Design of Projection Matrices for PRNU Compression

Luca Bondi[1], Fernando Pérez-González[2], Paolo Bestagini[1] and Stefano Tubaro[1]

[1]Dipartimento di Informazione, Elettronica e Bioingegneria, Politecnico di Milano, Italy
[2]Signal Theory and Communications Department, University of Vigo E. E. Telecomunicación, Spain

***Abstract*—Photo Response Non-Uniformity (PRNU) is the de-facto standard in image source identification, allowing scientists, researchers, forensics investigators and courts to bind a picture under investigation to the specific camera sensor that took the shot at first place. Caused by silicon sensor imperfections, PRNU is characterized as a Gaussian i.i.d weak multiplicative noise embedded into every digital photo at acquisition time. Despite PRNU nearly-flat spectral characteristics, it undergoes several interpolations steps while image is demosaicked and optionally JPEG compressed. In this paper we propose a novel approach to the design of projection matrices tailored to PRNU compression. Joint effect of interpolation and projection on cross-correlation test is first analyzed, in order to derive those conditions that maximize detection while reducing false-alarm probability. A design methodology to build effective projection matrices is then presented, taking into account computational complexity. Validation of the proposed approach is finally performed against state-of-the-art methods on a well known public image dataset.**

I. Introduction

Over the last decade, ownership attribution and origin verification of digital content has become of capital importance, due to the widespread diffusion of digital devices capable of acquiring images, videos and audio tracks, as well as sharing them over the Internet. The most widespread technique for camera device identification is based on the Photo Response Non-Uniformity (PRNU) [1]. This is a time invariant weak multiplicative signal introduced on every picture taken with a CCD/CMOS imaging device that acts as unique fingerprint for the sensor itself. Silicon imperfections occurring at sensor manufacturing process cause each pixel to have a slightly different area, thus the amount of light energy captured in a fixed time slot (i.e. exposure time) varies pixel-wise even under a perfectly uniform light field. The PRNU can be extracted, or at least estimated, by having access to a set of images captured by the same camera. PRNU knowledge allows to determine whether two images have been captured by the same device, link a picture to the specific camera that took it, and even detect forgeries [2]–[4].

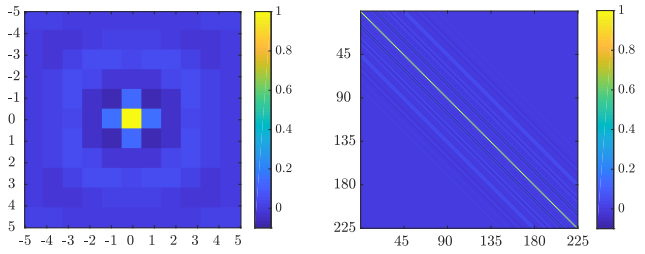
The PRNU fingerprint extraction process from natural and flatfield images has been thoroughly investigated [5]–[8] with the goal of extracting the fingerprint as accurately as possible. The overall goal was to improve matching and detection performance between fingerprints extracted from different images coming from the same device, while keeping a very low probability that estimated fingerprints coming from different sensors are matched.

PRNU robustness to scaling, cropping and compression [9], [10] is an important prerequisite to move toward large scale applications and scenarios [11], even when camera devices are not known in advance, but only image fingerprints are available [12]. One drawback of large scale approaches is the need to store in a central database, or transmit over bandwidth-limited channels, a huge amount of data. Indeed, PRNU fingerprints need to be extracted at higher resolutions, up to the size of the imaging sensor, to achieve better matching and detection performance and avoid false-alarms. A second issue arises in terms of computational complexity, when a query fingerprint needs to be matched against many device fingerprints stored in a central camera fingerprints database.

To overcome both rate and complexity issues, several PRNU compression methods have been recently proposed. The digest technique [13], aimed at preserving only the most relevant peaks from a PRNU fingerprint, proves to be really effective, even though it is bound to the knowledge of an affordable fingerprint estimate to preserve good matching performance. Under this condition, the fingerprint digest provides benefits both in terms of data and complexity reduction, because only a few peaks values and positions need to be preserved, thus reducing space requirements and number of operations to compute the matching between fingerprints. The main shortcoming of this technique in a server-client scenario comes from the need, at the client side, of knowing the peak positions. These are almost impossible to estimate reliably at the client side, so in order for the digest to work in this scenario, the server should send peak positions for every known camera device to each client, thus making this compression approach

This material is based on research sponsored by DARPA and Air Force Research Laboratory (AFRL) under agreement number FA8750-16-2-0173. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of DARPA and Air Force Research Laboratory (AFRL) or the U.S. Government. Work of F. Pérez-González is partially funded by GPSC, is funded by the Agencia Estatal de Investigacin (Spain) and the European Regional Development Fund (ERDF) under project WINTER (TEC2016-76409-C2-2-R). Also funded by the Xunta de Galicia and the European Union (European Regional Development Fund - ERDF) under projects Agrupacin Estratéxica Consolidada de Galicia accreditation 2016-2019 and Red Temática RedTEIC 2017-2018.

WIFS'2017, December, 4-7, 2017, Rennes, France.
978-1-5090-6769-5/17/$31.00



(a) (b)

Fig. 1: Estimated camera PRNU $\hat{\mathbf{K}}_c$ autocorrelation matrix (a) and the derived $\mathbf{H}_c$ approximation (b) as space-invariant bi-dimensional linear filter with $h = 11$ in a small-scale example with $N = 225$.

practically unfeasible. Another simple, yet effective, technique resorts to fingerprint binarization, allowing to save rate while moving from a cross-correlation test to an Hamming distance computation when it comes to comparing fingerprints. Fast search methods based on sub-linear hashing on reference fingerprint digest [14], [15] and minimization of number of observations required to reduce error probabilities below some pre-fixed misdetection rates [16] are again two methods aimed at reducing fingerprint identification and verification cost. In terms of data compression, the state-of-the-art for PRNU signals is achieved by binarized Gaussian Random Projections [17], [18], that exploit Johnson-Lindenstrauss Lemma to project fingerprints onto a reduced dimensionality subspace by means of Gaussian i.i.d. projection vectors.

In this paper we address the problem of PRNU compression by proposing a projection design methodology that takes into account interpolation effects on the PRNU. A Signal-to-Noise-Ratio (SNR) maximization problem for the alternative hypothesis of a fingerprint cross-correlation test is solved, taking into account the interpolation effect due to the post-acquisition operations a digital image undergoes, e.g. demosaicing, JPEG compression. We first provide a theoretical framework to establish near-optimal conditions for the projection matrix, then we provide a design methodology for such matrix. Finally, we validate the proposed approach on images from the Dresden Image Database [19] in comparison with state-of-the-art Gaussian Random Projections. Our approach not only yields better detection performance as compared to the state-of-the-art, but is much cheaper in computational terms, which is a critical indicator when huge databases need to be searched.

The rest of this work is organized as follows. Sect. II provides notation, background and problem formulation. Sect. III presents our theoretical analysis and gives near-optimal conditions for the projection matrix. Sect. IV offers a design methodology to satisfy those conditions. Sect. V and Sect. VI respectively introduce the experimental setup and results, while in Sect. VII conclusions are drawn.

II. Background and Problem Formulation

We first introduce the used notation, the necessary background about PRNU fingerprint estimation and matching, and then we formalize the problem under analysis.

Notation: Vectors are given by boldface letters, e.g., $\mathbf{x}$. The ith sample of $\mathbf{x}$ is represented by x_i. The Hadamard (sample-wise) product between $\mathbf{x}$ and $\mathbf{y}$ is denoted by $\mathbf{x} \circ \mathbf{y}$. Matrices are denoted by bold capital letters, e.g., $\mathbf{X}$, and the i, jth element is indicated by a subindex, e.g., $X_{i,j}$. $\mathbf{I}$, $\mathbf{0}$ and $\mathbf{1}$ denote the identity matrix, the all-zeros matrix, and the all-ones vector, respectively. The Kronecker product between two matrices $\mathbf{X}$ and $\mathbf{Y}$ is denoted by $\mathbf{X} \otimes \mathbf{Y}$.

PRNU estimation is based on a simplified linear model of the camera sensor output $\mathbf{y} = g^\gamma \left[(\mathbf{1} + \mathbf{k}) \circ \mathbf{f} + \mathbf{z}\right]^\gamma + \mathbf{n}_q$, where $\mathbf{y}$ is the one-dimensional representation of the acquired image, $\mathbf{f}$ is light field at the sensor, g is the color channel gain and γ is the gamma correction factor [3]. $\mathbf{k}$ is a zero-mean noise-like signal identified as the PRNU fingerprint, $\mathbf{z}$ is a combination of remaining noise sources (dark currents, read-out noise, shot noise), and $\mathbf{n}_q$ is quantization noise. The imaging model can be further simplified as $\mathbf{y} = \mathbf{x} + \mathbf{x} \circ \mathbf{k}_0 + \mathbf{n}$, where $\mathbf{x} = (g\mathbf{f})^\gamma$ is a noise-free version of $\mathbf{y}$, $\mathbf{k}_0 = \gamma \mathbf{k}$ and $\mathbf{n}$ condensates all the independent random noise components residuals. Here we consider, without loss of generality, the wavelet decomposition based denoising algorithm proposed in [20] and adopted in [2] to get $\mathbf{x}$. When a single query image $\mathbf{y}_q$ is available, we define its residual as $\mathbf{w}_q = \mathbf{y}_q - \mathbf{x}_q$. When a set of N_c images from the same device is available, we define the residual for each picture $\mathbf{y}_c$, $c \in [1, N_c]$ as $\mathbf{w}_c = \mathbf{y}_c - \mathbf{x}_c$ and we obtain the estimated PRNU fingerprint $\hat{\mathbf{k}}$ for the device as

$$\hat{\mathbf{k}} = \frac{\sum_{c=1}^{N_c} \mathbf{w}_c \circ \mathbf{y}_c}{\sum_{c=1}^{N_c} \mathbf{y}_c^2} \tag{1}$$

Wiener adaptive filtering in the Discrete Fourier Transform domain and mean subtraction are finally applied to $\hat{\mathbf{k}}$ in order to remove model or compression specific artifacts.

When a query image under investigation $\mathbf{y}_q$ and a camera device c, characterized by a PRNU fingerprint $\mathbf{k}_c$ are made available, a two-channel hypothesis testing problem is faced in order to determine whether query $\mathbf{y}_q$ has been shot by the device characterized by $\mathbf{k}_c$:

$$\begin{aligned} H_0 &: \mathbf{y}_q \text{ was not taken with camera } c; \\ &\quad \text{thus it does not contain } \mathbf{k}_c \\ H_1 &: \mathbf{y}_q \text{ was taken with camera } c; \\ &\quad \text{thus it contains } \mathbf{k}_c \end{aligned}$$

Detection of such matching is performed via cross-correlation test, defined as $r \doteq \langle \hat{\mathbf{k}}_c, \mathbf{w}_q \rangle$. When $r > \tau$ then H_1 hypothesis is verified and query image is verified to come from c. τ is a threshold on the cross-correlation test properly set in order to bound the false-alarm probability under a maximum desired value.

Problem formulation: Given a device whose estimated PRNU fingerprint is $\hat{\mathbf{k}}_c \in \mathbb{R}^N$ and a query image $\mathbf{y}_q \in \mathbb{R}^N$, our goal is to design a projection matrix $\mathbf{\Phi}$ of size $L \times N$, $L < N$ that under statistic

$$r \doteq \langle \mathbf{\Phi} \hat{\mathbf{k}}_c, \mathbf{\Phi} \mathbf{w}_c \rangle \tag{2}$$

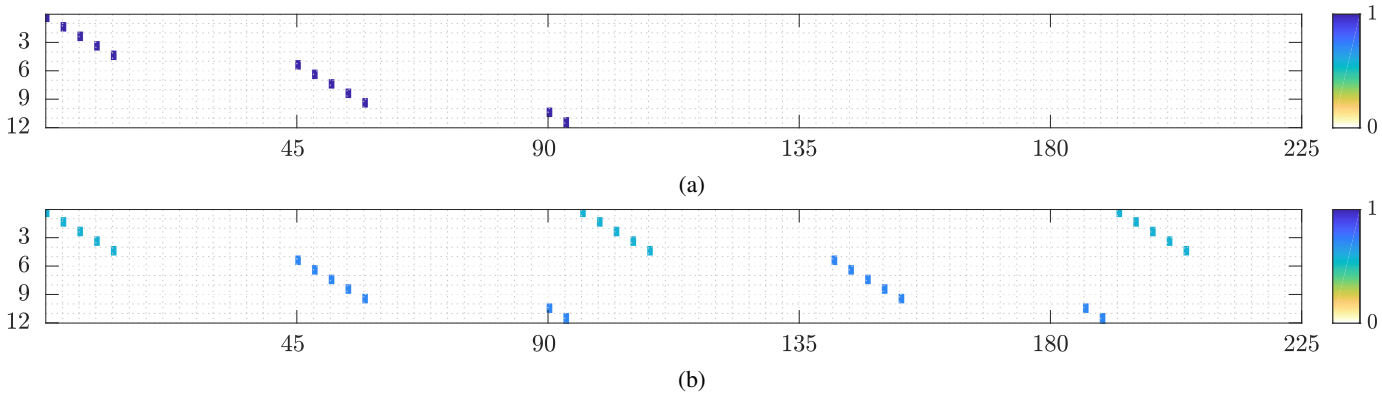


Fig. 2: Projection matrix $\mathbf{\Phi}$ designed as a bi-dimensional Sub-Sampling matrix $\mathbf{\Phi}^{ss}$ (a) and the derived Sub-Wrapping matrix $\mathbf{\Phi}^{sw}$ (b) in a small-scale example with $n = 15$, $L = 12$, $s = 3$.

enables dimensionality reduction of both $\hat{\mathbf{k}}_c$ and $\mathbf{w}_c$ while maximizing detection probability for H_1 hypothesis and minimizing false alarm probability for H_0 hypothesis.

III. Proposed projection method

We assume the existence of a zero-mean vector $\mathbf{k}_0$ with N i.i.d. components $\mathcal{N}(0, \sigma_k^2)$ describing the PRNU of a sensing device. The estimated device PRNU fingerprint $\hat{\mathbf{k}} \in \mathbb{R}^N$, extracted from a set $\mathcal{C}$ of images taken with the same camera device as from Eq. (1), is modeled as

$$\hat{\mathbf{k}} = \mathbf{H}_c \mathbf{k}_0 + \mathbf{n}_e \tag{3}$$

where $\mathbf{n}_e$ is the extraction noise which we assume zero-mean i.i.d., i.e. $\mathrm{E}\{\mathbf{n}_e \mathbf{n}_e^T\} = \sigma_e^2 \mathbf{I}$. Matrix $\mathbf{H}_c \in \mathbb{R}^{N \times N}$ accounts for demosaicing effects in the spatial domain and possibly the equivalent filtering performed in the DCT domain when images are compressed. All other effects due to compression are included in $\mathbf{n}_e$.

Given a query vector $\mathbf{y}$ from an image in the query images set $\mathcal{Q}$, and its denoised version $\mathbf{x}$, we model $\mathbf{x} = \mu_x \mathbf{1} + \tilde{\mathbf{x}}$, where $\tilde{\mathbf{x}}$ is assumed to be stationary with autocorrelation $r_{\tilde{x}}[i-j] \doteq E[\tilde{x}_i \tilde{x}_j]$.

The noise residual $\mathbf{y} - \mathbf{x}$ is modeled as $\mathbf{H}_q(\mathbf{k}_0 \circ \mathbf{x}) + \mathbf{n}_t$, where $\mathbf{n}_t$ is zero-mean i.i.d. noise, i.e. $\mathrm{E}\{\mathbf{n}_t \mathbf{n}_t^T\} = \sigma_t^2 \mathbf{I}$ and $\mathbf{H}_q$ plays a similar role to $\mathbf{H}_c$. In general $\mathbf{H}_c \neq \mathbf{H}_q$ because we assume that images from $\mathcal{C}$ and $\mathcal{Q}$ may be compressed differently. Also notice that more complicated noise models can be easily accommodated in our discussion, but we keep our choice for the sake of readability.

For dimensionality reduction, we assume that both $\hat{\mathbf{k}}$ and $\mathbf{y} - \mathbf{x}$ are projected using the dimensionality-reduction matrix $\mathbf{\Phi}$, so the test statistic is the one introduced in Eq. (2).

Under hypothesis H_1, we are interested in computing the mean and standard deviation of r, and find conditions on $\mathbf{\Phi}$ that maximize the Signal-to-Noise Ratio (SNR) defined as $\mathrm{SNR} = \mathrm{E}\{r\}/\sqrt{\mathrm{Var}\{r\}}$.

The mean term is

$$\mathrm{E}\{r\} = \mu_x \mathrm{E}\{\mathrm{Tr}[\mathbf{\Phi}_q \mathbf{k}_0 \mathbf{k}_0^T \mathbf{\Phi}_c^T]\} = \mu_x \sigma_k^2 \mathrm{Tr}[\mathbf{M}] \tag{4}$$

where $\mathbf{\Phi}_q \doteq \mathbf{\Phi}\mathbf{H}_q$, $\mathbf{\Phi}_c \doteq \mathbf{\Phi}\mathbf{H}_c$, and $\mathbf{M} \doteq \mathbf{\Phi}_q^T \mathbf{\Phi}_c$. We assume that projection vectors are normalized, so that each row of both $\mathbf{\Phi}_q$ and $\mathbf{\Phi}_c$ has l_2-norm equal to 1.

For the variance term, we show in [21] that

$$\begin{aligned}\mathrm{Var}\{r\} = \sigma_k^4 \Big(& \mu_x^2 \mathrm{Tr}[\mathbf{M}\mathbf{M}] + (\mu_x^2 + \sigma_x^2)\mathrm{Tr}[\mathbf{M}\mathbf{M}^T] \\ & + \sum_{i,j} r_{\tilde{x}}[i-j]\,(M_{i,i}M_{j,j} + M_{i,j}M_{j,i}) \Big) + \sigma_t^2 \sigma_k^2 L \\ & + \sigma_e^2 \sigma_k^2 (\mu_x^2 + \sigma_x^2) L + \sigma_e^2 \sigma_t^2 \mathrm{Tr}[(\mathbf{\Phi}\mathbf{\Phi}^T)^2]\end{aligned} \tag{5}$$

Recalling that $\mathrm{Tr}[\mathbf{\Phi}\mathbf{\Phi}^T] = L$, the last summand is minimized when $\mathbf{\Phi}\mathbf{\Phi}^T = \mathbf{I}$, that is when projection vectors are orthonormal. Minimization of (5) with respect to $\mathbf{M}$ is cumbersome; however, noticing that in practice μ_x^2 is often larger than σ_x^2, it makes sense to minimize instead $\sigma_k^4 \mu_x^2 \left(\mathrm{Tr}[\mathbf{M}\mathbf{M}] + \mathrm{Tr}[\mathbf{M}\mathbf{M}^T]\right)$, subject to $\mathrm{Tr}[\mathbf{M}] = L$.

Writing $\mathbf{M} = \mathbf{M}_S + \mathbf{M}_A$, where $\mathbf{M}_S$ and $\mathbf{M}_A$ are, respectively, symmetric and antisymmetric, the problem can be formulated as

$$\begin{aligned} \text{minimize} \quad & \mathrm{Tr}[\mathbf{M}\mathbf{M}_S] \\ \text{subject to} \quad & \mathrm{Tr}[\mathbf{M} + \mathbf{M}^T] = 2\mathrm{Tr}[\mathbf{M}_S] = 2L \end{aligned}$$

Observe that $\mathrm{Tr}[\mathbf{M}\mathbf{M}_S] = \mathrm{Tr}[(\mathbf{M}_S + \mathbf{M}_A)\mathbf{M}_S] = \mathrm{Tr}[\mathbf{M}_S\mathbf{M}_S]$. Then, we can pose the problem equivalently in terms of $\mathbf{M}_S$. In fact, if $\mathbf{\Lambda}$ is the diagonal matrix containing the eigenvalues of $\mathbf{M}_S$, the problem becomes

$$\begin{aligned} \text{minimize} \quad & \mathrm{Tr}[\mathbf{\Lambda}^2] \\ \text{subject to} \quad & \mathrm{Tr}[\mathbf{\Lambda}] = L \end{aligned} \tag{6}$$

From Cauchy-Schwarz inequality, the solution to this problem is achieved when all P non-null eigenvalues of $\mathbf{M}_S$ are identical and take the value L/P. Notice that $P = \mathrm{rank}(\mathbf{M}_S) \leq 2 \cdot \mathrm{rank}(\mathbf{M}) = 2L$. It is an ongoing problem to translate this general solution into design principles regarding $\mathbf{\Phi}_c$ and $\mathbf{\Phi}_q$. A particular approach consists in assuming that $\mathbf{M}$ is symmetric, in which case P above will be equal to L, and the eigenvalues of $\mathbf{\Phi}_c \mathbf{\Phi}_q^T$ will be all identical to 1. Therefore, we will seek projection matrices $\mathbf{\Phi}$ for which $\mathbf{\Phi}_c \mathbf{\Phi}_q^T = \mathbf{I}$.

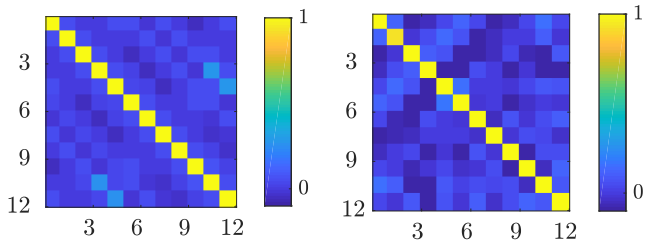


Fig. 3: $\mathbf{\Phi}_c\mathbf{\Phi}_q^T$ for $\mathbf{\Phi}$ is a Sub-Wrapping matrix (a) and when $\mathbf{\Phi}$ is a Gaussian matrix (b) in a small-scale example with $n = 15$, $L = 12$, $s = 3$ and real case $\mathbf{H}_c$ and $\mathbf{H}_q$.

On the other hand, for hypothesis H_0 we find that $\mathrm{E}\{r\} = 0$ and in [21] we show that $\mathrm{Var}\{r\} = (\mu_x^2+\sigma_x^2)\mathrm{Tr}[\mathbf{M}\mathbf{M}^T]$. Then, minimizing $\mathrm{Tr}[\mathbf{M}\mathbf{M}^T]$ also minimizes the variance for H_0.

IV. Projection matrix design

As shown in Sect. III, a condition to nearly maximize detection while minimizing false-alarms in a cross-correlation test is $\mathbf{\Phi}_c\mathbf{\Phi}_q^T = \mathbf{I}$. Since our goal is to design $\mathbf{\Phi}$, we first want to find the support of the point-spread function characterizing $\mathbf{H}_c$ and $\mathbf{H}_q$. To this end, we consider a bi-dimensional estimated camera PRNU fingerprint $\hat{\mathbf{K}}_c$ and compute its autocorrelation, as shown in Fig. 1a. If the estimated fingerprint were purely a white noise signal, its autocorrelation would result in a single peak at $(0, 0)$. However, demosaicing and low-pass filtering occurring during image compression lead the autocorrelation to be far from an impulse, thus with a support extended over more than one pixel. Approximating the PRNU autocorrelation as a bi-dimensional point-spread function with support $h \times h$, we can estimate $\mathbf{H}_c$ as a space-invariant bi-dimensional linear filter, as from the example in Fig. 1b. Following the same reasoning, it is possible to estimate the support of $\mathbf{H}_q$ via autocorrelation of $\mathbf{W}_q$, the bi-dimensional residual for image $\mathbf{Y}_q$. For the sake of simplicity, in the following we will consider $\mathbf{H}_q = \mathbf{H}_c$. The extension to the case where they are different is straightforward.

When it comes to the design of $\mathbf{\Phi}$, among the infinite set of solutions that lead to $\mathbf{\Phi}_c\mathbf{\Phi}_q^T = \mathbf{I}$, we propose two simple design strategies: i) Sub-Sampling; ii) Sub-Wrapping.

Sub-Sampling: The first proposed strategy is bi-dimensional subsampling with step-size s, cropped to the first L output coefficients. In order to build the projection matrix $\mathbf{\Phi}$ let us suppose we wish to project a bi-dimensional fingerprint $\hat{\mathbf{K}}_c \in \mathbb{R}^{n\times n}$, with $N = n^2$, L and n integer multiples of step-size s, such that $n = b \cdot s$, for some $b \in \mathbb{Z}$. Let us define

$$\mathbf{a} \doteq \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix} \in \{0,1\}^s$$
$$\mathbf{B} \doteq \begin{bmatrix} \mathbf{I}_{b\times b} & | & \mathbf{0}_{b\times b(s-1)} \end{bmatrix} \in \{0,1\}^{b\times bs}$$
$$\bar{\mathbf{\Phi}}^{ss} \doteq \mathbf{I}_{b\times b} \otimes (\mathbf{B} \otimes \mathbf{a}) \in \{0,1\}^{b^2\times N}$$

where $\mathbf{a}$ is a row vector of s elements, representing the structuring element for the subsampling matrix, $\mathbf{I}_{b\times b}$ is an identity matrix of size $b \times b$, $\mathbf{0}_{b\times b(s-1)}$ is an all-zero matrix of size $b \times b(s-1)$ and $\otimes$ represents the Kronecker product between matrices. The subsampling projection matrix $\mathbf{\Phi}^{ss} \in \{0,1\}^{L\times N}$ is then obtained by retaining the first L rows of $\bar{\mathbf{\Phi}}^{ss}$. An example for matrix $\mathbf{\Phi}^{ss}$ is shown in Fig. 2a.

Sub-Wrapping: The second strategy is an extension of subsampling with step-size s that takes into account the unused input coefficients due to cropping to length L. Cropped rows from $\bar{\mathbf{\Phi}}^{ss}$ are warped and summed to obtain $\mathbf{\Phi}^{sw}$, taking care of normalizing each row of $\mathbf{\Phi}^{sw}$. Formally, let us define

$$\delta \doteq Ls^2, \quad r \doteq \left\lceil \frac{b^2}{L} \right\rceil$$
$$\mathbf{Z}_{i,j}^{[z]} = \begin{cases} 1 & \text{if } i + z = j \\ 0 & \text{otherwise} \end{cases} \qquad i \in [1, N],\ j \in [1, N]$$
$$\mathbf{D} \doteq \sum_{z=0}^{r-1} \mathbf{Z}^{[z\delta]} \in \{0,1\}^{N\times N}$$

where δ is the offset, in input space, between successive warps, r is the maximum number of input samples contributing to a single output sample and $\mathbf{D}$ is an offset projection matrix. We can now define an over-sized projection matrix $\bar{\mathbf{\Phi}}^{sw}$ as

$$\bar{\mathbf{\Phi}}^{sw} \doteq \bar{\mathbf{\Phi}}^{ss}\mathbf{D} \in \{0,1\}^{b^2\times N}$$

and finally derive the actual subwrapping projection matrix $\mathbf{\Phi}^{sw} \in \mathbb{R}^{L\times N}$ by selecting the first L rows of $\bar{\mathbf{\Phi}}^{sw}$ and normalizing each row in l_2 norm. An example of the resulting projection matrix $\mathbf{\Phi}^{sw}$ is depicted in Fig. 2b.

Orthogonality. The design of $\mathbf{\Phi}^{ss}$and $\mathbf{\Phi}^{sw}$ presents two degrees of freedom, namely, s and $L \leq N/s^2$. When designing a Sub-Sampling or a Sub-Wrapping projection matrix, the rule that allows us to meet the $\mathbf{\Phi}_c\mathbf{\Phi}_q^T = \mathbf{I}$ requirement is $s \geq h$, since this way replicas of the point-spread function will not overlap in $\mathbf{\Phi}_c$ or $\mathbf{\Phi}_q$.

Fig. 3 shows an example of how Sub-Wrapping and Gaussian Random Projections affect $\mathbf{\Phi}_c\mathbf{\Phi}_q^T$ when real-world (i.e. with unlimited support) $\mathbf{H}_c$ and $\mathbf{H}_q$ are considered. With the same set of parameters of the figure ($n = 15, L = 12, s = 3$), the resulting values for $\mathrm{Tr}[\mathbf{\Lambda}^2]$ are 14.58, 14.92, 16.88 respectively for Sub-Sampling, Sub-Wrapping, Gaussian Random Projection. Recalling from Eq. (6) that the objective function to minimize is $\mathrm{Tr}[\mathbf{\Lambda}^2]$, we can see how Gaussian Random Projection, despite being built as an orthogonal basis, presents a higher value for $\mathrm{Tr}[\mathbf{\Lambda}^2]$ than Sub-Sampling and Sub-Wrapping, which leads to a worst expected performance, as we will confirm later.

Complexity. A final consideration about the design of $\mathbf{\Phi}$ is in terms of complexity. In the Sub-Sampling case, projection complexity is reduced only to samples selection, as no sum or multiplications are involved, thus $\mathcal{C}^{ss} = 1$. For the Sub-Wrapping case, sample selection is followed by summing of a maximum of r elements for each output sample, thus $\mathcal{C}^{sw} = L(r-1)$. When $\mathbf{\Phi}$ is built as a circulant matrix, as for the Gaussian Random Projections case, $\mathcal{C}^{RP} = 2N[2\log_2(N) + 3]$, considering a direct Fast-Fourier-Transform (FFT) of size N, a dot product between two complex vectors of length N and the final inverse FFT of size N. In a real-case application

with $n \approx 1,500$, $s = 3$, $b \approx 500$, $L \approx 150,000$ and $r = 2$, the complexity of Sub-Wrapping projection is $\mathcal{C}^{sw} \approx 150$K, less than 0.1% with respect to the complexity of Gaussian Random Projections $\mathcal{C}^{RP} \approx 200$M.

V. Experimental setup

In the following we provide details about evaluation metrics and datasets adopted to provide experimental evidences for the effectiveness of proposed projection methods. To ease the comparison among different camera models, we consider only aligned fingerprints and residuals at original resolution, meaning that we are not looking for rotation, cropping, or other affine transformations. Fingerprints and residuals have been center-cropped to 1500×1500 pixels. Projection performances are evaluated in terms of True-Positive Rate (TPR), at a specific false alarm probability of 0.05, against the fingerprint rate.

To verify the practical feasibility of the proposed methods, we perform two kinds of experiments. First, projected fingerprints and residuals are used without any kind of quantization, thus following the model presented in Sect. III, then we only keep the sign of projected fingerprints and residuals (i.e., one bit per projected component), to measure in a qualitative way the effect of quantization on the proposed model. Residuals rates are calculated as 32 bits per symbol when no quantization is applied, and as 1 bit per symbol when binarization is applied.

Datasets. Resorting to images from Dresden Image Database [19] we build two camera fingerprint datasets and two query residual datasets.

The first pair of datasets is built upon RAW images coming from 6 *Nikon* camera devices:

- $\mathcal{C}_f^{RAW}$ is composed of 6 camera fingerprints extracted from flatfield RAW images.
- $\mathcal{Q}_n^{RAW}$ is composed of 1317 query residuals extracted from natural RAW images.

The second pair of datasets is built upon JPEG images from 53 camera models, as from [18]:

- $\mathcal{C}_f^{JPG}$ is composed of 53 camera fingerprints extracted from flatfield JPEG images as encoded by the cameras' firmware.
- $\mathcal{Q}_n^{JPG}$ is composed of 9092 query residuals extracted from natural JPEG images as encoded by the cameras' firmware.

VI. Experiments

Experiments conducted in this Section aim at proving the compression effectiveness of the proposed PRNU fingerprint and residuals projection methods. Each reported plot presents four curves, each one related to a different compression pipeline:

- Crop (blue): a central square region of the fingerprint or residual is cropped and the rate is modulated by varying region size. This is a baseline reference method.
- Random Projections [RP] (green): Gaussian Random Projections implemented via circulant sensing matrices are used for fingerprint and residual projection. The rate is modulated by varying the projected subspace dimensionality.
- Sub-Sampling [SS] (purple): Bi-dimensional subsampling is implemented as depicted in Sect. IV. The rate is modulated by varying the downsamping step size s while keeping $L = bc$.
- Sub-Wrapping [SW] (red): Bi-dimensional subwrapping with $s = 3$. The rate is modulated by varying L.

Fig. 4 reports results obtained with camera fingerprints from $\mathcal{C}_f^{RAW}$ and query residuals from $\mathcal{Q}_f^{RAW}$. Images used to estimate camera fingerprints and query residuals have been affected only by demosaicing and neither projected fingerprints nor residuals are quantized. We observe how the SS, SW and RP perform at par, a symptom of almost diagonal $\mathbf{H}_c$ and $\mathbf{H}_q$ matrices due to reduced point-spread function support h.

Fig. 5 shows how SS and SW behavior changes when both camera fingerprints and query residuals are binarized after projection. Both SS and SW are now outperforming RP with a sensible compression improvement, e.g. a rate reduction in the order of 37% for a TPR of 0.99.

Fig. 6 reports results when images used to estimate camera fingerprints and query residuals have undergone JPEG compression by camera built-in firmware. In this case the point-spread function in $\mathbf{H}_c$ and $\mathbf{H}_q$ have surely larger support. This experiment resembles a real-world scenario, where camera fingerprints are taken from $\mathcal{C}_f^{JPG}$, query residuals are from $\mathcal{Q}_n^{JPG}$ and no quantization is applied to projected fingerprints and residuals. In this case SS is performing as RP, with better TPR for bitrates smaller than 1700kbit per residual. SW is instead always offering a TPR improvement in the order of 1% meaning a rate reduction of nearly 23% at a fixed TPR of 0.9.

Finally, Fig. 7 shows how binarization affects the aforementioned real case scenario on $\mathcal{C}_f^{JPG}$ and $\mathcal{Q}_f^{JPG}$. While with similar performance, among the three methods SW achieves better detection performances at equal bitrate with respect to RP and SS.

In any case, we once again remark that the complexity of the new projection methods is much lower than that of RP.

VII. Conclusions

In this work we have proposed a novel projection framework for PRNU fingerprint and residuals compression based on SNR near maximization of the alternative hypothesis in a cross-correlation statistical test. We have derived systematic design conditions for the projection matrix, taking into account local interpolation effects on PRNU due to demosaicing and JPEG compression. Two projection design strategies have been introduced and tested on real-world datasets, in comparison with state-of-the-art PRNU compression methods. The obtained experimental results confirm that proposed projection methods perform at par or better with state-of-the-art Gaussian Random Projection, with a signifcant reduction in terms of computational complexity.

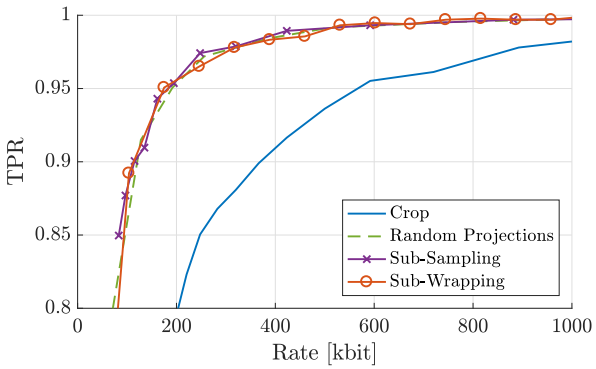


Fig. 4: Rate vs TPR when camera fingerprints and query residuals are extracted from RAW images. Fingerprints and residuals are not quantized.

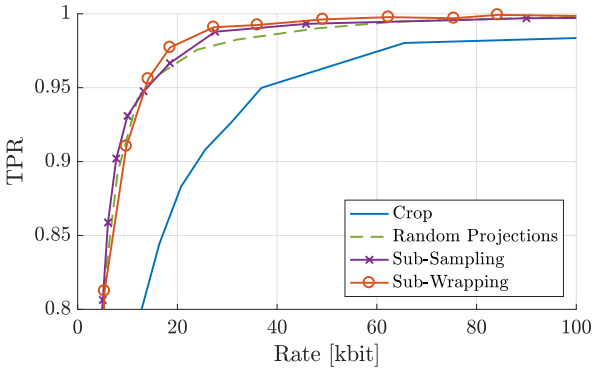


Fig. 5: Rate vs TPR when camera fingerprints and query residuals are extracted from RAW images. Fingerprints and residuals are binarized after projection.

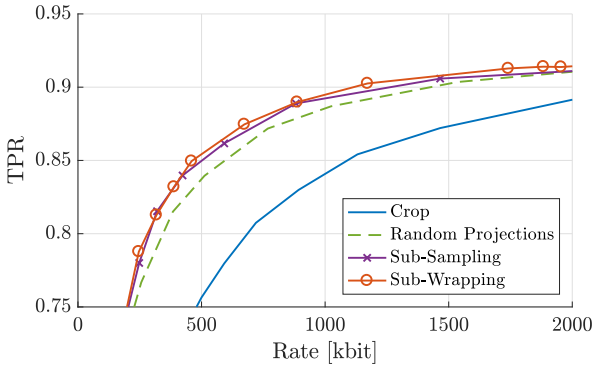


Fig. 6: Rate vs TPR when camera fingerprints and query residuals are extracted from JPEG images. Fingerprints and residuals are not quantized.

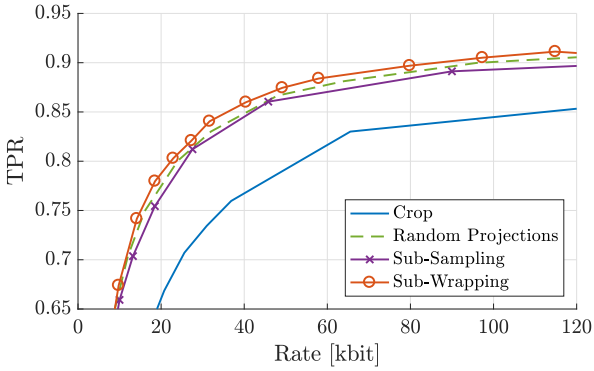


Fig. 7: Rate vs TPR when camera fingerprints and query residuals are extracted from JPEG images. Fingerprints and residuals are binarized after projection.

References

[1] J. Lukas, J. Fridrich, and M. Goljan, "Determining digital image origin using sensor imperfections," in *SPIE Conference on Image and Video Communications and Processing*, 2005.

[2] M. Goljan, M. Chen, and J. Fridrich, "Identifying common source digital camera from image pairs," in *IEEE International Conference on Image Processing (ICIP)*, 2007.

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

[4] J. Fridrich, "Digital image forensics," *IEEE Signal Processing Magazine*, vol. 26, no. 2, pp. 26–37, mar 2009.

[5] J. Lukas, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," *IEEE Transactions on Information Forensics and Security (TIFS)*, vol. 1, no. 2, pp. 205–214, jun 2006.

[6] I. Amerini, R. Caldelli, V. Cappellini, F. Picchioni, and A. Piva, "Analysis of denoising filters for photo response non uniformity noise extraction in source camera identification," in *International Conference on Digital Signal Processing (DSP)*, 2009.

[7] F. Gisolf, A. Malgoezar, T. Baar, and Z. Geradts, "Improving source camera identification using a simplified total variation based noise removal algorithm," *Digital Investigation*, vol. 10, no. 3, pp. 207–214, 2013.

[8] X. Lin and C.-T. Li, "Preprocessing reference sensor pattern noise via spectrum equalization," *IEEE Transactions on Information Forensics and Security (TIFS)*, vol. 11, no. 1, pp. 126–140, jan 2016.

[9] M. Goljan and J. Fridrich, "Camera identification from cropped and scaled images," in *SPIE Confernece on Security, Forensics, Steganography, and Watermarking of Multimedia Contents*, 2008.

[10] K. Rosenfeld and H. T. Sencar, "A study of the robustness of PRNU-based camera identification," *SPIE Conference on Media Forensics and Security*, 2009.

[11] M. Goljan, J. Fridrich, and T. Filler, "Large scale test of sensor fingerprint camera identification," *Proceedings of SPIE*, 2009.

[12] R. Caldelli, I. Amerini, F. Picchioni, and M. Innocenti, "Fast image clustering of unknown source images," in *IEEE International Workshop on Information Forensics and Security (WIFS)*, 2010.

[13] M. Goljan, J. Fridrich, and T. Filler, "Managing a large database of camera fingerprints," in *SPIE Conference on Media Forensics and Security*, 2010.

[14] Y. Hu, C.-T. Li, Z. Lai, and S. Zhang, "Fast camera fingerprint search algorithm for source camera identification," in *International Symposium on Communications, Control and Signal Processing*, 2012.

[15] Y. Hu, C.-T. Li, and Z. Lai, "Fast source camera identification using matching signs between query and reference fingerprints," *Multimedia Tools and Applications*, vol. 74, no. 18, pp. 7405–7428, may 2014.

[16] F. Perez-Gonzalez, M. Masciopinto, I. Gonzalez-Iglesias, and P. Comesana, "Fast sequential forensic detection of camera fingerprint," in *IEEE International Conference on Image Processing (ICIP)*, 2016.

[17] D. Valsesia, G. Coluccia, T. Bianchi, and E. Magli, "Large-scale image retrieval based on compressed camera identification," *IEEE Transactions on Multimedia (TMM)*, vol. 17, no. 9, pp. 1439–1449, sep 2015.

[18] ——, "Compressed fingerprint matching and camera identification via random projections," *IEEE Transactions on Information Forensics and Security (TIFS)*, vol. 10, no. 7, pp. 1472–1485, jul 2015.

[19] T. Gloe and R. Bhme, "The dresden image database for benchmarking digital image forensics," *Journal of Digital Forensic Practice*, vol. 3, no. 2-4, pp. 150–159, dec 2010.

[20] M. K. Mihcak, I. Kozintsev, K. Ramchandran, and P. Moulin, "Low-complexity image denoising based on statistical modeling of wavelet coefficients," *IEEE Signal Processing Letters (SPL)*, vol. 6, pp. 300–303, 1999.

[21] "Design of projection matrices for prnu compression," Tech. Rep. [Online]. Available: https://goo.gl/Gm238t