparisons between images having different size; on the other hand, it associates a monodimensional signal to the image, resulting computationally advantageous in the classification process. Section 4 is devoted to quantitative evaluate the ability of the proposed feature in assigning an image to the correct device, even under global image manipulations, as filtering, quantization, resizing.

Remark Although other geometrical transforms could be adopted for coding the geometrical pattern used as feature in this paper, the use of the Radon transform has been mainly motivated by the parameter-free definition. The dependence of the results on the selected transform or the definition of the best transform to use are out of the scope of the presented paper, whose aim was to investigate about the strength of 2D rearrangements in featuring source noise component; however, they can be the subject of future research.

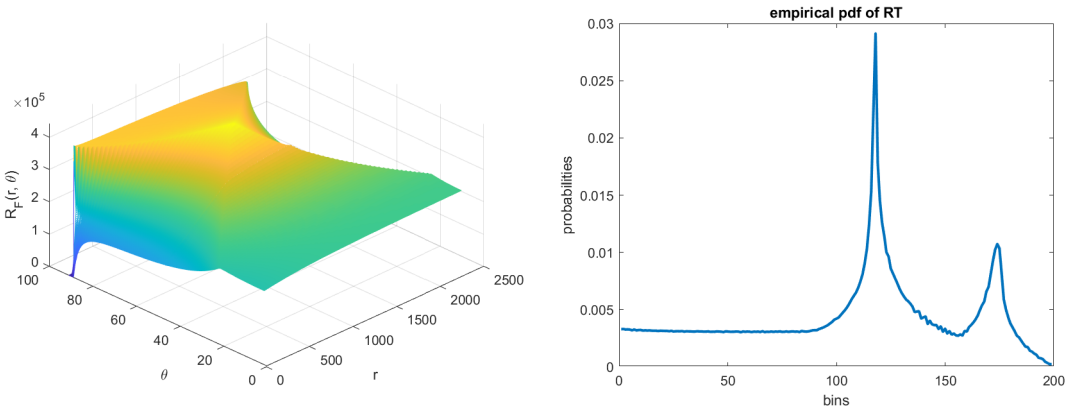

Figure 3: *Left*) Radon transform $\mathcal{R}_{J^*}$ of the rearranged image in Fig, 2.top; (*Right*) empirical probability of $\mathcal{R}_{J^*}$ values.

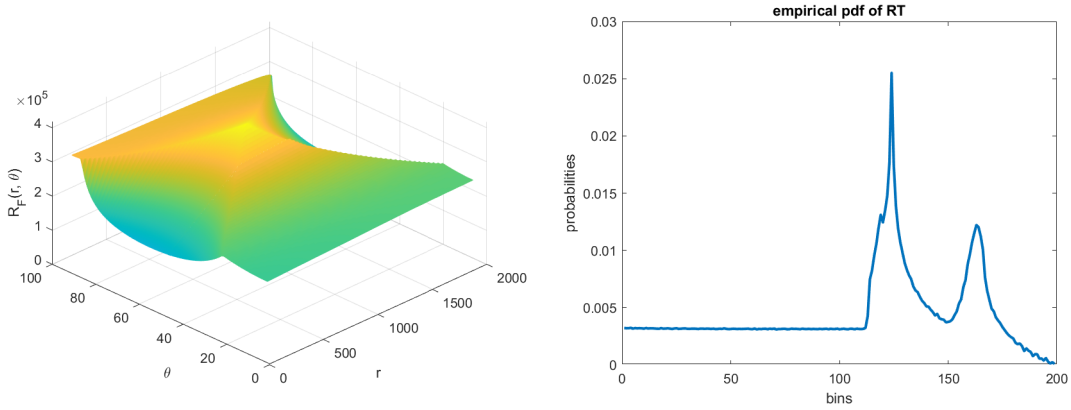

Figure 4: *Left*) Radon transform $\mathcal{R}_{J^*}$ of the rearranged image in Fig, 2.bottom; (*Right*) empirical probability of $\mathcal{R}_{J^*}$ values.

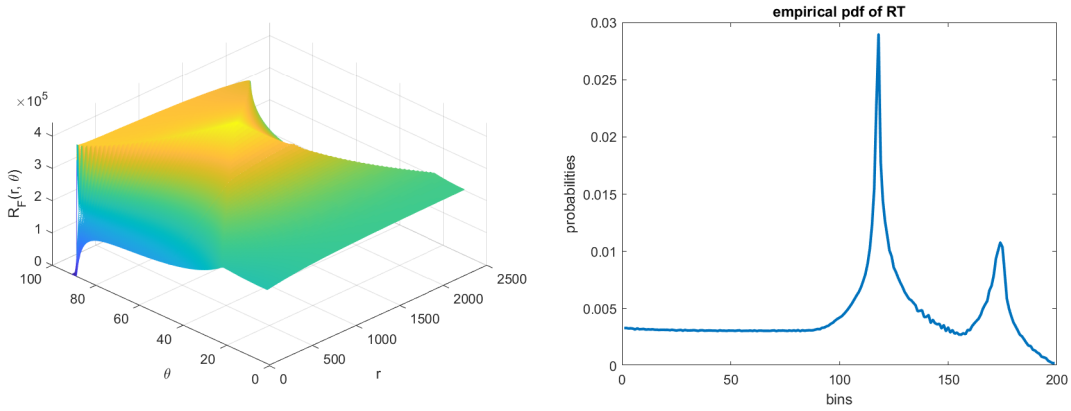

Figure 5: *Left*) Radon transform $\mathcal{R}_{J^*}$ of the rearranged image in Fig, 2.middle; (*Right*) empirical probability of $\mathcal{R}_{J^*}$ values.

3.2. The Algorithm: closed set scenario

In this section, the whole source recognition process in closed-set scenario is summarized. As we are dealing with FF images, the *image PRNU* is the image J as defined in eq. (8).

Algorithm for closed-set scenario

1. **Feature for *image PRNU***
   - Compute J^* as in eq. (16);
   - Compute the Radon Transform of J^*, i.e. $\mathcal{R}_{J^*}$;
   - Quantize $\mathcal{R}_{J^*}$ using N_B quantization bins and compute its empirical probability, i.e. $P_{\mathcal{R}_{J^*}}$;
   - Define the feature for the *image PRNU* as $P_{\mathcal{R}_{J^*}}$.
2. **Feature for *reference PRNU***
   - For each device $d_j, \quad j = 1, ..., M$
     - For each image J_{i,d_j} captured by the device d_j:
       * Compute J^*_{i,d_j} as in eq. (16);
       * Compute the Radon Transform of J^*_{i,d_j}, i.e. $\mathcal{R}_{J^*_{i,d_j}}$;
       * Quantize $\mathcal{R}_{J^*_{i,d_j}}$ using N_B quantization bins and compute its empirical probability, i.e. $P_{\mathcal{R}_{J^*_{i,d_j}}}$.
     - Define the feature for the *reference PRNU* as it follows

$$\bar{P}_{d_j} = \sum_{i=1}^{N_j} P_{\mathcal{R}_{J^*_{i,d_j}}} \tag{22}$$

3. **Source device identification**
   - Assign the image J to the device $d_{\tilde{j}}$ defined as it follows

$$\tilde{j} = argmin_{d_j} D(P_{\mathcal{R}_{J^*}}, \bar{P}_{d_j}),$$

where D is the selected similarity metric.

Table 1: Selected images and devices from Dresden database [25].

n	Device Model	no FF images	image size
1	Canon IXUS55	10	1944×2592
2	Canon IXUS70	50	2304×3072
3	Nikon D200	10	2592×3872
4	Nikon D70	10	2000×3008
5	Nikon D70S	10	2000×3008
6	Nikon S710	10	3264×4352
7	Olympus 1050SW	50	3648×2736
8	Panasonic DMC-FZ50	25	3648×2736
9	Pentax Optio A40	50	3648×2736
10	Ricoh Capilo GX100	25	3648×2736
11	Rollei RCP-7325XS	50	3072×2304
12	Samsung NV15	50	3648×2736
13	Samsung L74	50	2304×3072
14	Sony DSC-T77	50	3648×2736
15	Sony DSC-W170	50	3648×2736

4. Experimental results

This section is devoted to evaluate the accuracy of the proposed feature in discriminating images acquired by different devices. FF images from a selected subset of devices included in Dresden database [25] have been considered in all tests. This dataset includes hundred of images captured by several camera models and devices. It has been selected as if offers a considerable number of FF images. In order to test the robustness of the proposed feature, different image sizes have been considered as well as a different number of images per device. In particular, 15 devices and 500 FF images, listed in Table 1, have been considered in the tests presented in this section.

As we are in the closed-set scenario, a LOSO (Leave One Subject Out) validation approach has been used and the true positive rate (specificity) has been evaluated, i.e. $TPR = \frac{TP}{TP+FN}$, where TP and FN respectively denote correct assignments and failures. Three different metrics [42] for evaluating the distance between the empirical probability density functions have been adopted, i.e.

i) the *Euclidean distance*, i.e. $D_E(p,q) = \sqrt{\sum_{i=1}^{N_B} |p_i - q_i|^2}$;

ii) the *Hellinger distance*, i.e. $D_H(p,q) = \frac{1}{\sqrt{2}}\sqrt{\sum_{i=1}^{N_B}(\sqrt{p_i} - \sqrt{q_i})^2}$, and

iii) the *Triangular distance*, i.e. $D_T(p,q) = \sqrt{\sum_{i=1}^{N_B} \frac{|p_i - q_i|^2}{p_i + q_i}}$,

where p and q denote the two comparing empirical pdfs and N_B the number of bins. N_B has been set equal to 200 in all tests. In addition, in our tests $p = P_{\mathcal{R}_{J^*}}$ and $q = \bar{P}_{d_j}$. Each color channel is processed independently; the final similarity index is then defined as euclidean norm of the vector whose components are the distances evaluated for each color channel.

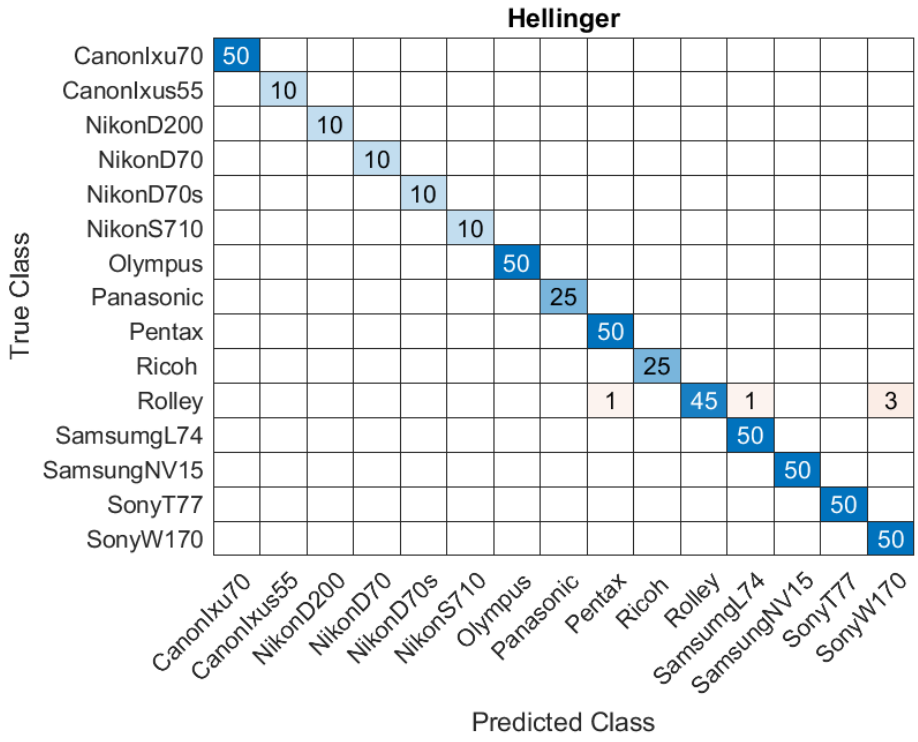

Figure 6: Confusion matrix of the classification results obtained using the Hellinger distance. False assignments do not belong to the diagonal. The True Positive Rate (TPR) is 99.00%.

Figs. 6, 7 and 8 contain the confusion matrices relative to the classification results obtained using the three selected metrics. As it can be observed, for all but one device all images have been assigned correctly. In addition, the Hellinger distance shows the best true positive rate as only 5 images over

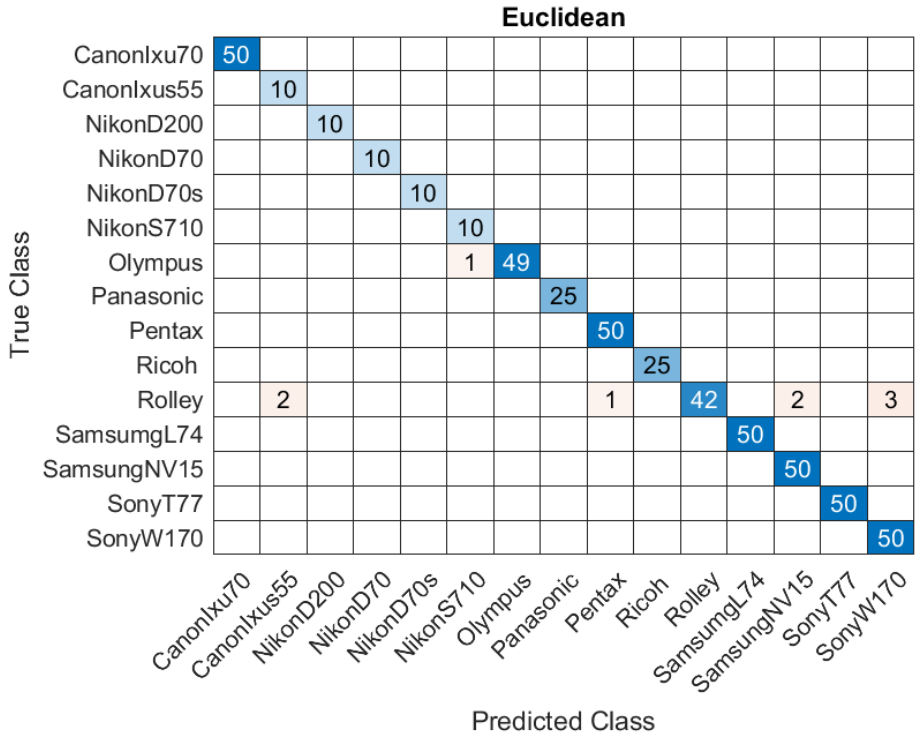

Figure 7: Confusion matrix of the classification results obtained using the Euclidean distance. False assignments do not belong to the diagonal. The True Positive Rate (TPR) is 98.20%.

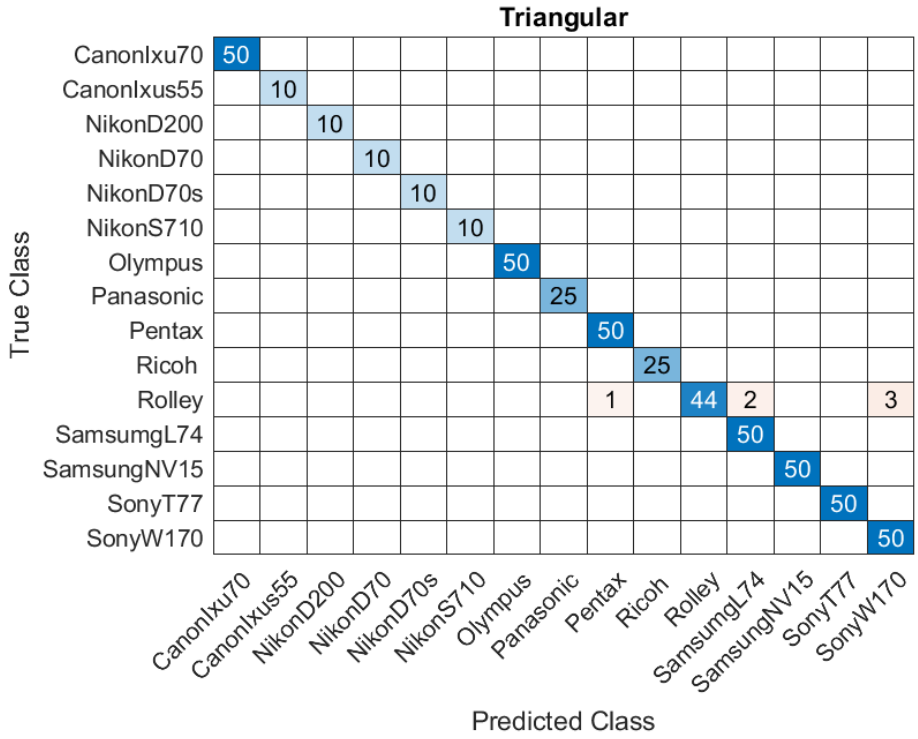

Figure 8: Confusion matrix of the classification results obtained using the Triangular distance. False assignments do not belong to the diagonal. The True Positive Rate (TPR) is 98.80%.

Table 2: True Positive Rate on different testing sets: comparisons between the proposed method, and the method in [30] where resizing is used for images having different size. The Table refers to the whole test dataset, and two of its subsets: Set I contains only images having the same size (3648×2736), while Set II contains the remaining ones. TPR refers to the Hellinger distance.

	Proposed	Method in [30]
all SET	**99.0**	96.4
SET I	**100**	**100**
SET II	**98.3**	92.9

500 have not been assigned to the corresponding source. It is also worth observing that the minimum distance value is measured independently of the number of FF images used for estimating the feature of the reference PRNU. Table 2 compares the proposed Radon-based feature method with the pioneering and reference approach in [30]. The latter uses an MLE strategy for estimating the reference PRNU and the normalized correlation coefficient ρ for assigning the source to each image, as described in Section 2 where $\mathcal{D}$ is a wavelet-based denoiser. As ρ is defined for images having the same size, resizing is used whenever images having different size have to be compared. Table 2 clearly gives evidence of the main advantage of the proposed method: it allows us to compare the *image PRNU* with any *reference PRNU*, independently of the image size while guaranteeing high recognition rates. In particular, whenever a testing set composed of images having the same size is used (SET I in the Table 2) the proposed approach has the same recognition rates of the competing method; on the contrary, for testing sets composed of images having different size (allSET and SET II in the Table 2), the proposed approach considerably outperforms the competing method that is penalized by the resizing operation. In addition, the proposed feature-based approach significantly reduces storage sources as a very short feature vector is required instead of the whole reference PRNU image.

To give evidence of the characterization of each single device, one FF image has been considered as *reference PRNU* and the LOSO analysis has been performed again. The boxplot plot of the assignments (evaluated using the Triangular distance) obtained by comparing the image of interest with the remaining test images is depicted in Fig. 9. As it can be observed, boxes do not overlap and the classes are clearly separated. This result is important as it makes the proposed feature useful whenever very few images

are available for each device, resulting in limited data for the estimation of the reference PRNU.

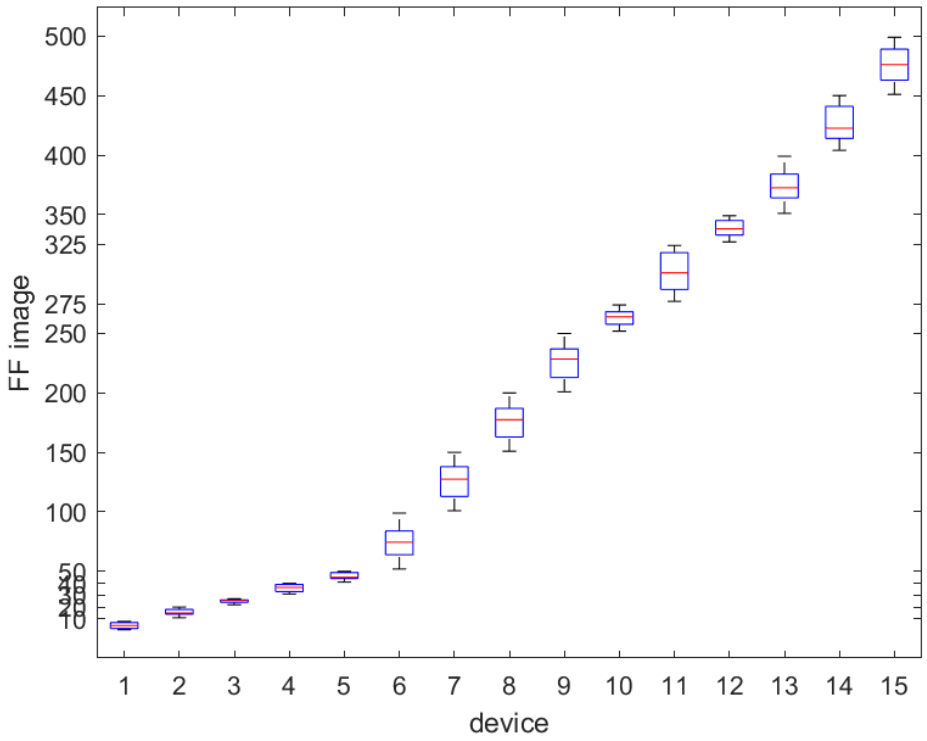

Figure 9: Boxplot of Triangular distance values obtained in the closed set scenario and using a single FF image as reference PRNU.

Finally, in order to evaluate the robustness of the proposed feature to image manipulations, test images have been modified using some standard operations involving all image pixels, such as sampling (S), interpolation (I), quantization (Q), blurring (B). They represent some standard and commonly adopted image manipulations, as for example when uploading or downloading images from social networks, when sending an image by email or when storing the image on a storage device. Table 3 contains the achieved true positive rates corresponding to sampling steps equal to 2, cubic 2D interpolation, blurring with a gaussian window having standard deviation equal to 2, quantization bin size equal to 10. As it can be observed, the number of false assignments slightly decreases for some operations like quantization, while it

increases in case of blurring. True positive rate decreases more whenever manipulations are jointly used, although still remaining greater than 92%. Fig. 10 shows the confusion matrix for the case blurring combined with quantization. As it can be observed, for most of the devices the assignments remain correct for all images; incorrect assignments mainly refer to few specific devices, as for example Olympus, Panasonic and Ricoh. It is worth observing that the result is not negligible if one considers that just a simple and global feature is employed. This analysis motivates future investigations. In fact, it proves that a simple linear filter applied to FF images does not alter the features that are captured by the RT of the sorted FF image. In addition, even though the Hellinger distance seems to provide the best recognition results on average, the Triangular one seems to be a little bit more robust to more complex manipulations. The role of the similarity metric will be one of the topic worth of attention in the future studies.

Table 3: True Positive Rate achieved by the proposed method using Euclidean, Hellinger and Triangular distances. Unaltered (U) and manipulated images using conventional operations have been considered: sampling (S), interpolation (I), blurring using gaussian filtering (B), quantization (Q), consecutive blurring and quantization (B+Q). For each column, the best result is in bold.

	U	S	I	B	Q	B + Q
TPR (D_E)	98.20	98.40	98.40	98.80	97.80	**93.00**
TPR (D_H)	**99.00**	**99.00**	**99.00**	**99.50**	**98.20**	91.20
TPR (D_T)	98.80	98.80	98.80	99.20	97.80	92.80

5. Conclusions

In this paper a study concerning some properties of PRNU components in an image have been investigated. Based on the relations between functions and their rearrangements, it has been shown that the Radon Transform of a proper rearrangement of PRNU matrix preserves some peculiarities of the original data. Preliminary tests prove that the empirical probability of the values of that transform is able to assign each FF image to its source device with high accuracy. In particular, the proposed feature offers different advantages: *i)* it allows for the direct comparison between PRNU components

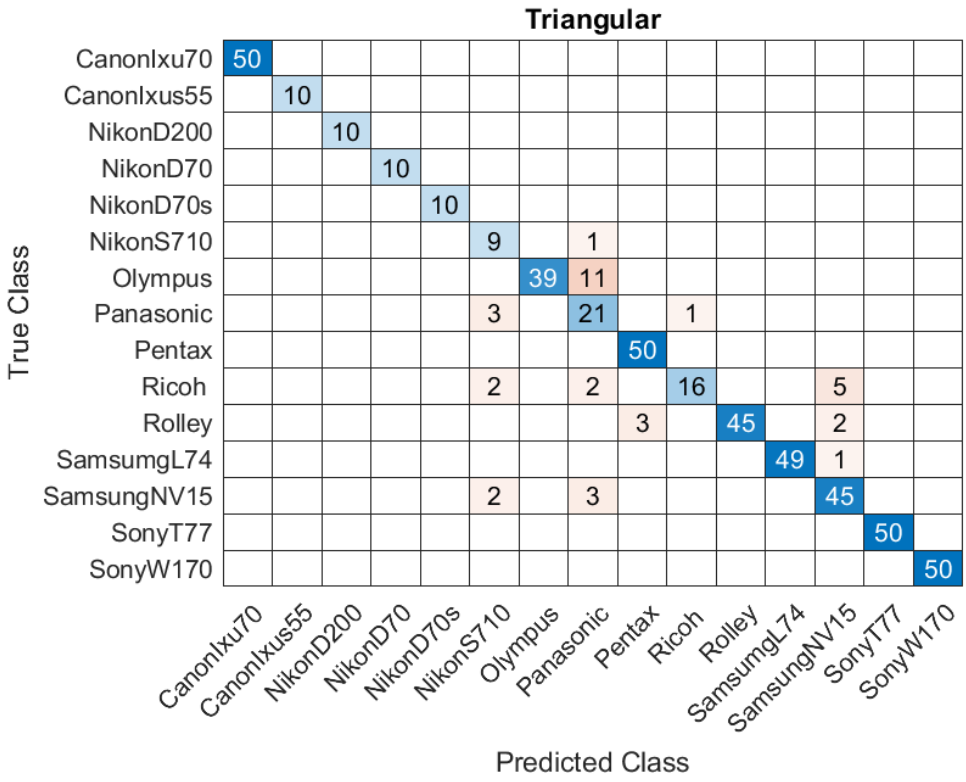

Figure 10: Confusion matrix of the classification results obtained using the Triangular distance on FF images that have been blurred and then quantized. The True Positive Rate (TPR) is 92.80%.

having different size; *ii)* it is a feature for each single image acquired by the device, as a result it can contribute in providing reliable responses even if very few images are available for the estimation of the reference PRNU; *iii)* it is robust to some standard image manipulations. Based on the promising preliminary results, future research will be devoted to the extraction of that feature from natural images by combining the theoretical properties of both the decreasing rearrangements and the Radon transform. In addition, features other than RT empirical probability will be investigated in order to further improve recognition accuracy.

6. Acknowledgments

This research was partially funded by the Italian national research group GNCS (INdAM). This research has been accomplished within RITA (Research ITalian network on Approximation).

References

[1] K.R. Akshatha, A.K. Karunakar, H. Anitha, U. Raghavendra, and D. Shetty, "Digital camera identification using PRNU: A feature based approach", Digital Investigation, 19 (2016), 69–77.

[2] M. Al-Ani, F. Khelifi, A. Lawgaly, 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) 1–5.

[3] M. Al-Ani and F. Khelifi, "On the SPN Estimation in Image Forensics: A Systematic Empirical Evaluation", IEEE Trans. on Inf. Forensics and Security, 12(5), (2017), 1067–1081.

[4] S. Barza, L.E. Persson, J. Soria, Multidimensional rearrangements and Lorentz spaces, Acta Mathenatica Hungarica, 104 (3), (2004) 203-224.

[5] A. P. Blozinski, Multivariate rearrangements and Banach function spaces with mixed norms, Transactions of the American Mathematical Society, 263(1), (1991), 149-167.

[6] V. Bruni, D. De Canditiis, D. Vitulano, Local Sorting for Adaptive Signal Regularization, IEEE Signal Processing Letters 17(7), (2010)

[7] V. Bruni and D. Vitulano, Signal and image denoising without regularization, IEEE International Conference on Image Processing, 539–542 (2013)

[8] V. Bruni, A. Salvi and D. Vitulano, Joint correlation measurements for PRNU-based source identification, Proc. of the 18th Int. Conf. on Computer Analysis of Images and Patterns, CAIP 2019, (2019).

[9] V. Bruni, S. Marconi, D. Vitulano, A novel Fourier based approach for camera identification, in Proceedings of the International Conference on Image Processing and Vision Engineering (IMPROVE 2021), (2021), 99–106.

[10] V. Bruni, M. Tartaglione and D. Vitulano, "Coherence of PRNU weighted estimations for improved source camera identification", Multimedia Tools and Applications, (2022).

[11] R. Caldelli, T. Amerini and C. Tsun Li, "PRNU-based Image Classification of Origin Social Network with CNN", Proc. of EUSIPCO (2018).

[12] A. Bruno, P. Capasso, G. Cattaneo, et al. A novel image dataset for source camera identification and image based recognition systems. Multimedia Tools Applications, (2022).

[13] M. Chen, J. Fridrich, M. Goljan and J. Lukas, "Determining Image Origin and Integrity Using Sensor Noise", IEEE Trans. on Inf. Forensics and Security, 3(1), (2008), 74–90.

[14] G. Chierchia, G. Poggi, C. Sansone, L. Verdoliva "A Bayesian-MRF Approach for PRNU-Based Image Forgery Detection", IEEE Trans. on Information Forensics and Security, 9(4), (2014), 554–567.

[15] D. Cozzolino, L. Verdoliva, Noiseprint: A CNN-Based Camera Model Fingerprint, IEEE Transactions on Information Forensics and Security, 15, (2020) 144-159.

[16] D. Cozzolino, F. Marra, D. Gragnaniello, G. Poggi, L. Verdoliva, Combining PRNU and noiseprint for robust and efficient device source identification, EURASIP Journal on Information Security, (2020).

[17] S.R. Deans, "The Radon transform and some of its applications". Courier Corporation, (2007).

[18] X. Ding, Y. Chen, Z. Tang, Y. Huang, Camera Identification based on Domain Knowledge-driven Deep Multi-task Learning, IEEE Access, 7, (2019) 25878–25890.

[19] G. F. D.Duff, "Differences, derivatives, and decreasing rearrangements", Canadian Journal of Mathematics, 19, (1967), 1153-1178.

[20] J. B. Epperson, A Class of Monotone Decreasing Rearrangements, Journal of Mathematical Analysis and Applications, 150, (1990), 224-236.

[21] P. Ferreira, Sorting continuous-time signals: analog median and median-type filters, IEEE Transactions on Signal Processing, 49(11), (2001), 2734-2744.

[22] J. Fridrich, " Digital image forensics", IEEE Signal Processing Magazine, 26(2), (2009), 26–37.

[23] C Galdi, F Hartung, J-L Dugelay, SOCRatES: a database of realistic data for SOurce Camera REcognition on Smartphones, Proc. of International Conference on Pattern Recognition Applications and Methods (ICPRAM 2019), (2019)

[24] S. Georgievska, R. Bakhshi, A. Gavai, A. Sclocco, B. van Werkhoven, Clustering image noise patterns by embedding and visualization for common source camera detection, Digital Investigation, 23, (2017), 22-30

[25] T. Gloe and R. Bhme, "The Dresden image database for benchmarking digital image forensics", J. of Dig. For. Prac., 3, (2010), 150–159.

[26] B. Gupta, M. Tiwari, Improving source camera identification performance using dct based image frequency components dependent sensor pattern noise extraction method, Digital Investigation, 24, (2018), 121–127.

[27] Y. Huang, J. Zhang and H., Huang "Camera Model Identification With Unknown Models", IEEE Trans. on Inf. Forernics and Sec., 10(12), (2015), 2692–2704.

[28] X. Kang, J. Chen, K. Lin, A. Peng, A context-adaptive SPN predictor for trustworthy source camera identification, EURASIP Journal on Image and Video Processing, 1, (2014), 19-30.

[29] Y. Liu, Z. Zou, Y. Yang, N.-Law, A.A. Bharath, Efficient source camera identification with diversity-enhanced patch selection and deep residual prediction, Sensors 21(14), (2021).

[30] J. Lukas, J. Fridrich and M., Goljan, "Digital camera identification from sensor pattern noise", IEEE Trans. on Inf. Forensics and Sec., 1, (2006), 205–214.

[31] F. Marra, G. Poggi, C. Sansone and L. Verdoliva, "Blind PRNU-Based Image Clustering for Source Identification", IEEE Trans. on Inf. For. and Sec., 12(9), (2017), 2197–2211.

[32] S.F. Menduina, F. P. Gonzalez, M. Masciopinto, Source camera attribution via PRNU emphasis: Towards a generalized multiplicative model, Signal Processing: Image Communication, in press, (2023)

[33] D. Freire-Obregón, F. Narducci, S. Barra, M. Castrillón-Santana, Deep learning for source camera identification on mobile devices, Pattern Recognition Letters, 126, (2019), 86–91.

[34] A. Papoulis, S. Pillai, Probability, Random Variables, and Stochastic Processes, McGraw Hill, Boston, (2002).

[35] Q. Rao, J. Wang, L. Zhang, Enhancing Source Camera Identification Based on Multiplicative Denoising Filter, IEEE Trustcom/BigDataSE/ISPA, Tianjin, (2016), 983-988.

[36] J.A. Redi, W. Taktak and J. Dugelay, "Digital image forensics: a booklet for beginners", Multimedia Tools and Applications 51, (2011), 133–162.

[37] L. Ruizhe, L. Chang-Tsun, G. Yu, Inference of a compact representation of sensor fingerprint for source camera identification, Pattern Recognition, 74, (2018), 556-567.

[38] D. Shullani, M. Fontani, M. Iuliani, O. Al Shaya and A. Piva, "VISION: A video and image dataset for source identification", EURASIP J. Inf. Secur., vol. 2017(1), (2017).

[39] T. H. Thaia, F. Retraintband, R. Cogranne, Camera Model Identification Based on the Generalized Noise Model in Natural Images, Digital Signal Processing, 00, (2015), 1-15.

[40] M. Tiwari, B. Gupta, Image features dependant correlation-weighting function for efficient PRNU based source camera identification, Forensic Science International, 285, (2018), 111–120.

[41] M.Tiwari, B. Gupta, Efficient PRNU extraction using joint edge-preserving filtering for source camera identification and verification, Proc. of IEEE Applied Signal Processing Conference (ASPCON) (2018)

[42] F. Topsoe, "Some Inequalities for Information Divergence and Related Measures of Discrimination", IEEE Trans. on Information Theory, 46(4), (2000), 1602–1609.

[43] A. Tuama, F. Comby, M. Chaumont, Camera model identification with the use of deep convolutional neural networks, Proc. of the 2016 IEEE International workshop on information forensics and security WIFS, (2016), 1-6.

[44] P. Yang, R. Ni, Y. Zhao, W. Zhao, Source camera identification based on content-adaptive fusion residual networks, Pattern Recognition Letters, 119, (2019), 195–204.

[45] P. Yang, D. Baracchi, R. Ni, Y. Zhao, F. Argenti, A. Piva, A Survey of Deep Learning-Based Source Image Forensics, Journal of Imaging, 6(9), (2020),

[46] H. Yao, T. Qiao, M. Xu, N. Zheng, Robust multi-classifier for camera model identification based on convolution neural network, IEEE Access, 6, (2018), 24973–24982.

[47] Y. Zhao, N. Zheng, T., Qiao and M. Xu, M. "Source camera identification via low dimensional PRNU features", Multimedia Tools and Applications, 78, (2019), 8247–8269.