Chapter 4
Sensor Fingerprints: Camera Identification and Beyond

Matthias Kirchner

Every imaging sensor introduces a certain amount of noise to the images it captures—slight fluctuations in the intensity of individual pixels even when the sensor plane was lit absolutely homogeneously. One of the breakthrough discoveries in multimedia forensics is that photo-response non-uniformity (PRNU), a multiplicative noise component caused by inevitable variations in the manufacturing process of sensor elements, is essentially a sensor fingerprint that can be estimated from and detected in arbitrary images. This chapter reviews the rich body of literature on camera identification from sensor noise fingerprints with an emphasis on still images from digital cameras and the evolving challenges in this domain.

4.1 Introduction

Sensor noise fingerprints have been a cornerstone of media forensics ever since Lukáš et al. (2005) observed that digital images can be traced back to their sensor based on unique noise characteristics. Minute manufacturing imperfections are believed to make every sensor physically unique, leading to the presence of a weak yet deterministic sensor pattern noise in images captured by the camera (Fridrich 2013). This fingerprint, commonly referred to as photo-response non-uniformity (PRNU), can be estimated from images captured by a specific camera for the purpose of source camera identification. As illustrated in Fig. 4.1, a noise signal is extracted in this process at test time from a probe image of unknown provenance, which can then be compared against pre-computed fingerprint estimates from a set of candidate cameras.

M. Kirchner (✉)
Kitware, Inc., Clifton Park, NY, USA
e-mail: matthias.kirchner@kitware.com

H. T. Sencar et al. (eds.), *Multimedia Forensics*, Advances in Computer Vision and Pattern Recognition, https://doi.org/10.1007/978-981-16-7621-5_4

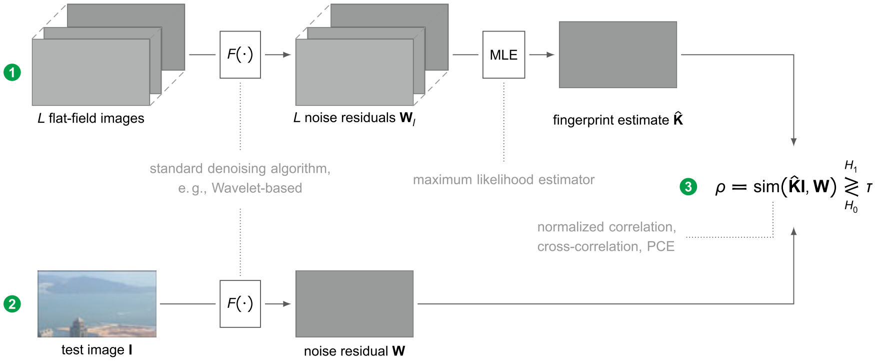

1
L flat-field images
F(·)
L noise residuals $\mathbf{W}_l$
MLE
fingerprint estimate $\hat{\mathbf{K}}$
standard denoising algorithm,
e. g., Wavelet-based
maximum likelihood estimator
3
$\rho = \text{sim}(\hat{\mathbf{K}}\mathbf{I}, \mathbf{W}) \underset{H_0}{\overset{H_1}{\gtrless}} \tau$
normalized correlation,
cross-correlation, PCE
2
test image $\mathbf{I}$
F(·)
noise residual $\mathbf{W}$

Fig. 4.1 Basic camera identification from PRNU-based sensor noise fingerprints (Fridrich 2013). (1) A camera fingerprint $\hat{\mathbf{K}}$ is estimated from a number flatfield images taken by the camera of interest. (2) At test time, a noise residual $\mathbf{W}$ is extracted from a probe image of unknown provenance. (3) A detector decides whether or not the probe image originates from the camera of interest by evaluating a suitable similarity score, ρ

Because PRNU emerges from physical noise-like properties of individual sensor elements, it has a number of attractive characteristics for source device attribution (Fridrich 2013). The fingerprint signal appears random and is of high dimensionality, which makes the probability of two sensors having the same fingerprint extremely low (Goljan et al. 2009). At the same time, it can be assumed that all common imaging sensor types exhibit PRNU, that each sensor output contains a PRNU component, except for completely dark or saturated images, and that PRNU fingerprints are stable over time (Lukás et al. 2006). Finally, various independent studies have found that the fingerprint is highly robust to common forms of post-processing, including lossy compression and filtering.

The goal of this chapter is not to regurgitate the theoretical foundations of the subject at length, as these have been discussed coherently before, including in a dedicated chapter in the first edition of this book (Fridrich 2013). Instead, we hope to give readers an overview of the ongoing research and the evolving challenges in the domain while keeping the focus on conveying the important concepts.

So why is it that there is still an abundance of active research when extensive empirical evidence (Goljan et al. 2009) has already established the feasibility of highly reliable PRNU-based consumer camera identification at scale? The simple answer is technological progress. Modern cameras, particularly those installed in smartphones, go to great lengths to produce visually appealing imagery. Imaging pipelines are ever evolving, and computational photography challenges our understanding of what a “camera-original” image looks like. On the flip side, many of these new processing steps interfere with the underlying assumptions at the core of PRNU-based camera identification, which requires strictly that the probe image and the camera fingerprint are spatially aligned with respect to the camera sensor elements. Techniques such as lens distortion correction (Goljan and Fridrich 2012), electronic image stabilization (Taspinar et al. 2016), or high dynamic range imaging (Shaya et al. 2018) have all been found to impede camera identification if not accounted for through spatial resynchronization. Robustness to low resolution and strong compression is another concern (van Houten and Geradts 2009; Chuang et al. 2011; Goljan et al. 2016; Altinisik et al. 2020, i. a.) that has been gaining more and more practical relevance due the widespread sharing of visual media through online social networks (Amerini et al. 2017; Meij and Geradts 2018). At the same time, the remarkable success of PRNU-based camera identification has also surfaced concerns for the anonymity of photographers, who may become identifiable through the analysis and combination of information derived from one or multiple images. As a result, the desire to protect anonymous image communication, e.g., in the case of journalism, activism, or legitimate whistle-blowing, has brought counter-forensic techniques (Böhme and Kirchner 2013) to suppress traces of origin in digital images to the forefront.

We discuss these and related topics with a specific focus on still camera images in more detail below. Our treatment of the subject here is complemented by dedicated chapters on video source attribution (Chap. 5) and large scale camera identification (Chap. 6). An overview of computational photography is given in Chap. 3. Sections 4.2 and 4.3 start with the basics of sensor noise fingerprint estimation and camera identification, before Sect. 4.4 delves into the challenges related to spatial sensor misalignment. Section 4.5 focuses on recent advances in image manipulation localization based on PRNU fingerprints. Counter-forensic techniques for fingerprint removal and copying are discussed in Sect. 4.6, while Sect. 4.7 highlights early attempts to apply tools from the field of deep learning to domain-specific problems. A brief overview of relevant public datasets in Sect. 4.8 follows, before Sect. 4.9 concludes the chapter.

4.2 Sensor Noise Fingerprints

State-of-the-art sensor noise forensics assumes a simplified imaging model for single-channel images $\mathbf{I}(m, n)$, $0 \leq m < M$, $0 \leq n < N$,

$$\mathbf{I} = \mathbf{I}^{(o)}(\mathbf{1} + \mathbf{K}) + \boldsymbol{\theta}\,, \tag{4.1}$$

in which the multiplicative PRNU factor $\mathbf{K}$ modulates the noise-free image $\mathbf{I}^{(o)}$, while $\boldsymbol{\theta}$ comprises a variety of additive noise components (Fridrich 2013). Ample empirical evidence suggests that signal $\mathbf{K}$ is a unique and robust camera fingerprint (Goljan et al. 2009). It can be estimated from a set of L images taken with the specific sensor of interest. The standard procedure relies on a denoising filter $F(\cdot)$ to obtain a noise residual,

$$\mathbf{W}_l = \mathbf{I}_l - F(\mathbf{I}_l)\,, \tag{4.2}$$

from the l-th image $\mathbf{I}_l$, $0 \leq l < L$. The filter acts mainly as a means to increase the signal-to-noise ratio between the signal of interest (the fingerprint) and the observed image. To date, most works still resort to the wavelet-based filter as adopted by Lukáš et al. (2006) for its efficiency and generally favorable performance, although a number of studies have found that alternative denoising algorithms can lead to moderate improvements (Amerini et al. 2009; Cortiana et al. 2011; Al-Ani and Khelifi 2017; Chierchia et al. 2014, i. a.). In general, it is accepted that noise residuals obtained from off-the-shelf filters are imperfect by nature and that they are contaminated non-trivially by remnants of image content. Salient textures or quantization noise may exacerbate the issue. For practical applications, a simplified modeling assumption

$$\mathbf{W}_l = \mathbf{K}\mathbf{I}_l + \boldsymbol{\eta}_l \tag{4.3}$$

with i. i. d. Gaussian noise $\boldsymbol{\eta}_l$ leads to a maximum likelihood estimate (MLE) $\hat{\mathbf{K}}$ of the PRNU fingerprint of the form (Fridrich 2013)

$$\hat{\mathbf{K}} = \frac{\sum_l \mathbf{W}_l \mathbf{I}_l}{\sum_l \mathbf{I}_l^2} \,. \tag{4.4}$$

In this procedure, it is assumed that all images $\mathbf{I}_l$ are spatially aligned so that the pixel-wise operations are effectively carried out over the same physical sensor elements.

If available, it is beneficial to use flatfield images for fingerprint estimation to minimize contamination from image content. The quality of the estimate in Eq. (4.4) generally improves with L, but a handful of homogeneously lit images typically suffices in practice. Uncompressed or raw sensor output is preferable over compressed images, as it will naturally reduce the strength of unwanted nuisance signals in Eq. (4.1). When working with raw output from a sensor with a color filter array (CFA), it is advisable to subsample the images based on the CFA layout (Simon et al. 2009).

Practical applications warrant a post-processing step to clean the fingerprint estimate from non-unique artifacts, which may otherwise increase the likelihood of false fingerprint matches. Such artifacts originate from common signal characteristics that occur consistently across various devices, for instance, due to the distinctively structured layout of the CFA or block-based JPEG compression. It is thus strongly recommended to subject $\hat{\mathbf{K}}$ to zero-meaning and frequency-domain Wiener filtering, as detailed in Fridrich (2013). Additional post-processing operations have been discussed in the literature (Li 2010; Kang et al. 2012; Lin and Li 2016, i. a.), although their overall merit is often marginal (Al-Ani and Khelifi 2017). Other non-trivial non-unique artifacts have been documented as well. Cases of false source attribution linked to lens distortion correction have been reported when images from a different camera were captured at a focal length that was prominently featured during fingerprint estimation (Goljan and Fridrich 2012; Gloe et al. 2012). Non-unique artifacts introduced by advanced image enhancement algorithms in modern devices are a subject of ongoing research (Baracchi et al. 2020; Iuliani et al. 2021).

For multi-channel images, the above procedures can be applied to each color channel individually before averaging the obtained signals into a single-channel fingerprint estimate (Fridrich 2013).

4.3 Camera Identification

For a given probe image **I** of unknown provenance, camera identification can be formulated as a hypothesis testing problem:

$$\begin{aligned} &H_0 : \mathbf{W} = \mathbf{I} - F(\mathbf{I}) \text{ does not contain the fingerprint of interest, } \mathbf{K} \\ &H_1 : \mathbf{W} \text{ does contain the fingerprint } \mathbf{K}\,; \end{aligned}$$

i. e., the probe is attributed to the tested camera if H_1 holds. In practice, the test can be decided by evaluating a similarity measure ρ,

$$\rho = \mathrm{sim}(\mathbf{W}, \hat{\mathbf{K}}\mathbf{I}) \gtrless_{H_0}^{H_1} \tau\,. \tag{4.5}$$

for a suitable threshold τ. Under the modeling assumptions adopted in the literature (Fridrich 2013), the basic building block for this is the normalized cross-correlation (NCC), which is computed over a grid of shifts $\boldsymbol{s} = (s_1, s_2)$, $0 \leq s_1 < M$, $0 \leq s_2 < N$, for two matrices **A**, **B** as

$$\mathbf{NCC}_{\mathbf{A},\mathbf{B}}(s_1, s_2) = \frac{\sum_{m,n} \left(\mathbf{A}(m,n) - \bar{\mathbf{A}}\right) \left(\mathbf{B}(m+s_1, n+s_2) - \bar{\mathbf{B}}\right)}{\left\|\mathbf{A} - \bar{\mathbf{A}}\right\| \left\|\mathbf{B} - \bar{\mathbf{B}}\right\|}\,. \tag{4.6}$$

We assume implicitly that matrices **A** and **B** are of equal dimension and that zero-padding has been applied to assert this where necessary. In practice, the above expression is evaluated efficiently in the frequency domain. It is common in the field to approximate Eq. (4.6) by working with the circular cross-correlation, i. e., by operating on the FFTs of matrices **A**, **B** without additional zero-padding. Taking the maximum NCC over a set $\mathcal{S}$ of admissible shifts as similarity,

$$\rho_{\mathrm{ncc}}^{(\mathcal{S})} = \max_{\boldsymbol{s} \in \mathcal{S}} \mathbf{NCC}_{\mathbf{W},\hat{\mathbf{K}}\mathbf{Y}}(\boldsymbol{s})\,, \tag{4.7}$$

can conveniently account for potential sensor misalignment between the tested fingerprint and the probe as they would result from different sensor resolutions and/or cropping. Peak-to-correlation energy (PCE) has been proposed as an alternative that mitigates the need for sensor-specific thresholds

$$\rho_{\mathrm{pce}}^{(\mathcal{S})} = \frac{(MN - |\mathcal{N}|) \cdot \left(\rho_{\mathrm{ncc}}^{(\mathcal{S})}\right)^2}{\sum_{\mathbf{s} \notin \mathcal{N}} \mathbf{NCC}^2_{\mathbf{W},\hat{\mathbf{K}}\mathbf{Y}}(\boldsymbol{s})}\,. \tag{4.8}$$

In the equation above, $\mathcal{N}$ denotes a small neighborhood around the peak NCC. It is commonly set to a size of 11×11 (Goljan et al. 2009). In practice, camera identification must account for the possibility of mismatching sensor orientations in the probe image and in the fingerprint estimate, so the test for the presence of the fingerprint should be repeated with one of the signals rotated by 180° if the initial test stays below the threshold.

The choice of set $\mathcal{S}$ crucially impacts the characteristics of the detector, as a larger search grid will naturally increase the variance of detector responses on true negatives. If it can be ruled out a priori that the probe image underwent cropping, the "search" can be confined to $\mathcal{S} = \{(0, 0)\}$ to reduce the probability of false alarms. Interested readers will find a detailed error analysis for this scenario in Goljan et al. (2009), where the authors determined a false alarm rate of 2.4×10^{-5} while attributing 97.62% of probes to their correct source in experiments with more than one million images from several thousand devices. The PCE threshold for this operating point is 60, which is now used widely in the field as a result.

A number of variants of the core camera identification formulation can be addressed with only a little modification (Fridrich 2013). For instance, it can be of interest to determine whether two arbitrary images originate from the same device, without knowledge or assumptions about associated camera fingerprints (Goljan et al. 2007). In a related scenario, the goal is to compare and match a number of fingerprint estimates, which becomes particularly relevant in setups with the objective of clustering a large set of images by their source device (Bloy 2008; Li 2010; Amerini et al. 2014; Marra et al. 2017, i.a.).

Specific adaptations also exist for testing against large databases of camera fingerprints, where computational performance becomes a relevant dimension to monitor in practice. This includes efforts to reduce the dimensionality of camera fingerprints (Goljan et al. 2010; Bayram et al. 2012; Valsesia et al. 2015, i.a.) and protocols for efficient fingerprint search and matching (Bayram et al. 2015; Valsesia et al. 2015; Taspinar et al. 2020, i.a.). We refer the reader to Chap. 6 in this book which discusses these techniques in detail.

4.4 Sensor Misalignment

Camera identification from PRNU-based sensor fingerprints can only succeed if the probe image and the tested fingerprint are spatially aligned. While cross-correlation can readily account for translation and cropping, additional steps become inevitable if more general geometric transformations have to be considered. This includes, for instance, combinations of scaling and cropping when dealing with the variety of image resolutions and aspect ratios supported by modern devices (Goljan and Fridrich 2008; Tandogan et al. 2019). Specifically, assuming that a probe image $\mathbf{I}$ underwent a geometric transform $T_{\mathbf{u}^*}(\cdot)$ with parameters $\mathbf{u}^*$, we want to carry out the above basic procedures on $\mathbf{I} \leftarrow T_{\mathbf{u}^*}^{-1}(\mathbf{I})$. If the parameters of the transform are unknown, the problem essentially translates into a search over a set $\mathcal{U}$ of admissible candidate transforms $\mathbf{u}$. In practice, the maximum similarity over all candidate transforms will determine whether or not an image contains the tested fingerprint,

$$\rho = \max_{\mathbf{u}\in\mathcal{U}} \operatorname{sim}\left(T_{\mathbf{u}}^{-1}(\mathbf{I}) - F\left(T_{\mathbf{u}}^{-1}(\mathbf{I})\right), \hat{\mathbf{K}}T_{\mathbf{u}}^{-1}(\mathbf{I})\right) \gtrless_{H_0}^{H_1} \tau\,, \tag{4.9}$$

and the detector threshold should be adjusted accordingly compared to the simpler case above to maintain a prescribed false alarm rate. Unfortunately, this search can quickly become computationally expensive. A number of approximations have been proposed to speed up the search (Goljan and Fridrich 2008), including applying the inverse transform to a precomputed noise residual $\mathbf{W}$ instead of recomputing the noise residual for each candidate transform, and evaluating $\operatorname{sim}(T_{\mathbf{u}}^{-1}(\mathbf{WI}), \hat{\mathbf{K}})$ instead of $\operatorname{sim}(T_{\mathbf{u}}^{-1}(\mathbf{W}), \hat{\mathbf{K}}T_{\mathbf{u}}^{-1}(\mathbf{I}))$.[1] As for the search itself, a coarse-to-fine grid search is recommended by Goljan and Fridrich (2008) for potentially scaled and cropped images, while Mandelli et al. (2020) adopt a particle swarm optimization technique. Gradient-based search methods generally do typically not apply to the problem due to the noise-like characteristics of detector responses outside of a very sharp peak around the correct candidate transform (Fridrich 2013).

Sensor misalignment may not only be caused by "conventional" image processing. With camera manufacturers constantly striving to improve the visual image quality that their devices deliver to the customer, advanced in-camera processing and the rise of computational photography in modern imaging pipelines can pose significant challenges in that regard. One of the earliest realizations along those lines was that computational lens distortion correction can introduce non-linear spatial misalignment that needs to be accounted for (Goljan and Fridrich 2012; Gloe et al. 2012). Of particular interest is lens radial distortion, which lets straight lines in a scene appear curved in the captured image. This type of distortion is especially prominent for zoom lenses as they are commonly available for a wide variety of consumer cameras. For a good trade-off between lens size, cost and visual quality, cameras correct for lens radial distortion through warping the captured image according to a suitable compensation model that effectively inverts the nuisance warping introduced by the lens. This kind of post-processing will displace image content relative to the sensor elements. It impairs camera identification because the strength of lens radial distortion depends on the focal length, i.e., images taken by the same camera at different focal lengths will undergo different warping in the process of lens distortion correction. As a result, a probe image captured at a certain focal length that was not (well) represented in the estimation of the camera fingerprint $\hat{\mathbf{K}}$ in Eq. (4.4) may not be associated with its source device inadvertently.

A simple first-order parametric model to describe and invert radially symmetric barrel/pincushion distortion facilitates camera identification via a coarse-to-fine search over a single parameter (Goljan and Fridrich 2012), similar to the handling of general geometric transformations in Eq. (4.9) above. The model makes a number

[1] The two expressions are equivalent when used in combination with the PCE whenever it can be assumed that $T_{\mathbf{u}}^{-1}(\mathbf{W}) \cdot T_{\mathbf{u}}^{-1}(\mathbf{I}) \approx T_{\mathbf{u}}^{-1}(\mathbf{WI})$.

of simplifying assumptions that may not always hold in practice to avoid a higher dimensional search space, including the assumed concurrence of the optical center and the image center. Crucially, the practical applicability of such an approach first and foremost depends on the validity of the assumed distortion correction model in real cameras. In most circumstances, it is ultimately not fully known which design choices camera manufacturers make, and we are not aware of published large-scale evaluations that span a significant number of different devices/lenses. There are incidental observations of missed detections that suggest deviations from a purely radial distortion correction model (Gloe et al. 2012; Goljan and Fridrich 2014), however.

One of the reasons why interest in the effects of lens distortion correction and their remedies seems to have waned over the past years may be the enormous gain in the popularity of smartphones and similar camera-equipped mobile devices. These devices operate under very different optical and computational constraints than "conventional" consumer cameras and have introduced their own set of challenges to the field of PRNU-based camera identification. The source attribution of video data, and in particular the effects of electronic image stabilization (EIS), have arguably been the heavyweight in this research domain, and we direct readers to Chap. 5 of this book for a dedicated exposition. In the context of this chapter, it suffices to say that EIS effectively introduces gradually varying spatial misalignment to sequences of video frames, which calls for especially efficient computational correction approaches (Taspinar et al. 2016; Iuliani et al. 2019; Mandelli et al. 2020; Altinisik and Sencar 2021, i.a.).

Other types of in-camera processing are more and more moving into the focus as well. For instance, high dynamic range (HDR) imaging (Artusi et al. 2017) is routinely supported on many devices today and promises enhanced visual quality, especially also under challenging light conditions. In smartphones, it is usually realized by fusing a sequence of images of a scene, each taken at a different exposure setting in rapid succession. This requires registration of the images to mitigate global misalignment due to camera motion and local misalignment due to moving objects in the scene. In practice, the HDR image will be a content-dependent weighted mixture of the individual exposures, and different regions in the image may have undergone different geometric transformations. Without additional precautions, camera identification may fail under these conditions (Shaya et al. 2018). While it seems infeasible to express such complex locally varying geometric transformations in a parametric model, it is possible to conduct a search over candidate transformation on local regions of the image. Empirical evidence from a handful of smartphones suggests that the contributions of the individual exposures can be locally synchronized via cross-correlation, after correcting for a global, possibly anisotropic rescaling operation (Hosseini and Goljan 2019). Future research will have to determine whether such approach generalizes.

Sensor misalignment can be expected to remain a practical challenge, and it is likely to take on new forms and shapes as imaging pipelines continue to evolve. We surmise that the nature of misalignments will broaden beyond purely geometrical characteristics with the continued rise of computational photography and camera manufacturers allowing app developers access to raw sensor measurements. Empirical reports of impaired camera identification across mismatching imaging pipelines may be taken as first cautionary writing on the wall (Joshi et al. 2020).

4.5 Image Manipulation Localization

When a region of an image is replaced with content from elsewhere, the new content will lack the characteristic camera PRNU fingerprint one would expect to find otherwise. This is true irrespective of whether the inserted content has been copied from within the same image, or from a different image. Recasting camera identification as a local test for the presence of an expected fingerprint thus allows for the detection and localization of image manipulations (Chen et al. 2008; Fridrich 2013).

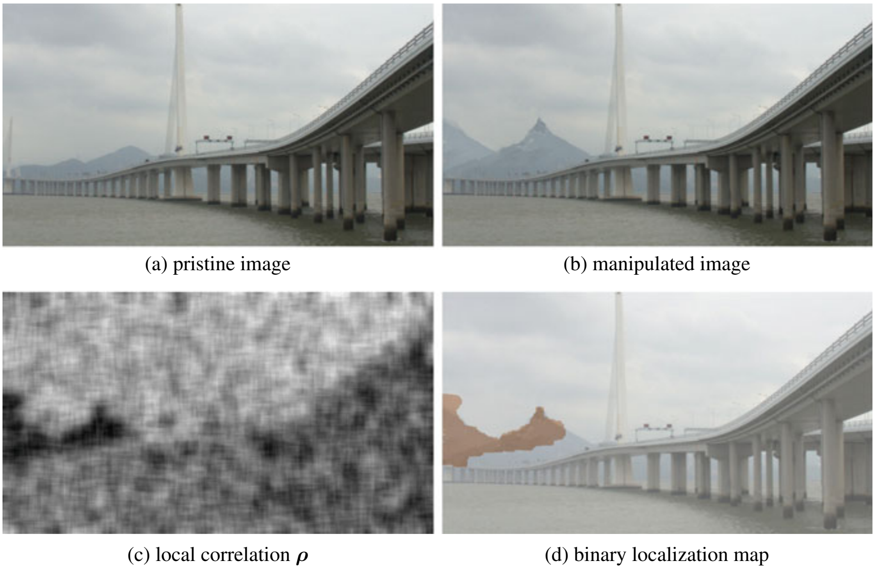
A straightforward approach is to examine the probe image $\mathbf{I}$ by sliding an analysis window of size $B \times B$ over the probe image, and to assign a binary label $\mathbf{Y}(m,n) \in \{-1, 1\}$,

$$\mathbf{Y}(m,n) = \operatorname{sgn}\left(\boldsymbol{\rho}(m,n) - \tau\right) \ , \tag{4.10}$$

to the window centered around location (m,n), with $\boldsymbol{\rho}(m,n)$ obtained by evaluating Eq. (4.5) for the corresponding analysis window. The resulting binary map $\boldsymbol{Y}$ will then be indicative of local manipulations, with $\mathbf{Y}(m,n) = -1$ corresponding to the absence of the fingerprint in the respective neighborhood. The literature mostly resorts to the normalized correlation for this purpose, $\rho = \rho_{\text{ncc}}^{(\{\mathbf{0}\})}$. It can be computed efficiently in one sweep for all sliding windows by implementing the necessary summations as linear filtering operations on the whole image.

The localization of small manipulated regions warrants sufficiently small analysis windows, which impacts the ability to reliably establish whether or not the expected fingerprint is present negatively. The literature often finds a window size of $B = 64$ as a reasonable trade-off between resolution and accuracy (Chierchia et al. 2014; Chakraborty and Kirchner 2017; Korus and Huang 2017). A core problem for analysis windows that small is that the measured local correlation under H_1 depends greatly on local image characteristics. One possible remedy is to formulate a camera-specific correlation predictor $\hat{\boldsymbol{\rho}}(m,n)$ that uses local image characteristics to predict how strongly the noise residual in a particular analysis window is expected to correlate with the purported camera fingerprint under H_1 (Chen et al. 2008). The decision whether to declare the absence of the tested fingerprint can then be conditioned on the expected correlation.

Adopting the rationale that more conservative decisions should be in place when the local correlation cannot be expected to take on large values per se, an adjusted binary labeling rule decides $\mathbf{Y}(m,n) = -1$ iff $\boldsymbol{\rho}(m,n) \leq \tau$ and $\hat{\boldsymbol{\rho}}(m,n) > \lambda$, where threshold λ effectively bounds the probability of missed fingerprint detection under H_1 (Chen et al. 2008). To ensure that the binary localization map mostly contains connected regions of a minimum achievable size, a pruning and post-processing step with morphological filters is recommended.


(a) pristine image

(b) manipulated image

(c) local correlation ρ

(d) binary localization map

Fig. 4.2 Image manipulation localization. The local correlation between the camera fingerprint and the noise residual extracted from the manipulated image is lowest (darker) in the manipulated region. It was computed from sliding windows of size 64 × 64. The binary localization map (overlayed on top of the manipulated image) was obtained with a conditional random field approach that evaluates the difference between measured and predicted correlation, $\rho - \hat{\rho}$ (Chakraborty and Kirchner 2017). Images taken from the Realistic Tampering Dataset (Korus and Huang 2017)

Significant improvements have been reported when explicitly accounting for the observation that local decisions from neighboring sliding windows are interdependent. A natural formulation follows from approaching the problem in a global optimization framework with the objective of finding the optimal mapping

$$\mathbf{Y}^* = \arg\max_{\mathbf{Y}} p\left(\mathbf{Y}|\rho, \hat{\rho}\right) . \tag{4.11}$$

This sets the stage for rewarding piecewise constant label maps via a variety of probabilistic graphical modeling techniques such as Markov random fields (Chierchia et al. 2014) and conditional random fields (Chakraborty and Kirchner 2017; Korus and Huang 2017). Figure 4.2 gives an example result.

A general downside of working with fixed-sized sliding windows is that $\rho(m, n)$ will naturally change only very gradually, which makes the detection of very small manipulated regions challenging (Chierchia et al. 2011). Multi-scale reasoning over various analysis window sizes can mitigate this to some extent (Korus and Huang 2017). Favorable results have also been reported for approaches that incorporate image segmentation (Korus and Huang 2017; Zhang et al. 2019; Lin and Li 2020) or guided filtering (Chierchia et al. 2014) to adaptively adjust analysis windows based on image characteristics.

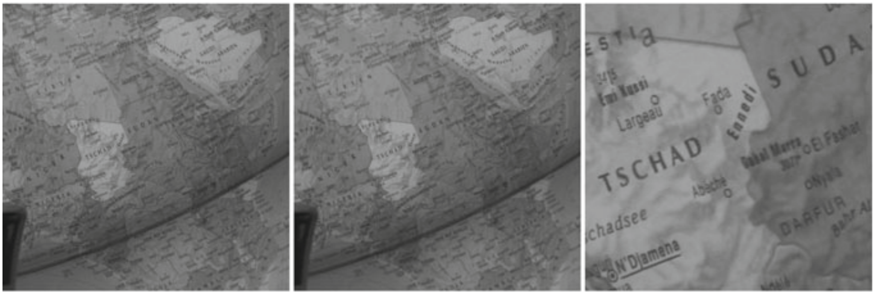
Surprisingly, very few notable updates to the seminal linear correlation predictor $\hat{\rho}$ from simple intensity, flatness, and texture features by Chen et al. (2008) have surfaced in the literature, despite its paramount role across virtually all PRNU-based localization approaches and a generally rather mixed performance (Quan and Li 2021). We highlight here the replacement of the original linear regression model with a feed-forward neural network by Korus and Huang (2017), and the observation that a more accurate prediction can be achieved when the camera's ISO speed is taken into account (Quan and Li 2021). Attempts to leverage a deep learning approach to obtain potentially more expressive features Chakraborty (2020) are commendable but require a more thorough evaluation for authoritative conclusions.

Overall, image manipulation localization based on camera sensor noise has its place in the broader universe of digital media forensics when there is a strong prior belief that the probe image indeed originates from a specific camera. If the manipulated region is sufficiently small, this can be established through conventional full-frame camera identification. Non-trivial sensor misalignment as discussed in Sect. 4.4 can be expected to complicate matters significantly for localization, but we are not aware of a principled examination of this aspect to date.

4.6 Counter-Forensics

The reliability and the robustness of camera identification based on sensor noise have been under scrutiny ever since seminal works on sensor noise forensics surfaced over 15 years ago. As a result, it is widely accepted that PRNU fingerprints survive a variety of common post-processing operations, including JPEG compression and resizing. Counter-forensics focuses on more deliberate attempts to impair successful camera identification by acknowledging that there are scenarios where intelligent actors make targeted efforts to induce a certain outcome of forensic analyses (Böhme and Kirchner 2013). In this context, it is instructive to distinguish between two major goals of countermeasures, fingerprint removal and fingerprint copying (Lukás et al. 2006; Gloe et al. 2007). The objective of fingerprint removal is the suppression of a camera's fingerprint to render source identification impossible. This can be desirable in efforts of protecting the anonymity of photographers, journalists, or legitimate whistleblowers in threatening environments (Nagaraja et al. 2011). Fingerprint copying attempts to make an image plausibly appear as if it was captured by a different camera, which is typically associated with nefarious motives. It strictly implies the suppression of the original fingerprint and it is thus generally a harder problem. The success of such counter-forensic techniques is to a large degree bound by the admissible visual quality of the resulting image. If an image purports to be camera-original but has suffered from noticeable degradation, it will likely raise suspicion. If anonymity is of utmost priority, strong measures that go along with a severe loss of image resolution are more likely acceptable.

Existing fingerprint removal methods can be categorized under two general approaches (Lukás et al. 2006). Methods of the first category are side-informed



(a) original image (b) modified image (c) close-up

Fig. 4.3 Fingerprint removal with PatchMatch replaces local image content with visually similar content from elsewhere in the image (Entrieri and Kirchner 2016). The PCE, computed considering all possible cross-correlation shifts, decreases from 5,617 for the original image to 32 after the modification. The “anonymized” image has a PSNR of 38.3 dB. Original image size: 2,000×2,000 pixels

in the sense that they use an estimate of the sensor noise fingerprint to ensure a detector output below the identification threshold. Flatfielding—a denoising technique that targets the general imaging model in Eq. (4.1)—is known to remove the multiplicative noise term **K** effectively, but it ideally requires access to the raw sensor measurements (Lukás et al. 2006; Gloe et al. 2007). Adaptive fingerprint removal techniques explicitly attempt to minimize Eq. (4.5) by finding a noise sequence that cancels out the multiplicative fingerprint term in Eq. (4.3) (Karaküçük and Dirik 2015; Zeng et al. 2015). This works best when exact knowledge of the detector (and thus the images used to estimate $\hat{\mathbf{K}}$) is available.

Uninformed techniques make less assumptions and directly address the robustness of the sensor noise fingerprint. Methods of this category apply post-processing to the image until the noise pattern is too corrupted to correlate with the fingerprint. No specific knowledge of the camera, the camera’s fingerprint, or the detector is assumed in this process. However, the high robustness of the sensor fingerprint makes this a non-trivial problem (Rosenfeld and Sencar 2009), and solutions may often come with a more immediate loss of image quality compared to side-informed methods. One promising direction is to induce irreversible sensor misalignment. Seam-carving—a form of content-adaptive resizing that shrinks images by removing low-energy “seams” (Avidan and Shamir 2007)—is an effective candidate operation in this regard (Bayram et al. 2013), although a considerable amount of seams must be removed to successfully desynchronize the sensor fingerprint (Dirik et al. 2014). This bears a high potential for the removal of “important” seams, degrading image quality and resolution. In many ways, recomposing the image entirely instead of removing a large number of seams is thus a more content-preserving alternative. The idea here is to replace local content with content from elsewhere in the image with the objective of finding replacements that are as similar to the original content as possible while lacking the telltale portion of the fingerprint. A modified version of the PatchMatch algorithm (Barnes et al. 2009) has been demonstrated to produce viable results (Entrieri and Kirchner 2016), while a later variant employed inpainting for this purpose (Mandelli et al. 2017). Figure 4.3 showcases an example of successful fingerprint removal with the PatchMatch algorithm. All three strategies, seam-carving, PatchMatch, and inpainting, can reliably prevent PRNU-based camera identification from a single image. An aggregation of fingerprint traces from multiple "anonymized" images from the same camera can reestablish the link to the common source device to some extent, however (Taspinar et al. 2017; Karaküçük and Dirik 2019).

In a fingerprint copy attack, a nefarious actor Eve operates with the goal of making an arbitrary image **J** look as if it was captured by an innocent user Alice's camera. Eve may obtain an estimate $\hat{\boldsymbol{K}}_E$ of Alice's camera fingerprint from a set of publicly available images and leverage the multiplicative nature of the PRNU to obtain (Lukás et al. 2006)

$$\mathbf{J}' = \mathbf{J}(1 + \alpha\hat{\boldsymbol{K}}_E)\,, \tag{4.12}$$

where the scalar factor $\alpha > 0$ determines the fingerprint strength. Attacks of this type have been demonstrated to be effective, in the sense that they can successfully mislead a camera identification algorithm in the form of Eq. (4.5). The attack's success generally depends on a good choice of α: too low values mean that the bogus image $\mathbf{J}'$ may not be assigned to Alice's camera; a too strong embedding will make the image appear suspicious (Goljan et al. 2011; Marra et al. 2014). In practical scenarios, Eve may have to apply further processing to make her forgery more compelling, e.g., removing the genuine camera fingerprint, synthesizing color filter interpolation artifacts (Kirchner and Böhme 2009), and removing or adding traces of JPEG compression (Stamm and Liu 2011).

Under realistic assumptions, it is virtually impossible to prevent Eve from forcing a high similarity score in Eq. (4.5). All is not lost, however. Alice can utilize that noise residuals computed with practical denoising filters are prone to contain remnants of image content. The key observation here is that the similarity between a noise residual $\mathbf{W}_\mathbf{I}$ from an image **I** taken with Alice's camera and the noise residual $\mathbf{W}_{\mathbf{J}'}$ due to a common attack-induced PRNU term will be further increased by some shared residual image content, if **I** contributed to Eve's fingerprint estimate $\hat{\mathbf{K}}_E$. The so-called triangle test (Goljan et al. 2011) picks up on this observation by also considering the correlation between Alice's own fingerprint and both $\mathbf{W}_\mathbf{I}$ and $\mathbf{W}_{\mathbf{J}'}$ to determine whether the similarity between $\mathbf{W}_\mathbf{I}$ with $\mathbf{W}_{\mathbf{J}'}$ is suspiciously large. A pooled version of the test establishes whether any images in a given set of Alice's images have contributed to $\hat{\mathbf{K}}_E$ (Barni et al. 2018), without determining which. Observe that in either case Alice may have to examine the entirety of images ever made public by her. On Eve's side, efforts to create a fingerprint estimate $\hat{\mathbf{K}}_E$ in a procedure that deliberately suppresses telltale remnants of image content can thwart the triangle test's success (Caldelli et al. 2011; Rao et al. 2013; Barni et al. 2018), thereby only setting the stage for the next iteration in the cat-and-mouse game between attacks and defenses.

The potentially high computational (and logistical) burden and the security concerns around the triangle test can be evaded in a more constrained scenario. Specifically, assume that Eve targets the forgery of an uncompressed image but only has access to images shared in a lossy compression format when estimating $\hat{\mathbf{K}}_E$. Here, it can be sufficient to test for the portion of the camera fingerprint that is fragile to lossy compression to establish that Eve's image $\mathbf{J}'$ does not contain a complete fingerprint (Quiring and Kirchner 2015). This works because the PRNU in uncompressed images is relatively uniform across the full frequency spectrum, whereas lossy compression mainly removes high-frequency information from images. As long as Alice's public images underwent moderately strong compression, such as JPEG at a quality factor of about 85, no practicable remedies for Eve to recover the critically missing portion of her fingerprint estimate are known at the time of this writing (Quiring et al. 2019).

4.7 Camera Fingerprints and Deep Learning

The previous sections have hopefully given the reader the impression that research around PRNU-based camera fingerprints is very much alive and thriving, and that new (and old) challenges continue to spark the imagination of academics and practitioners alike. Different from the broader domain of media forensics, which is now routinely drawing on deep learning solutions, only a handful of works have made attempts to apply ideas from this rapidly evolving field to the set of problems typically discussed in the context of device-specific (PRNU-based) camera fingerprints. We can only surmise that this in part due to the robust theoretical foundations that have defined the field and that have ultimately led to the wide acceptance of PRNU-based camera fingerprints in practical forensic casework, law enforcement, and beyond. Data-driven "black-box" solutions may thus appear superfluous to many.

However, one of the strengths that deep learning techniques can bring to the rigid framework of PRNU-based camera identification is in fact their very nature: they are data-driven. The detector in Sect. 4.3 was originally derived under a specific set of assumptions with respect to the imaging model in Eq. (4.1) and the noise residuals in Eq. (4.3). There is good reason to assume that real images will deviate from these simplified models to some degree. First and foremost, noise residuals from a single image will always suffer from significant and non-trivial distortion, if we accept that content suppression is an ill-posed problem in the absence of viable image models (and possibly even of the noise characteristics itself (Masciopinto and Pérez-González 2018)). This opens the door to potential improvements from data-driven approaches, which ultimately do not care about modeling assumptions but rather learn (hopefully) relevant insights from the training data directly.

One such approach specifically focuses on the extraction of a camera signature from a single image $\mathbf{I}$ at test time (Kirchner and Johnson 2019). Instead of relying on a "blind" denoising procedure as it is conventionally the case, a convolutional neural network (CNN) can serve as a flexible non-linear optimization tool that learns how to obtain a better approximation of $\mathbf{K}$. Specifically, the network is trained to extract a noise pattern $\tilde{\mathbf{K}}$ to minimize $\|\hat{\mathbf{K}} - \tilde{\mathbf{K}}\|_2^2$, as the pre-computed estimate $\hat{\mathbf{K}}$ is the best available approximation of the actual PRNU signal under the given imaging

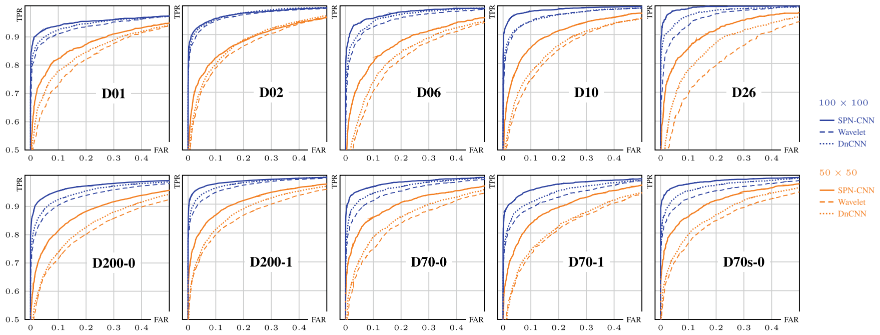


Fig. 4.4 Camera identification from noise signals computed with a deep-learning based fingerprint extractor (coined "SPN-CNN") (Kirchner and Johnson 2019). The ROC curves were obtained by thresholding the normalized correlation, $\rho^{(\mathbf{0})}_{ncc}$, between the extracted SPN-CNN noise patterns, $\tilde{\mathbf{K}}$, and "conventional" camera fingerprints, $\hat{\mathbf{K}}$, for patches of size 100 × 100 (blue) and 50 × 50 (orange). Device labels correspond to devices in the VISION (Shullani et al. 2017) (top) and Dresden Image Databases (Gloe and Böhme 2010) (bottom). Curves for the standard Wavelet denoiser and an off-the-shelf DnCNN denoiser (Zhang et al. 2017) are included for comparison

model. The trained network replaces the denoiser $F(\cdot)$ at test time, and the fingerprint similarity is evaluated directly for $\hat{\mathbf{K}}$ instead of $\hat{\mathbf{K}}\mathbf{I}$ in Eq. (4.5). Notably, this breaks with the tradition of employing the very same denoiser for both fingerprint estimation and detection. Empirical results suggest that the resulting noise signals have clear benefits over conventional noise residuals when used for camera identification, as showcased in Fig. 4.4. A drawback of this approach is that it calls for extraction CNNs trained separately for each relevant camera (Kirchner and Johnson 2019).

The similarity function employed by the detector in Eq. (4.5) is another target for potential improvement. If model assumptions do not hold, the normalized cross-correlation may no longer be a good approximation of the optimal detector (Fridrich 2013), and a data-driven approach may be able to reach more conclusive decisions. Along those lines, a Siamese network structure trained to compare spatially aligned patches from a fingerprint estimate $\hat{\mathbf{K}}$ and the noise residual $\mathbf{W}$ from a probe image $\mathbf{I}$ was reported to greatly outperform a conventional PCE-based detector across a range of image sizes (Mandelli et al. 2020). Different from the noise extraction approach by Kirchner and Johnson (2019), experimental results suggest that no device-specific training is necessary. Both approaches have not yet been subjected to more realistic settings that include sensor misalignment.

While the two works discussed above focus on specific building blocks in the established camera identification pipeline, we are not currently aware of an end-to-end deep learning solution that achieves source attribution at the level of individual devices at scale. There is early evidence, however, to suggest that camera model traces derived from a deep model can be a beneficial addition to conventional camera fingerprints, especially when the size of the analyzed images is small (Cozzolino et al. 2020).

CNNs have also been utilized in the context of counter-forensics, where the twofold objective of minimizing fingerprint similarity while maximizing image quality almost naturally invite data-driven optimization solutions to the problem of side-informed fingerprint removal. An early proposal trains an auto-encoder inspired anonymization operator by encoding the stated objectives in a two-part cost function (Bonettini et al. 2018). The solution, which has to be retrained for each new image, relies on a learnable denoising filter as part of the network, which is used to extract noise residuals from the "anonymized" images during training. The trained network is highly effective at test time as long as the detector uses the same denoising function, but performance dwindles when a different denoising filter, such as the standard Wavelet-based approach, is used instead. This limits the practical applicability of the fingerprint removal technique for the time being.

Overall, it seems almost inevitable that deep learning will take on a more prominent role across a wide range of problems in the field of (PRNU-based) device identification. We have included these first early works here mainly to highlight the trend, and we invite the reader to view them as important stepping stones for future developments to come.

4.8 Public Datasets

Various publicly available datasets have been compiled over time to advance research on sensor-based device identification, see Table 4.1. Not surprisingly, the evolution of datasets since the release of the trailblazing Dresden Image Database (Gloe and Böhme 2010) mirrors the general consumer shift from dedicated digital cameras to mobile devices. There are now also several diverse video datasets with a good coverage across different devices available. In general, a defining quality of a good dataset for source device identification is not only the number of available images per unique device, but also whether multiple instances of the same device model are represented. This is crucial for studying the effects of model-specific artifacts on false alarms, which may remain unidentified when only one device per camera model is present. Other factors worth considering may be whether the image/video capturing protocol controlled for certain exposure parameters, or whether all devices were used to capture the same set (or type) of scenes.

Although it has become ever more easy to gather suitably large amounts of data straight from some of the popular online media sharing platforms, dedicated custom-made research datasets offer the benefit of a well-documented provenance while fostering the reproducibility of research and mitigating copyright concerns. In addition, many of these datasets include, by design, a set of flatfield images to facilitate the computation of high-quality sensor fingerprints. However, as we have seen throughout this chapter, a common challenge in this domain is to keep pace with the latest technological developments on the side of camera manufacturers. This translates into a continued need for updated datasets to maintain practical relevance and timeliness. Results obtained on an older dataset may not hold up well on data from newer

Table 4.1 Public datasets with a special focus on source device identification. The table lists the number of available images and videos per dataset, the number of unique devices and camera models covered, as well as the types of cameras included (C: consumer, D: DSLR, M: mobile device)

Dataset	Year	Images	Videos	Devices	Models	Cameras
Dresden Gloe and Böhme (2010)*	2010	14k+		73	25	C D
RAISE Dang-Nguyen et al. (2015)	2015	8k+		3	3	D
VISION Shullani et al. (2017)**	2017	11k+	0.6k	35	30	M
HDR Shaya et al. (2018)	2018	5k+		23	22	M
SOCRatES Galdi et al. (2019)	2019	9k+	1k	103	60	M
video-ACID Hosler et al. (2019)	2019		12k+	46	36	C D M
NYUAD-MMD Taspinar et al. (2020)	2020	6k+	0.3k	78	62	M
Daxing Tian et al. (2019)	2020	43k+	1k+	90	22	M
Warwick Quan et al. (2020)*	2020	58k+		14	11	C D
Forchheim Hadwiger and Riess (2020)**	2020	3k+		27	25	M

*Includes multiple images of the same scene with varying exposure settings
**Provides auxiliary data from sharing the base data on various social network sites

devices due to novel sensor features or acquisition pipelines. A good example is the recent release of datasets with a specific focus on high dynamic range (HDR) imaging (Shaya et al. 2018; Quan et al. 2020), or the provision of annotations for videos that underwent electronic image stabilization (Shullani et al. 2017). With the vast majority of media now shared in significantly reduced resolution through online social network platforms or messaging apps, some of the more recent datasets also consider such common modern-day post-processing operations explicitly (Shullani et al. 2017; Hadwiger and Riess 2020).

4.9 Concluding Remarks

Camera-specific sensor noise fingerprints are a pillar of media forensics, and they are unrivaled when it comes to establishing source device attribution. While our focus in this chapter has been on still camera images (video data is covered in Chap. 5 of this book), virtually all imaging sensors introduce the same kind of noise fingerprints as we discussed here. For example, line sensors in flatbed scanners received early attention (Gloe et al. 2007; Khanna et al. 2007), and recent years have also seen biometric sensors move into the focus (Bartlow et al. 2009; Kauba et al. 2017; Ivanov and Baras 2017, 2019, i. a.).

Although keeping track of advances by device manufacturers and novel imaging pipelines is crucial for maintaining this status, an active research community has so far always been able to adapt to new challenges. The field has come a long way over the past 15 years, and new developments such as the cautious cross-over into the world of deep learning promise a continued potential for fruitful exploration. New perspectives and insights may also arise from applications outside the realm of media forensics. For example, with smartphones as ubiquitous companions in our everyday life, proposals to utilize camera fingerprints as building blocks to multi-factor authentication let users actively provide their device's sensor fingerprint via a captured image to be granted access to a web server (Valsesia et al. 2017; Quiring et al. 2019; Maier et al. 2020, i. a.). This poses a whole range of interesting practical challenges on its own, but it also invites a broader discussion about what camera fingerprints mean to an image-saturated world. At the minimum, concerns over image anonymity must be taken seriously in situations that call for it, and so we also see counter-forensics as part of the bigger picture unequivocally.

Acknowledgements Parts of this work were supported by AFRL and DARPA under Contract No. FA8750-16-C-0166. Any findings and conclusions or recommendations expressed in this material are solely the responsibility of the authors and do not necessarily represent the official views of AFRL, DARPA, or the U.S. Government.

References

Al-Ani M, Khelifi F (2017) On the SPN estimation in image forensics: a systematic empirical evaluation. IEEE Trans Inf Forensics Secur 12(5):1067–1081

Altinisik E, Sencar HT (2021) Source camera verification for strongly stabilized videos. IEEE Trans Inf Forensics Secur 16:643–657

Altinisik E, Tasdemir K, Sencar HT (2020) Mitigation of H.264 and H.265 video compression for reliable PRNU estimation. IEEE Trans Inf Forensics Secur 15:1557–1571

Amerini I, Caldelli R, Cappellini V, Picchioni F, Piva A (2009) Analysis of denoising filters for photo response non uniformity noise extraction in source camera identification. In: 16th international conference on digital signal processing, DSP 2009, Santorini, Greece, July 5-7, 2009. IEEE, pp 1–7

Amerini I, Caldelli R, Crescenzi P, del Mastio A, Marino A (2014) Blind image clustering based on the normalized cuts criterion for camera identification. Signal Process Image Commun 29(8):831–843

Amerini I, Caldelli R, Del Mastio A, Di Fuccia A, Molinari C, Rizzo AP (2017) Dealing with video source identification in social networks. Signal Process Image Commun 57:1–7

Artusi A, Richter T, Ebrahimi T, Mantiuk RK (2017) High dynamic range imaging technology [lecture notes]. IEEE Signal Process Mag 34(5):165–172

Avidan S, Shamir A (2007) Seam carving for content-aware image resizing. ACM Trans Graph 26(3):10

Baracchi D, Iuliani M, Nencini AG, Piva A (2020) Facing image source attribution on iphone X. In: Zhao X, Shi Y-Q, Piva A, Kim HJ (eds) Digital forensics and watermarking - 19th international workshop, IWDW 2020, Melbourne, VIC, Australia, November 25–27, 2020, Revised Selected Papers, vol 12617. Lecture notes in computer science. Springer, pp 196–207

Barni M, Nakano-Miyatake M, Santoyo-Garcia H, Tondi B (2018) Countering the pooled triangle test for prnu-based camera identification. In: 2018 IEEE international workshop on information forensics and security, WIFS 2018, Hong Kong, China, December 11-13, 2018. IEEE, pp 1–8

Barni M, Santoyo-Garcia H, Tondi B (2018) An improved statistic for the pooled triangle test against prnu-copy attack. IEEE Signal Process Lett 25(10):1435–1439

Bartlow N, Kalka ND, Cukic B, Ross A (2009) Identifying sensors from fingerprint images. In: IEEE conference on computer vision and pattern recognition, CVPR workshops 2009, Miami, FL, USA, 20–25 June 2009. IEEE Computer Society, pp 78–84

Bayram S, Sencar HT, Memon ND (2013) Seam-carving based anonymization against image & video source attribution. In: 15th IEEE international workshop on multimedia signal processing, MMSP 2013, Pula, Sardinia, Italy, September 30 - Oct. 2, 2013. IEEE, pp 272–277

Bayram S, Sencar HT, Memon ND (2012) Efficient sensor fingerprint matching through fingerprint binarization. IEEE Trans Inf Forensics Secur 7(4):1404–1413

Bayram S, Sencar HT, Memon ND (2015) Sensor fingerprint identification through composite fingerprints and group testing. IEEE Trans Inf Forensics Secur 10(3):597–612

Bloy Greg J (2008) Blind camera fingerprinting and image clustering. IEEE Trans Pattern Anal Mach Intell 30(3):532–534

Böhme R, Kirchner M (2013) Counter-forensics: attacking image forensics. In: Sencar HT, Memon N (eds) Digital image forensics: there is more to a picture than meets the eye. Springer, pp 327–366

Bonettini N, Bondi L, Mandelli S, Bestagini P, Tubaro S, Guera D (2018) Fooling PRNU-based detectors through convolutional neural networks. In: 26th European signal processing conference, EUSIPCO 2018, Roma, Italy, September 3-7, 2018. IEEE, pp 957–961

Caldelli R, Amerini I, Novi A (2011) An analysis on attacker actions in fingerprint-copy attack in source camera identification. In: 2011 ieee international workshop on information forensics and security, WIFS 2011, Iguacu Falls, Brazil, November 29 - December 2, 2011. IEEE Computer Society, pp 1–6

Chakraborty S (2020) A CNN-based correlation predictor for prnu-based image manipulation localization. In: Alattar AM, Memon ND, Sharma G (eds) Media watermarking, security, and forensics 2020, Burlingame, CA, USA, 27-29 January 2020

Chakraborty S, Kirchner M (2017) PRNU-based image manipulation localization with discriminative random fields. In: Alattar AM, Memon ND (eds) Media Watermarking, Security, and Forensics 2017, Burlingame, CA, USA, 29 January 2017 - 2 February 2017, pages 113–120. Ingenta, 2017

Chen M, Fridrich JJ, Goljan M, Lukáš J (2008) Determining image origin and integrity using sensor noise. IEEE Trans Inf Forensics Secur 3(1):74–90

Chierchia G, Cozzolino D, Poggi G, Sansone C, Verdoliva L (2014) Guided filtering for PRNU-based localization of small-size image forgeries. In: IEEE international conference on acoustics, speech and signal processing, ICASSP 2014, Florence, Italy, May 4-9, 2014. IEEE, pp 6231–6235

Chierchia G, Parrilli S, Poggi G, Verdoliva L, Sansone C (2011) PRNU-based detection of small-size image forgeries. In: 17th international conference on digital signal processing, DSP 2011, Corfu, Greece, July 6-8, 2011. IEEE, pp 1–6

Chierchia G, Poggi G, Sansone C, Verdoliva L (2014) A Bayesian-MRF approach for PRNU-based image forgery detection. IEEE Trans Inf Forensics Secur 9(4):554–567

Chuang W-H, Su H, Wu M (2011) Exploring compression effects for improved source camera identification using strongly compressed video. In: Macq B, Schelkens P (eds) 18th IEEE international conference on image processing, ICIP 2011, Brussels, Belgium, September 11-14, 2011. IEEE, pp 1953–1956

Connelly B, Eli S, Adam F, Goldman Dan B (2009) Patchmatch: a randomized correspondence algorithm for structural image editing. ACM Trans Graph 28(3):24

Cortiana A, Conotter V, Boato G, De Natale FGB (2011) Performance comparison of denoising filters for source camera identification. In: Memon ND, Dittmann J, Alattar AM, Delp EJ III (eds) Media forensics and security III, San Francisco Airport, CA, USA, January 24-26, 2011, Proceedings, SPIE Proceedings, vol 7880. SPIE, p 788007

Cozzolino D, Marra F, Gragnaniello D, Poggi G, Verdoliva L (2020) Combining PRNU and noiseprint for robust and efficient device source identification. EURASIP J Inf Secur 2020:1

Dang-Nguyen D-T, Pasquini C, Conotter V, Boato G (2015) RAISE: a raw images dataset for digital image forensics. In: Ooi WT, Feng W-c, Liu F (eds), Proceedings of the 6th ACM multimedia systems conference, MMSys 2015, Portland, OR, USA, March 18-20, 2015. ACM, pp 219–224

Dirik AE, Sencar HT, Memon ND (2014) Analysis of seam-carving-based anonymization of images against PRNU noise pattern-based source attribution. IEEE Trans Inf Forensics Secur 9(12):2277–2290

Entrieri J, Kirchner M (2016) Patch-based desynchronization of digital camera sensor fingerprints. In: Alattar AM, Memon ND (eds) Media watermarking, security, and forensics 2016, San Francisco, California, USA, February 14-18, 2016. Ingenta, pp 1–9

Fridrich J (2013) Sensor defects in digital image forensic. In: Sencar HT, Memon N (eds) Digital Image forensics: there is more to a picture than meets the eye. Springer, Berlin, pp 179–218

Galdi C, Hartung F, Dugelay J-L (2019) SOCRatES: A database of realistic data for source camera recognition on smartphones. In: De Marsico M, di Baja GS, Fred ALN (eds) Proceedings of the 8th international conference on pattern recognition applications and methods, ICPRAM 2019, Prague, Czech Republic, February 19-21, 2019. SciTePress, pp 648–655

Gloe T, Böhme R (2010) The Dresden image database for benchmarking digital image forensics. J Digit Forensic Pract 3(2–4):150–159

Gloe T, Franz E, Winkler A (2007) Forensics for flatbed scanners. In: Delp EJ III, Wong PW (eds.) Security, steganography, and watermarking of multimedia contents IX, San Jose, CA, USA, January 28, 2007, SPIE Proceedings, vol 6505. SPIE, p. 65051I

Gloe T, Kirchner M, Winkler A, Böhme R (2007) Can we trust digital image forensics? In: Lienhart R, Prasad AR, Hanjalic A, Choi S, Bailey BP, Sebe N (eds) Proceedings of the 15th international conference on multimedia 2007, Augsburg, Germany, September 24-29, 2007. ACM, pp 78–86

Gloe T, Pfennig S, Kirchner M (2012) Unexpected artefacts in PRNU-based camera identification: a 'Dresden image database' case-study. In: Li C-T, Dittmann J, Katzenbeisser S, Craver S (eds) Multimedia and security workshop, MM&Sec 2012, Coventry, United Kingdom, September 6-7, 2012. ACM, pp 109–114

Goljan M, Chen M, Comesaña P, Fridrich JJ (2016) Effect of compression on sensor-fingerprint based camera identification. In: Alattar AM, Memon ND (eds) Media watermarking, security, and forensics 2016, San Francisco, California, USA, February 14-18, 2016. Ingenta, pp 1–10

Goljan M, Chen M, Fridrich JJ (2007) Identifying common source digital camera from image pairs. In: Proceedings of the international conference on image processing, ICIP 2007, September 16-19, 2007, San Antonio, Texas, USA. IEEE, pp 125–128

Goljan M, Fridrich JJ, Filler T (2009) Large scale test of sensor fingerprint camera identification. In: Delp EJ, Dittmann J, Memon ND, Wong PW (eds) Media forensics and security I, part of the IS&T-SPIE electronic imaging symposium, San Jose, CA, USA, January 19-21, 2009, Proceedings, SPIE Proceedings, vol 7254. SPIE, p. 72540I

Goljan M, Fridrich JJ, Filler T (2010) Managing a large database of camera fingerprints. In: Memon ND, Dittmann J, Alattar AM, Delp EJ (eds) Media forensics and security II, part of the IS&T-SPIE electronic imaging symposium, San Jose, CA, USA, January 18-20, 2010, Proceedings, SPIE Proceedings, vol 7541. SPIE, pp 754108

Goljan M, Fridrich JJ (2008) Camera identification from cropped and scaled images. In: Delp EJ III, Wong PW, Dittmann J, Memon ND (eds) Security, forensics, steganography, and watermarking of multimedia contents X, San Jose, CA, USA, January 27, 2008, SPIE Proceedings, vol 6819. SPIE, pp 68190E

Goljan M, Fridrich JJ (2012) Sensor-fingerprint based identification of images corrected for lens distortion. In: Memon ND, Alattar AM, Delp EJ III (eds) Media watermarking, security, and forensics 2012, Burlingame, CA, USA, January 22, 2012, Proceedings, SPIE Proceedings, vol 8303. SPIE, p 83030H

Goljan M, Fridrich JJ (2014) Estimation of lens distortion correction from single images. In: Alattar AM, Memon ND, Heitzenrater C (eds) Media watermarking, security, and forensics 2014, San Francisco, CA, USA, February 2, 2014, Proceedings, SPIE Proceedings, vol 9028. SPIE, p 90280N

Goljan M, Fridrich JJ, Chen M (2011) Defending against fingerprint-copy attack in sensor-based camera identification. IEEE Trans Inf Forensics Secur 6(1):227–236

Hadwiger B, Riess C (2021) The Forchheim image database for camera identification in the wild. In: Del Bimbo A, Cucchiara R, Sclaroff S, Farinella GM, Mei T, Bertini M, Escalante HJ, Vezzani R (eds) Pattern recognition. ICPR international workshops and challenges - virtual event, January 10-15, 2021, Proceedings, Part VI, Lecture Notes in Computer Science, vol 12666. Springer, pp 500–515

Hosler BC, Zhao X, Mayer O, Chen C, Shackleford JA, Stamm MC (2019) The video authentication and camera identification database: a new database for video forensics. IEEE Access 7:76937–76948

Hosseini MDM, Goljan M (2019) Camera identification from HDR images. In: Cogranne R, Verdoliva L, Lyu S, Troncoso-Pastoriza JR, Zhang X (eds) Proceedings of the ACM workshop on information hiding and multimedia security, IH&MMSec 2019, Paris, France, July 3-5, 2019. ACM, pp 69–76

Iuliani M, Fontani M, Piva A (2019) Hybrid reference-based video source identification. Sensors 19(3):649

Iuliani M, Fontani M, Piva A (2021) A leak in PRNU based source identification - questioning fingerprint uniqueness. IEEE Access 9:52455–52463

Ivanov VI, Baras JS (2017) Authentication of swipe fingerprint scanners. IEEE Trans Inf Forensics Secur 12(9):2212–2226

Ivanov VI, Baras JS (2019) Authentication of area fingerprint scanners. Pattern Recognit 94:230–249

Joshi S, Korus P, Khanna N, Memon ND (2020) Empirical evaluation of PRNU fingerprint variation for mismatched imaging pipelines. In: 12th IEEE international workshop on information forensics and security, WIFS 2020, New York City, NY, USA, December 6-11, 2020. IEEE, pp 1–6

Kang X, Li Y, Qu Z, Huang J (2012) Enhancing source camera identification performance with a camera reference phase sensor pattern noise. IEEE Trans Inf Forensics Secur 7(2):393–402

Karaküçük A, Dirik AE (2015) Adaptive photo-response non-uniformity noise removal against image source attribution. Digit Investigat 12:66–76

Karaküçük A, Dirik AE (2019) PRNU based source camera attribution for image sets anonymized with patch-match algorithm. Digit Investig 30:43–51

Kauba C, Debiasi L, Uhl A (2017) Identifying the origin of iris images based on fusion of local image descriptors and PRNU based techniques. In: 2017 IEEE international joint conference on biometrics, IJCB 2017, Denver, CO, USA, October 1-4, 2017. IEEE, pp 294–301

Khanna N, Mikkilineni AK, Chiu GT-C, Allebach JP, Delp EJ (2007) Scanner identification using sensor pattern noise. In: Delp EJ III, Wong PW (eds) Security, steganography, and watermarking of multimedia contents IX, San Jose, CA, USA, January 28, 2007. SPIE Proceedings, vol 6505, p 65051K. SPIE

Kirchner M, Böhme R (2009) Synthesis of color filter array pattern in digital images. In: Delp EJ, Dittmann J, Memon ND, Wong PW (eds) Media forensics and security I, part of the IS&T-SPIE electronic imaging symposium, San Jose, CA, USA, January 19-21, 2009, Proceedings, SPIE Proceedings, vol 7254. SPIE, p 72540K

Kirchner M, Johnson C (2019) SPN-CNN: boosting sensor-based source camera attribution with deep learning. In: IEEE international workshop on information forensics and security, WIFS 2019, Delft, The Netherlands, December 9-12, 2019. IEEE, pp 1–6

Knight S, Moschou S, Sorell M (2009) Analysis of sensor photo response non-uniformity in RAW images. In: Sorell M (ed) Forensics in telecommunications, information and multimedia, second international conference, e-forensics 2009, Adelaide, Australia, January 19–21, 2009, Revised Selected Papers, vol 8. Lecture Notes of the Institute for Computer Sciences. Springer, Social Informatics and Telecommunications Engineering, pp 130–141

Korus P, Huang J (2017) Multi-scale analysis strategies in prnu-based tampering localization. IEEE Trans Inf Forensics Secur 12(4):809–824

Li C-T (2010) Source camera identification using enhanced sensor pattern noise. IEEE Trans Inf Forensics Secur 5(2):280–287

Li C-T (2010) Unsupervised classification of digital images using enhanced sensor pattern noise. International symposium on circuits and systems (ISCAS 2010), May 30 - June 2, 2010. France. IEEE, Paris, pp 3429–3432

Lin X, Li C-T (2016) Preprocessing reference sensor pattern noise via spectrum equalization. IEEE Trans Inf Forensics Secur 11(1):126–140

Lin X, Li C-T (2020) PRNU-based content forgery localization augmented with image segmentation. IEEE Access 8:222645–222659

Lukás J, Fridrich JJ, Goljan M (2005) Determining digital image origin using sensor imperfections. In: Said A, Apostolopoulos JG (eds) Electronic Imaging: Image and Video Communications and Processing 2005, San Jose, California, USA, 16–20 January 2005, SPIE Proceedings, vol 5685. SPIE

Lukáš J, Fridrich JJ, Goljan M (2006) Digital camera identification from sensor pattern noise. IEEE Trans Inf Forensics Secur 1(2):205–214

Maier D, Erb H, Mullan P, Haupert V (2020) Camera fingerprinting authentication revisited. In: 23rd international symposium on research in attacks, intrusions and defenses ({RAID} 2020), pp 31–46

Mandelli S, Bestagini P, Verdoliva L, Tubaro S (2020) Facing device attribution problem for stabilized video sequences. IEEE Trans Inf Forensics Secur 15:14–27

Mandelli S, Bondi L, Lameri S, Lipari V, Bestagini P, Tubaro S (2017) Inpainting-based camera anonymization. In: 2017 IEEE international conference on image processing, ICIP 2017, Beijing, China, September 17-20, 2017. IEEE, pp 1522–1526

Mandelli S, Cozzolino D, Bestagini P, Verdoliva L, Tubaro S (2020) CNN-based fast source device identification. IEEE Signal Process Lett 27:1285–1289

Marra F, Poggi G, Sansone C, Verdoliva L (2017) Blind PRNU-based image clustering for source identification. IEEE Trans Inf Forensics Secur 12(9):2197–2211

Marra F, Roli F, Cozzolino D, Sansone C, Verdoliva L (2014) Attacking the triangle test in sensor-based camera identification. In: 2014 IEEE international conference on image processing, ICIP 2014, Paris, France, October 27-30, 2014. IEEE, pp 5307–5311

Masciopinto M, Pérez-González F (2018) Putting the PRNU model in reverse gear: findings with synthetic signals. In: 26th European signal processing conference, EUSIPCO 2018, Roma, Italy, September 3-7, 2018. IEEE, pp 1352–1356

Meij C, Geradts Z (2018) Source camera identification using photo response non-uniformity on WhatsApp. Digit Investig 24:142–154

Nagaraja S, Schaffer P, Aouada D (2011) Who clicks there!: anonymising the photographer in a camera saturated society. In: Chen Y, Vaidya J (eds) Proceedings of the 10th annual ACM workshop on Privacy in the electronic society, WPES 2011, Chicago, IL, USA, October 17, 2011. ACM, pp 13–22

Quan Y, Li C-T, Zhou Y, Li L (2020) Warwick image forensics dataset for device fingerprinting in multimedia forensics. In: IEEE international conference on multimedia and expo, ICME 2020, London, UK, July 6-10, 2020. IEEE, pp 1–6

Quan Y, Li C-T (2021) On addressing the impact of ISO speed upon PRNU and forgery detection. IEEE Trans Inf Forensics Secur 16:190–202

Quiring E, Kirchner M (2015) Fragile sensor fingerprint camera identification. In: 2015 IEEE international workshop on information forensics and security, WIFS 2015, Roma, Italy, November 16-19, 2015. IEEE, pp 1–6

Quiring E, Kirchner M, Rieck K (2019) On the security and applicability of fragile camera fingerprints. In: Sako K, Schneider SA, Ryan PYA (eds) Computer security - ESORICS 2019–24th European symposium on research in computer security, Luxembourg, September 23–27, 2019, Proceedings, Part I, vol 11735. Lecture notes in computer science. Springer, pp 450–470

Rao Q, Li H, Luo W, Huang J (2013) Anti-forensics of the triangle test by random fingerprint-copy attack. In: Computational visual media conference

Rosenfeld K, Sencar HT (2009) A study of the robustness of PRNU-based camera identification. In: Delp EJ, Dittmann J, Memon ND, Wong PW (eds), Media forensics and security I, part of the IS&T-SPIE electronic imaging symposium, San Jose, CA, USA, January 19-21, 2009, Proceedings, SPIE Proceedings, vol 7254. SPIE, p 72540M

Shaya OA, Yang P, Ni R, Zhao Y, Piva A (2018) A new dataset for source identification of high dynamic range images. Sensors 18(11):3801

Shullani D, Fontani M, Iuliani M, Shaya OA, Piva A (2017) VISION: a video and image dataset for source identification. EURASIP J Inf Secur 2017:15

Stamm MC, Liu KJR (2011) Anti-forensics of digital image compression. IEEE Trans Inf Forensics Secur 6(3–2):1050–1065

Tandogan SE, Altinisik E, Sarimurat S, Sencar HT (2019) Tackling in-camera downsizing for reliable camera ID verification. In: Alattar AM, Memon ND, Sharma G (eds) Media watermarking, security, and forensics 2019, Burlingame, CA, USA, 13–17 January 2019. Ingenta

Taspinar S, Mohanty M, Memon ND (2016) Source camera attribution using stabilized video. In: IEEE international workshop on information forensics and security, WIFS 2016, Abu Dhabi, United Arab Emirates, December 4-7, 2016. IEEE, pp 1–6

Taspinar S, Mohanty M, Memon ND (2017) PRNU-based camera attribution from multiple seam-carved images. IEEE Trans Inf Forensics Secur 12(12):3065–3080

Taspinar S, Mohanty M, Memon ND (2020) Camera fingerprint extraction via spatial domain averaged frames. IEEE Trans Inf Forensics Secur 15:3270–3282

Taspinar S, Mohanty M, Memon ND (2020) Camera identification of multi-format devices. Pattern Recognit Lett 140:288–294

Tian H, Xiao Y, Cao G, Zhang Y, Xu Z, Zhao Y (2019) Daxing smartphone identification dataset. IEEE. Access 7:101046–101053

Valsesia D, Coluccia G, Bianchi T, Magli E (2015) Compressed fingerprint matching and camera identification via random projections. IEEE Trans Inf Forensics Secur 10(7):1472–1485

Valsesia D, Coluccia G, Bianchi T, Magli E (2015) Large-scale image retrieval based on compressed camera identification. IEEE Trans Multim 17(9):1439–1449

Valsesia D, Coluccia G, Bianchi T, Magli E (2017) User authentication via PRNU-based physical unclonable functions. IEEE Trans Inf Forensics Secur 12(8):1941–1956

van Houten W, Geradts ZJMH (2009) Source video camera identification for multiply compressed videos originating from youtube. Digit Investig 6(1–2):48–60

Zeng H, Chen J, Kang X, Zeng W (2015) Removing camera fingerprint to disguise photograph source. In: 2015 IEEE international conference on image processing, ICIP 2015, Quebec City, QC, Canada, September 27-30, 2015. IEEE, pp 1687–1691

Zhang W, Tang X, Yang Z, Niu S (2019) Multi-scale segmentation strategies in PRNU-based image tampering localization. Multim Tools Appl 78(14):20113–20132

Zhang K, Zuo W, Chen Y, Meng D, Zhang L (2017) Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising. IEEE Trans Image Process 26(7):3142–3155