A Bayesian-MRF Approach for PRNU-based Image Forgery Detection

Giovanni Chierchia, Giovanni Poggi, Carlo Sansone, *Member, IEEE*, and Luisa Verdoliva

***Abstract*—Graphics editing programs of the last generation provide ever more powerful tools which allow to retouch digital images leaving little or no traces of tampering. The reliable detection of image forgeries requires, therefore, a battery of complementary tools that exploit different image properties. Techniques based on the photo-response non-uniformity (PRNU) noise are among the most valuable such tools, since they do not detect the inserted object but rather the absence of the camera PRNU, a sort of camera fingerprint, dealing successfully with forgeries that elude most other detection strategies.**

In this work we propose a new approach to detect image forgeries using sensor pattern noise. Casting the problem in terms of Bayesian estimation, we use a suitable Markov random field prior to model the strong spatial dependencies of the source, and take decisions jointly on the whole image rather than individually for each pixel. Modern convex optimization techniques are then adopted to achieve a globally optimal solution and PRNU estimation is improved by resorting to nonlocal denoising. Large-scale experiments on simulated and real forgeries show that the proposed technique largely improves upon the current state of the art, and that it can be applied with success to a wide range of practical situations.

***Index Terms*—Image forgery detection, PRNU, sensor noise, Bayesian approach, Markov random fields.**

I. Introduction

Digital images are more and more frequently used to support important decisions. This is especially true in the forensic field where, to make just a few examples, images are routinely used to describe the scene of a crime, or to define responsibilities in road accidents. Unfortunately, with the wide availability of sophisticated image manipulation tools, modifying a digital photo, with little or no obvious signs of tampering, has become easier than ever before [1]. Therefore, it is important to devise tools that help deciding on the authenticity of a digital image, which raises attention on the image forgery detection field.

Several approaches have been proposed in the literature to detect image alterations under a variety of scenarios. A first category comprises active techniques, for image authentication, based on the use of watermarks [2] and signatures [3], [4]. In the first case, the watermark is embedded into the image (possibly originating small distortions), while in the latter, the signature is attached to the image as side information. Although these methods are very effective, they can be applied only when the digital source is protected at the origin, which is probably a minority of the cases of interest. Therefore, there has been a steadily growing interest on passive techniques which retrieve traces of manipulations from the image itself, with no need of collaboration on the part of the user.

Some techniques are specifically tailored to copy-move forgeries, where portions of the image are cut and pasted elsewhere in the same image to duplicate or hide objects of interest. Duplicated parts are discovered by block-based processing or, more efficiently, by means of suitable invariant features [5], [6], [7]. A more general approach considers physical inconsistencies, such as the lighting of objects, shadows, or geometric features (dimension, position, etc.) of objects w.r.t. the camera [8], [9], [10]. Also, as many images are saved in some compressed JPEG format, several forgery detection techniques rely on the traces left by multiple JPEG compression. In fact, when a JPEG image is modified and saved again in JPEG format, specific artifacts appear as a result of the multiple quantization processes, suggesting the presence of some forms of tampering [11], [12], [13], [14].

Another valuable source of information is the acquisition phase, which often leaves peculiar traces, related to lens characteristics [15], [16], the color filter array (CFA) pattern [17], [18], [19], or the sensor array [20], [21], that can be used to discover image manipulations. In this latter context, the photo-response non uniformity (PRNU) noise appears as one of the most promising tools at hand. The PRNU arises from tiny imperfections in the silicon wafer used to manufacture the imaging sensor [22]. These physical differences provide a unique sensor pattern, specific of each individual camera, constant in time and independent of the scene. It can be therefore considered as a sort of camera fingerprint and used as such to accomplish forgery detection or image identification tasks. Indeed, the most common forms of image forgery, like copy-move or splicing, delete the original camera PRNU from the target region, a fact that can be detected through suitable analyses, provided the camera PRNU is available. Note that, unlike with most other approaches, the detection of tampering is based on the *absence* of the fingerprint, hence does not depend on the specific type of forgery. On the other hand, the PRNU pattern is fairly robust to several common forms of image processing, such as JPEG compression, filtering, or gamma correction [20], [21].

In [20] a PRNU-based technique for camera identification and forgery detection has been proposed, then refined in [21]. Given the potential of this algorithm [23], many research

G.Chierchia is with Institut Mines-Télécom; Télécom ParisTech; CNRS LTCI, 75014 Paris, France. G.Poggi, C.Sansone and L.Verdoliva are with the DIETI, Università Federico II di Napoli, Naples, Italy.

E-mail: chierchi@telecom-paristech.fr, {giovanni.poggi, carlo.sansone, luisa.verdoliva}@unina.it.

groups have soon started working, trying to improve the estimate of the image PRNU, to define better decision statistics, or better decision strategies.

In fact, since the PRNU is a very weak signal, its reliable estimation is crucial for the algorithm success. Typically, the estimate is computed by subtracting a filtered version of the image from the observed one, obtaining a residual where the PRNU is present but the image (seen as noise in this context) is mostly removed. However, the residual contains also traces of the signal, especially at high frequencies, due to the imperfection of the filtering process or to in-camera processing, such as JPEG compression, CFA interpolation, or vignetting [24]. In [25] the denoising filter used in the original technique has been replaced with a state-of-the-art nonlocal filter, with significant performance improvements. In [26] a strategy to reduce the interference of scene details on the PRNU is proposed, based on selective attenuation of wavelet transform coefficients. A major challenge is the suppression of the so-called non-unique artifacts [27], specific to a camera model or manufacturer. These include, for example, JPEG block artifacts, and CFA interpolation artifacts, both characterized by regular "linear" spatially periodic patterns, relatively easy to correct [28]. Non-unique artifacts may lead to wrong results, especially in camera identification, because of the increased similarity between the PRNU fingerprints of a different devices with similar characteristics. Recently, non-linear artifacts have also been reported, due to correction of radial lens distortion [29] and other advanced in-camera processing procedures [30].

As for decision statistics, since the normalized correlation used in the original algorithm was very sensitive to artifacts a more stable statistics, the peak-to-correlation energy (PCE) ratio, is proposed in [27], [31] for camera identification purposes, further modified in [32] to lower the false positive rate. A different approach is followed in [33] where the adoption of canonical correlation analysis is proposed. In [34] it is proposed to compute statistics, and possibly take decisions, based on a prior segmentation of the image, thus moving towards an object-oriented processing. The original signal is used also in [35], where only regions characterized by higher signal quality are used, discarding those regions heavily deteriorated by irrelevant noise. Some papers, finally, focus on computational/storage complexity, certainly an issue for applications that involve large-scale databases, proposing the use of a quantized [36], or spatially limited [37], or hashed [38] PRNU.

In this work we propose a new PRNU-based forgery detection algorithm[1] which relies on the same general structure and basic tools as [21] but improves upon it under several respects: *i)* first of all, we abandon the constant false alarm rate decision strategy to adopt a more flexible Bayesian rule; *ii)* more important, decisions are now made jointly on the whole image rather than individually for each pixel; *iii)* to single out the global solution we resort to convex-optimization tools, which guarantee convergence to an optimum in a limited time; *iv)* the strong spatial dependencies of the source are taken into account by modeling the data through a suitable Markov random field; *v)* finally, the quality of the basic data is improved by using a nonlocal denoising algorithm. Experiments prove that the proposed algorithm outperforms the reference one, with a very limited increase in the computational burden.

In the following, after thoroughly revising the reference algorithm in Section II, in Section III we motivate and describe in detail the proposed improvements. In Section IV we analyze performance by means of simulation experiments. Section V draws conclusions and outlines future research.

II. Background

Let $y \in \mathbb{R}^N$ be a digital image, defined on a rectangular lattice Ω, with y_i the value at site $i \in \Omega$, observed at the camera output, either as a single color band or the composition of multiple color bands. Let us assume, in a simplified model [21], [22], that y can be written as

$$y = (1 + k)x + \theta \tag{1}$$

where x is the ideal noise-free image, k the camera PRNU, and θ an additive noise term which accounts for all types of disturbances. Products between images, unless otherwise stated, are pixel-wise. By rewriting (1) as

$$y = xk + x + \theta \tag{2}$$

we stress from the beginning that the PRNU, k, is the only signal of interest in all our analyses, and our goal is to decide whether or not it comes from the camera under test, in all or part of the image, so as to detect possible forgeries. All other terms assume the role of unwanted disturbance, including the ideal image x, which is in fact estimated and subtracted from the original image, obtaining a more tractable *noise residual*. Even so, Since the PRNU is typically very small, except for possible faulty sensors, we work in a very hostile environment, with a very low signal-to-noise (SNR) ratio.

In the following, we describe the image integrity verification procedure proposed in [21] comprising four basic steps:

A. estimation of the camera PRNU (off-line);
B. computation of image noise residual and of derived statistics;
C. sliding-window pixel-wise forgery detection test;
D. morphological processing of test result map.

A. PRNU estimation

The first step of the procedure is not a hard task because one is supposed to have either a large number of images taken by the camera of interest or the camera itself and hence any reasonable estimation technique will provide, by sheer brute force, a good result. The maximum likelihood estimate of the PRNU from M given images is computed in [21] as

$$\widehat{k} = \sum_{m=1}^{M} y_m r_m / \sum_{m=1}^{M} y_m^2 \tag{3}$$

where the weighting terms y_m account for the fact that dark areas of the image present an attenuated PRNU and hence should contribute less to the overall estimate. In the following, for the sake of simplicity, we will neglect the estimation error and will assume to know the camera PRNU perfectly, that is $\widehat{k} = k$.

[1] Some preliminary results are reported in [39].

B. Computation of noise residual

In the second step of the algorithm we estimate the ideal image x by means of a denoising filter f

$$\widehat{x} = f(y) \tag{4}$$

and compute the noise residual

$$r = y - \widehat{x} = yk + (x-y)k + (x-\widehat{x}) + \theta = yk + n \tag{5}$$

where, for convenience, k multiplies the observed image y rather than the unknown original x, and the small difference term $(x-y)k$ has been included with the denoising error $(x-\widehat{x})$ and other disturbances in a single noise term n.

Even in the best case, with perfect denoising, the noise term is likely to dominate r which, therefore, will be only weakly correlated with the camera PRNU. In the presence of textured areas, however, denoising is typically less accurate and some signal components leak into the residual. This event lowers, even dramatically, the operative SNR, and makes detection virtually impossible. Especially in these areas, the effectiveness of the denoising algorithm becomes crucial for the overall performance.

C. Forgery detection test

The detection problem can be formulated as a binary hypothesis test:

$$\begin{cases} H_0 & r_i = n_i \\ H_1 & r_i = z_i + n_i \end{cases} \tag{6}$$

with $z = yk$, between hypothesis H_0, that the camera PRNU is absent, hence the pixel has been tampered and hypothesis H_1, that it is present, hence the pixel is genuine[2]. The true and estimated pixel classes will be denoted by u_i and $\widehat{u}_i$, both defined in $\{0,1\}$.

The detection test is based on the normalized correlation index [21]

$$\rho_i = \mathrm{CORR}(r_{W_i}, z_{W_i}) \tag{7}$$

between r_{W_i} and z_{W_i}, the restrictions of r and z, respectively, to a window W_i centered on the target pixel. Pixel labeling is obtained by comparing the decision statistic with a threshold γ_1

$$\widehat{u}_i = \begin{cases} 0 & \rho_i < \gamma_1 \\ 1 & \text{otherwise} \end{cases} \tag{8}$$

and the threshold is selected with a Neyman-Pearson approach to obtain the desired false acceptance rate (FAR), that is, a suitably small probability that a tampered pixel be wrongly identified as genuine. The pdf of ρ under hypothesis H_0 is estimated by computing the correlation between the camera PRNU and a large amount of noise residuals coming from other cameras, and using standard density fitting techniques. Rather large square blocks must be considered in order to obtain a reliable statistic; [21] chooses a size of 128×128 pixels.

Even in the absence of forgery, the correlation might happen to be very low when the image is dark (since y multiplies the PRNU), saturated (because of intensity clipping), or when denoising does not perform well and some image content leaks into the noise residual. In [21] this problem is addresses by means of a "predictor" which, based on local images features, such as texture, flatness and intensity, computes the expected value $\widehat{\rho}_i$ of the correlation index under hypothesis H_1. When $\widehat{\rho}_i$ is too low, the pixel is labeled as genuine, the less risky decision, irrespective of the value of ρ_i. In this case, in fact, even for a genuine pixel one cannot expect a correlation index much larger than 0. Therefore, the test becomes

$$\widehat{u}_i = \begin{cases} 0 & \rho_i < \gamma_1 \ \text{AND} \ \widehat{\rho}_i > \gamma_2 \\ 1 & \text{otherwise} \end{cases} \tag{9}$$

The second threshold γ_2 is chosen heuristically by the user and separates, in practice, reliable regions from problematic ones. It is worth underlining that the refined decision test (9) can only reduce the false alarm rate but does not increase (actually it might reduce) the probability of detecting an actual forgery. In addition, the choice of the threshold itself is not obvious and can significantly impact on the performance.

D. Label map post-processing

Morphological filtering operates on the map output by the previous test, working on the regions of would-be forgeries over a background of genuine pixels. In [21] all regions smaller than 64×64 pixels, one fourth of the window size, are attributed to random errors and removed. Finally, the surviving regions are dilated with a structured element of radius 20 pixel to restore, approximately, the shape of the forged region, since many border points go lost because their correlation index is computed on mixed (forged/genuine) blocks.

III. Proposed algorithm

The algorithm proposed here relies on [21], using the same hypotheses (*e.g.*, known camera), basic approach, and tools, but differs profoundly from it in the formulation of the problem and, consequently, in the solution techniques. We follow a Bayesian approach which allows us to better balance the observed statistics and the prior knowledge on the image. As a consequence, we obtain an improved performance, that is, an increased likelihood to reveal small forgeries (hence, an higher spatial resolution) and a much lower probability of declaring forgery in genuine regions. In addition, as proposed in [25], we replace the wavelet-based filter used in [21] with a better nonlocal filter, which provides us with more reliable data for the subsequent decision phase.

A. Bayesian Formulation

Our goal is to find the label map $\widehat{u} \in \{0,1\}^N$ which has the maximum probability[3] to occur given the observed image

[2] Notice that, since we focus on the detection of forgeries, denoted by the *absence* of the PRNU, the role of two two hypotheses is inverted w.r.t. what is usual. For example we will talk of False Alarm when H_0 is accepted but H_1 holds (a forged pixel is declared genuine), and of Missing Detection when H_0 is accepted but H_1 holds (a genuine pixel is declared forged).

[3] In the following, we model all quantities of interest, except for the deterministic PRNU, as random variables/fields, using the same symbol p for probability density/mass functions, being obvious from the context which one applies.

y

$$\widehat{u} = \arg\max_{u\in\{0,1\}^N} \; p(u|y). \tag{10}$$

Like in [21], however, we consider the noise residual $r = y - f(y)$ in place of the original image, because of its reduced noise power, although this is sub-optimal, in principle, because r is not a sufficient statistic [40] for our decision problem. From r we then compute the decision statistics ρ. Under the hypothesis H_0 (forged pixel) the expected value of ρ_i is zero, since noise n is supposed not to depend on z. Under H_1, instead, the expected value of ρ_i is larger than zero, but not known, as it depends in a complex way on local signal features. Since this information is necessary to make any sensible decision, we resort to a further statistic, the predictor $\widehat{\rho}_i$ proposed in [21], and assume

$$\widehat{\rho}_i \simeq E[\rho_i|H_1]. \tag{11}$$

Our problem becomes therefore

$$\begin{aligned}\widehat{u} &= \arg\max_{u\in\{0,1\}^N} \; p(u|\rho,\widehat{\rho}) \\ &= \arg\max_{u\in\{0,1\}^N} \; p(\rho|u,\widehat{\rho})p(u|\widehat{\rho}) \\ &= \arg\max_{u\in\{0,1\}^N} \; p(\rho|u,\widehat{\rho})p(u),\end{aligned} \tag{12}$$

in which the second equality comes from the application of the conditional Bayes law and the last from the fact that $\widehat{\rho}$ does not depend on u, but only on the image content, be it genuine or forged. The first term in (12), $p(\rho|u,\widehat{\rho})$, is the conditional likelihood of observing ρ and is the only term depending on the data, while the second, $p(u)$, accounts for the prior probability of the labels. Equation (12) provides some insight about the strength of the Bayesian approach. The prior term, in fact, allows us to take into account all available knowledge on the expected forgery map and guide the decision process towards reasonable results. In the absence of such a term, decisions are typically taken independently for each pixel, which could generate fragmented and inconsistent maps calling for intense *ad hoc* postprocessing. On the contrary, by choosing a suitable model for u, which takes into account the strong spatial dependencies exhibited by natural images, decisions are taken jointly, penalizing ultimately maps with isolated points or many small regions and providing a smooth output.

Obviously, the choice of the prior model plays a pivotal role for performance and, not least, for the complexity of the optimization algorithm, which is why we use a Markovian prior. Markov random fields (MRF) [41], [42], [43], [44] are relatively simple and effective tools to model the prior distribution of an image. An image u, defined on a suitable domain Ω, is said to be a MRF with respect to the neighborhood system η when each pixel u_i depends on the rest of the image only through a selected group of neighbors $u_{\eta(i)}$,

$$p(u_i|u_{\Omega-i}) = p(u_i|u_{\eta(i)}). \tag{13}$$

Thanks to the Markovian property one avoids the challenging problem of assigning a global prior, and specifies only local statistics of the image $p(u_i|u_{\eta(i)})$. It can be proved that any positive MRF has a Gibbs probability law

$$p(u) = \frac{1}{Z}\exp[-U(u)] = \frac{1}{Z}\exp[-\sum_{c\in\mathcal{C}} V_c(u)], \tag{14}$$

where Z is a normalizing constant, and $U(u)$ is a suitable energy function depending, in turn, on the potentials $V_c(\cdot)$. The potentials are defined on small groups of neighboring pixels, called cliques c, and are ultimately responsible for the MRF properties. Consider for example a two-pixel clique $c = (i,j)$, with associated potential $V_c(u_i,u_j) = |u_i - u_j|$ and $u_i,u_j \in \mathbb{R}$. If $u_i = u_j$ the clique c will contribute a unitary factor to the overall probability of the image, while if $|u_i - u_j| = \Delta \gg 0$ it will contribute a factor $e^{-\Delta}$ very close to 0. Therefore, with this choice of the local potential, images with a sharp transition between pixels i and j are made very unlikely *a priori*, and smooth images are highly preferred.

To model our label field we resort to the popular Ising model [43] where only single-site cliques $\{c'\}$, and 4-connected two-site cliques $\{c''\}$ are considered and the potentials are

$$V_{c'}(u_i) = \begin{cases} -\alpha/2 & \text{if } u_i = 0 \\ +\alpha/2 & \text{if } u_i = 1 \end{cases} \tag{15}$$

$$V_{c''}(u_i,u_j) = \begin{cases} \beta & \text{if } u_i \neq u_j \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

Single-site potentials are directly related to the prior probability p_0 and p_1 of the classes, in particular

$$\alpha = \log\frac{p_0}{p_1}. \tag{17}$$

Two-site potentials, instead, penalize label transitions (remember the minus sign before the energy) between 4-connected sites, enforcing a bias toward smooth images whose strength depends on the edge-penalty parameter β. For $\beta = 0$ there is no interaction between pixels, no matter how close they are, while the bias against label transition grows stronger with increasing values of β, and extends well beyond a local clique thanks to chain propagation. Therefore, in the absence of observable data, the prior probability is maximized by a flat map, with all labels equal to the one most probable *a priori*.

Turning to the likelihood, ρ, we assume conditional independence (this and other hypotheses will be discussed later) by which, after taking the negative log, we can rewrite the problem as

$$\begin{aligned}\widehat{u} = \arg\min_{u\in\{0,1\}^N} \Bigg\{ &-\sum_{i=1}^{N}\log p(\rho_i|u_i,\widehat{\rho}_i) + \\ &+\sum_{c'\in\mathcal{C}} V_{c'}(u_i) + \sum_{c''\in\mathcal{C}} V_{c''}(u)\Bigg\}\end{aligned} \tag{18}$$

that is

$$\begin{aligned}\widehat{u} = \arg\min_{u\in\{0,1\}^N} \Bigg\{ &-\sum_{i=1}^{N}\log p(\rho_i|u_i,\widehat{\rho}_i) + \\ &+\alpha\sum_{i=1}^{N} u_i + \beta\sum_{i=1}^{N}\sum_{j\in\mathcal{N}_i}|u_j - u_i|\Bigg\}\end{aligned} \tag{19}$$

where the regularization term $R(u) = \sum_{i=1}^{N} \sum_{j \in \mathcal{N}_i} |u_j - u_i|$, with $\mathcal{N}_i$ the set of four-connected neighbors of i, is the sum of all class transitions over all 4-connected cliques of the image, hence it counts the number of edges in the map or, alternatively, the length of all boundaries between different regions.

Now, since a generic function $g(u_i)$ of the binary-valued variable u_i can be written alternatively as $g(u_i) = u_i[g(1) - g(0)] + g(0)$, and hence $\arg\min_{u_i} g(u_i) = \arg\min_{u_i} u_i[g(1) - g(0)]$, we can rewrite (19) as

$$\widehat{u} = \arg\min_{u \in \{0,1\}^N} \left\{ -\sum_{i=1}^{N} u_i[\Lambda(\rho_i|\widehat{\rho}_i)] + \sum_{i=1}^{N} u_i \alpha + \beta R(u) \right\}, \tag{20}$$

where

$$\Lambda(\rho_i|\widehat{\rho}_i) = \log \frac{p(\rho_i|u_i = 1, \widehat{\rho}_i)}{p(\rho_i|u_i = 0, \widehat{\rho}_i)} \tag{21}$$

is the log likelihood ratio between the two hypotheses. Finally, we assume the likelihood to be Gaussian under both hypotheses, with zero mean and variance σ_0^2 under H_0, and mean $\widehat{\rho}_i$ and variance σ_1^2 under H_1, by which we readily obtain

$$\Lambda(\rho_i|\widehat{\rho}_i) = \frac{\rho_i^2}{2\sigma_0^2} - \frac{(\rho_i - \widehat{\rho}_i)^2}{2\sigma_1^2} + \log \frac{\sigma_0}{\sigma_1} \tag{22}$$

and, recalling also (17), eventually rewrite (20) in explicit form as

$$\widehat{u} = \arg\min_{u \in \{0,1\}^N} \left\{ \sum_{i=1}^{N} u_i \left[\frac{(\rho_i - \widehat{\rho}_i)^2}{2\sigma_1^2} - \frac{\rho_i^2}{2\sigma_0^2} + \right. \right.$$
$$\left. \left. - \log \frac{\sigma_0}{\sigma_1} - \log \frac{p_1}{p_0} \right] + \beta R(u) \right\} \tag{23}$$

This is the final expression to consider to look for the optimal $\widehat{u}$ which, by itself, is by no means a trivial task. Before that, however, let us gain some insight into the meaning of (23). Assume $\beta = 0$, for the time being, which means that decisions are taken independently for each pixel, selecting $u_i = 1$ if the term between square brackets, call it *biased likelihood*, is positive, and $u_i = 0$ otherwise. If we assume also $\sigma_0^2 = \sigma_1^2$ and $p_0 = p_1$ (no bias) the algorithm reduces to comparing the correlation index ρ_i with a threshold $\widehat{\rho}_i/2$ placed halfway between the two means. Taking into account unequal prior probabilities and unequal variances modifies somewhat the decision regions, favoring for example the most probable class, without altering, however, the essence of the procedure. If we now consider $\beta > 0$, decisions are not independent anymore, as transitions are penalized, and a decision can be reverted if convenient. This can happen especially if the biased likelihood is small, that is, typically, when $\widehat{\rho}_i$ is small. However, if the biased likelihood keeps the same sign over a relatively large and compact area, there is no reason to change decisions, or refrain from taking a decision at all, even if absolute values are small. Therefore, it becomes possible to detect relatively small forgeries even in dark and textured regions of the image. This is the fundamental improvement w.r.t. the original algorithm, which did not trust at all the decision statistic in problematic regions, throwing in the towel. Both algorithms enforce some compactness constraints, but in [21] this is done only by morphological filtering of the label map, after irreversible hard decisions have been already taken. Here, on the contrary, likelihood and prior are weighted optimally (in the limit of the accuracy of the model) before taking any decision, and the compactness constraint is taken into account through the regularization term $\beta R(u)$.

A few words are due about the hypotheses. The conditional independence of the likelihood does not hold true if correlation indexes are computed on largely overlapping sliding windows. A necessary condition to restore it is to use disjoint blocks, but this would entail an annoying loss of spatial resolution. Therefore, as a reasonable compromise, we use a limited subsampling (8×8, only in the optimization phase) which guarantees a good spatial resolution, allows for a significantly saving in CPU time, and reduces the spatial correlation of the likelihood field. On the other hand, the residual correlation, which modifies the absolute values of the global likelihood, is automatically taken into account through the edge penalty parameter β, set by means of preliminary experiments. Concerning the Gaussian model, it was observed in [21] that it fits experimental data very accurately under H_0 but not under H_1, where a generalized Gaussian (GG) model was preferred. However, the choice of the GG model is strongly influenced by a small number of outliers that lie on the right tail of the distribution (large ρ_i) of little interest for the decision. In the region where a good modeling is more important, between 0 and $\widehat{\rho}$, the simpler Gaussian fitting seems accurate enough. The ratio between the prior probabilities p_1/p_0 can be based on actual observations or, more practically, based on the different risks associated with false alarm and missing detection errors. This choice can be loosely related, therefore, with the choice of the constant FAR threshold of [21]. Finally, the edge penalty (actually, the weight associated with it in the convex optimization procedure) is set based on the outcome of a number of controlled preliminary experiments.

B. Optimization

Let us rewrite the problem of (23) more compactly as

$$\widehat{u} = \arg\min_{u \in \{0,1\}^N} \left\{ \sum_{i=1}^{N} u_i F_i + \beta R(u) \right\}, \tag{24}$$

where F_i is the biased likelihood term associated with site i.

This is a classical problem in the MRF literature, and in general it is NP-hard even for the simplest models. Stochastic sampling algorithms, like the Simulated Annealing [45], are able to find global minima but require a prohibitively slow convergence schedule, providing controversial results if faster schedules are used. The research has then focused on suboptimal methods, like the popular Iterated Conditional Modes (ICM), used already in [42], which locates local optima or deterministic algorithms such as Graph Cuts [46] or Belief Propagation [47].

A powerful alternative approach, proposed in [48] for two-class problems and extended in [49] to multiple classes, relies on the concept of functional "lifting", used to convert a non-convex problem into a convex one admitting the same solution, to be solved by means of suitable optimization techniques.

Here, we follow this approach, and therefore need to address the following steps: *i)* associate a suitable convex problem with (24) and *ii)* solve it.

First of all, to formulate our problem in terms of convex optimization we must replace the discrete domain $\{0,1\}^N$, corresponding to a combinatorial optimization, with the convex domain $[0,1]^N$, the unitary hypercube in $\mathbb{R}^N$, obtaining eventually

$$\widehat{u} = \underset{u\in[0,1]^N}{\arg\min} \left\{ \sum_{i=1}^{N} u_i F_i + \beta \sum_{i=1}^{N} \sum_{j\in\mathcal{N}_i} |u_j - u_i| \right\} \qquad (25)$$

It can be proved [48], [49] that (25) is equivalent to (24), because a solution to the integer-valued problem can be obtained by thresholding at any level $\mu \in]0,1[$ the solution of the real-valued problem.

Among the many approaches proposed to solve this class of problems, we resort here to proximal methods which can handle a wide class of convex optimization problems involving non-differentiable functions. Proximal methods guarantee convergence (under weak conditions) in a reasonable time even for large-scale problems, and have shown good performance and robustness to numerical errors (we refer to [50] for a survey on proximal algorithms and their applications, while more advanced applications are considered in [51]). In particular, we use the class of primal-dual algorithms [52], [53], [54], [55], which are able to address the following general convex optimization problem

$$\widehat{u} = \arg \min_{u\in\mathbb{R}^N} \left\{ f(u) + \sum_{k=1}^{K} g_k(L_k u) + h(u) + a^\top u \right\}, \qquad (26)$$

where $f : \mathbb{R}^N \mapsto \mathbb{R}_+$ and $g_k : \mathbb{R}^{M_k} \mapsto \mathbb{R}_+$ are proper lower-semicontinuous convex functions, $L_k \in \mathbb{R}^{M_k \times N}$ are linear operators, $h : \mathbb{R}^N \mapsto \mathbb{R}_+$ is a differentiable convex function with Lipschitzian gradient, $a \in \mathbb{R}^N$ is a constant vector, and u is regarded as a column vector. Indeed, our minimization problem fits nicely into this framework. First of all we set $a = F$. As for the regularization term, it can be expressed as $\beta R(u) = g(Lu)$, where L is a linear operator such that Lu is the vector that stacks all prime differences between four-connected neighbors in u, while $g(\cdot) = \beta\|\cdot\|_1$ is the sum of absolute values (i.e. the ℓ_1-norm) scaled by β. The third term is simply neglected, $h(u) = 0$, and, finally, we set $f(u) = \iota_{[0,1]^N}(u)$, where ι_C denotes the indicator function of a convex subset $C \subset \mathbb{R}^N$, equal to zero for $u \in C$ and unbounded outside, allowing us to force the solution to belong to the unitary hypercube (note that in (26) the min is over all $\mathbb{R}^N$). With these positions, we can resort, for example, to the M+LFBF algorithm proposed in [54], summarized above, where

$$P_C(u) = \underset{v\in C}{\arg\min}\ \|u - v\| \qquad (27)$$

is the orthogonal projection operator onto the convex set $C \subset \mathbb{R}^N$ and hence $P_{[u',u'']^N}(u) = (\max\{u', \min\{u_i, u''\}\})_{1\leq i\leq N}$. The algorithm stops when either the convergence criterion

$$\|u^{(t+1)} - u^{(t)}\|/\|u^{(t)}\| < 10^{-5} \qquad (28)$$

is satisfied or the maximum allowed number of iterations, $N_{\rm it}$=1000, is reached. Experiments show that 1000 iterations are always enough for a good solution. To speed up convergence, we choose the largest possible updating step $\gamma = 1/4 - \epsilon$.

Algorithm 1 Iterations of M+LFBF [54] for Problem (25)

```
set $\gamma \in ]0, 1/4[$,   $u^{(0)} \in \mathbb{R}^N$,   $v^{(0)} \in \mathbb{R}^{4N}$
for $t = 0 : N_{\rm it}$ do
    $\widetilde{u}^{(t)} = P_{[0,1]^N} \left[u^{(t)} - \gamma(F + L^\top v^{(t)})\right]$
    $\widetilde{v}^{(t)} = P_{[-\beta,\beta]^N} \left[v^{(t)} + \gamma\, L u^{(t)}\right]$
    $u^{(t+1)} = \widetilde{u}^{(t)} - \gamma\, L^\top \left[\widetilde{v}^{(t)} - v^{(t)}\right]$
    $v^{(t+1)} = \widetilde{v}^{(t)} + \gamma\, L \left[\widetilde{u}^{(t)} - u^{(t)}\right]$
end for
```

C. Nonlocal denoising

Although the proposed Bayesian formulation of the problem with the associated global optimization procedure will very likely improve the reliability of forgery detection, performance ultimately depends mostly on the quality of the original data. Recalling the expression of the noise residual, rewritten here for the sake of clarity,

$$r = y - \widehat{x} = yk + (x - y)k + (x - \widehat{x}) + \theta = yk + n \quad (29)$$

we see that the local signal-to-noise ratio depends primarily on image intensity y, which multiplies the PRNU and is outside of the designer's control, and the power of the three noise terms. Among them, the denoising error $(x - \widehat{x})$, due to imperfect treatment of edges and textures, typically predominates, and therefore, the quality of denoising impacts immediately and strongly on the detection performance.

Most denoising algorithms separate signal from noise based on their spectral properties, assuming the signal to be dominant at the lower spatial frequencies and noise to prevail at the higher ones. This approach inspires both spatial-domain and transform-domain filters, in more or less explicit forms, including those operating in the wavelet domain. Unfortunately, such a spectral separation holds very well in flat areas of the image, but much less so in textured areas, where the signal has significant components at the higher frequencies of the spectrum. As a result, some signal components are treated as noise and filtered, causing smoothing and blurring in the output image and, what is worse in our point of view, contributing significantly to noise power in the residual.

This is made clear by Fig. 1 showing, on the left, a texture-rich image, y, and the noise residual, $r = y - \widehat{x}$, obtained by subtracting from y the estimate $\widehat{x}$ of the clean image provided by the Mihcak algorithm used in [21]. With perfect denoising ($\widehat{x} = x$), the residual contains only the PRNU and white noise. In textured areas, however, denoising does not perfectly separate signal from noise. Therefore, part of the signal remains in the noise residual, as visible in the figure, and part of the PRNU is removed together with the signal. As a consequence, the PRNU (by now, our useful signal) is attenuated, while the leaked signal (by now, just

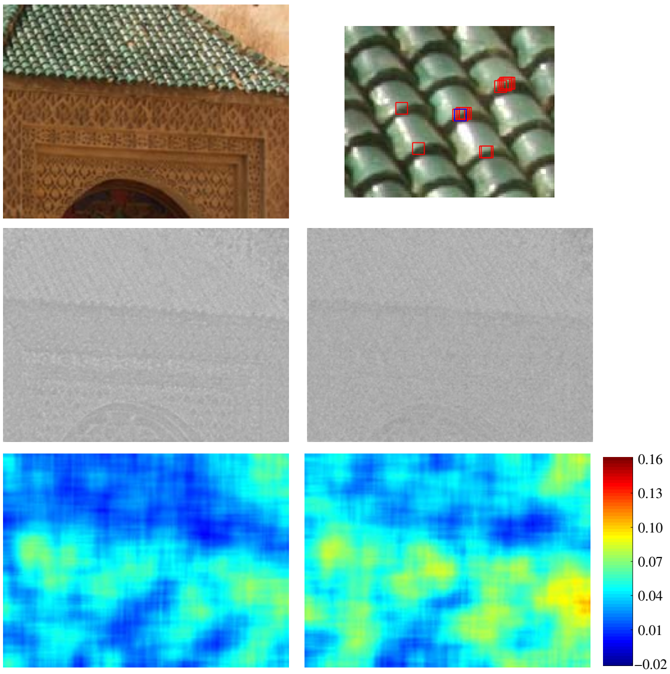

Fig. 1. Effects of image denoising. Left: an original texture-rich image, the noise residual (enhanced) obtained through Mihcak denoising, and the corresponding correlation index field, shown in false colors with blue[red] indicating small[large] correlation indexes. Right: example of similar patches used in joint nonlocal filtering, the noise residual (enhanced) obtained through BM3D denoising, and the corresponding correlation index field.

additional noise) contributes to the overall noise power. In summary, textured areas present a lower SNR than expected and, therefore, the correlation index field ρ, shown bottom-left in false colors, is quite low in all these textured areas, increasing the risk of declaring a forgery when there is none.

To improve the quality of the noise residual, in the proposed algorithm we resort to nonlocal denoising [56] and in particular to the state-of-the-art BM3D algorithm [57]. As the name suggests, to estimate the true value of a given pixel, nonlocal filtering[4] does not rely on its neighbors, but rather on pixels taken anywhere in the image but statistically homogeneous with the target. Thanks to the inherent self-similarity of images, there are several image patches in the surroundings of the target that happen to be very similar to it, and therefore can be assumed to be good predictors. By jointly filtering such patches one obtains a good estimate of the target, with little or no spatial smearing, mimicking a true statistical, rather than spatial, average. As an example, Fig. 1 right shows, on top, a close up of the image considered before, with the target patch enclosed in a blue box, and the patches most similar to it enclosed in red boxes. By jointly processing similar patches, the original image can be well estimated also in the presence of strong texture. Indeed, the noise residual obtained using BM3D (center) presents limited traces of signal, leading to a better correlation index field, especially in the textured areas.

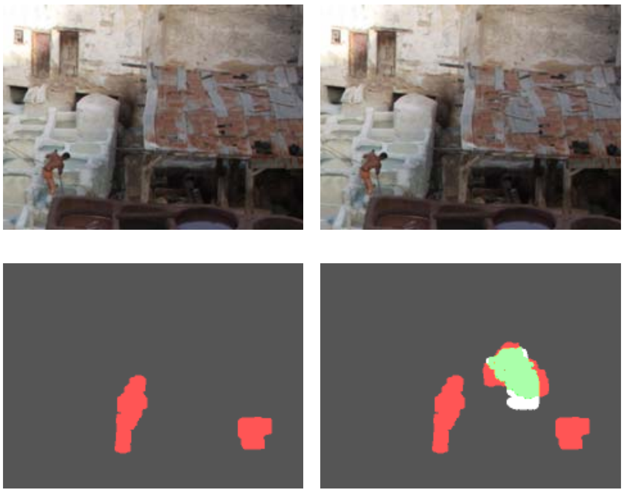

Fig. 2. Color coded detection masks. A genuine image with a tampered version (top), and the corresponding color coded detection masks (bottom). Gray: genuine pixel declared genuine, red: genuine pixel declared tampered (error), white: tampered pixel declared genuine (error), green: tampered pixel declared tampered.

IV. Experimental results

In this section we assess experimentally the performance of the proposed technique. Experiments are carried out on four cameras, a Canon EOS 450D, a Canon IXUS 95IS, a Nikon D200, and a Nikon Coolpix S5100. For each camera we use a first training set of 200 images to estimate the PRNU pattern and another training set of 20 images to design the predictor. Performance indicators are then computed on a larger set, disjoint from the training sets, comprising 600 images for each camera. All training and test images have the same size of 768×1024 pixels, and are cropped from the same region of the original images output by the camera. For each test image we consider both the genuine version, used to look for false alarms (wrongly declared forgeries), and a forged version, used to look for correct detections (correctly declared forgeries). To create the forged version we replace a square at the center of the image with an equal-size square taken at random (but in a reproducible way) from another image of the same or different camera. To study how performance depends on forgery size, we use forgeries of three sizes, $128{\times}128$, $256{\times}256$, and $384{\times}384$ pixels, creating thus three test subsets of 200 images each.

In order to assess separately the improvements due to the Bayesian formulation and to the improved filtering we consider all four combinations in our experiments, that is, the original technique with constant false acceptance rate (CFAR) decision rule and either Mihcak or BM3D denoising, and the proposed version with the Bayesian decision rule and, again, Mihcak or BM3D denoising. In the following we will call these techniques for short CFAR-Mihcak, CFAR-BM3D, Bayes-Mihcak and Bayes-BM3D, respectively.

Fig. 2 shows an example genuine image together with a tampered version with a medium-size realistic forgery at the center and, on the bottom row, the corresponding color-coded decision masks output by the CFAR-Mihcak algorithm. The first mask, obtained for the genuine image, is used to compute

[4] We convey just the basic concepts, here, thus providing a necessarily imprecise description. The interested reader is referred to the specific literature [56], [57], [58] for a more thorough treatment. Some preliminary results on the use of nonlocal denoising for PRNU detection are reported in [25].

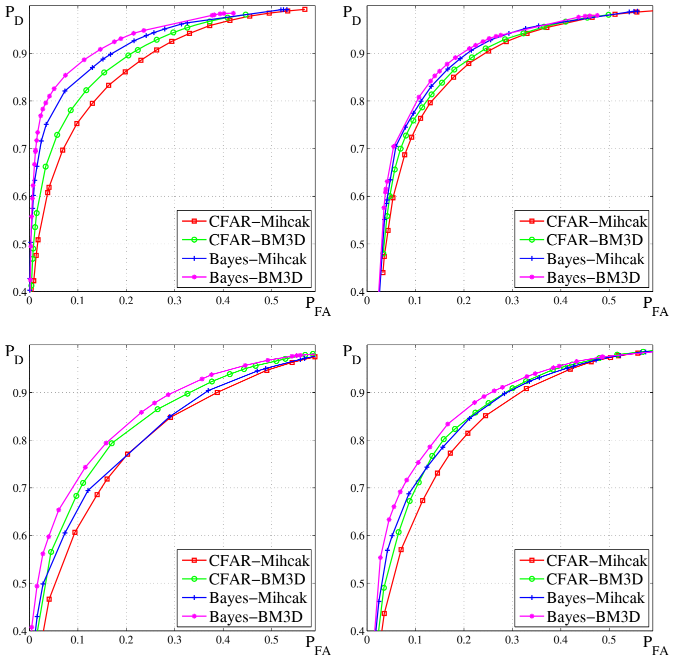

Fig. 3. Pixel-level ROCs (close-up) for all cameras under test: Canon EOS and IXUS (top), Nikon D200 and Coolpix (bottom).

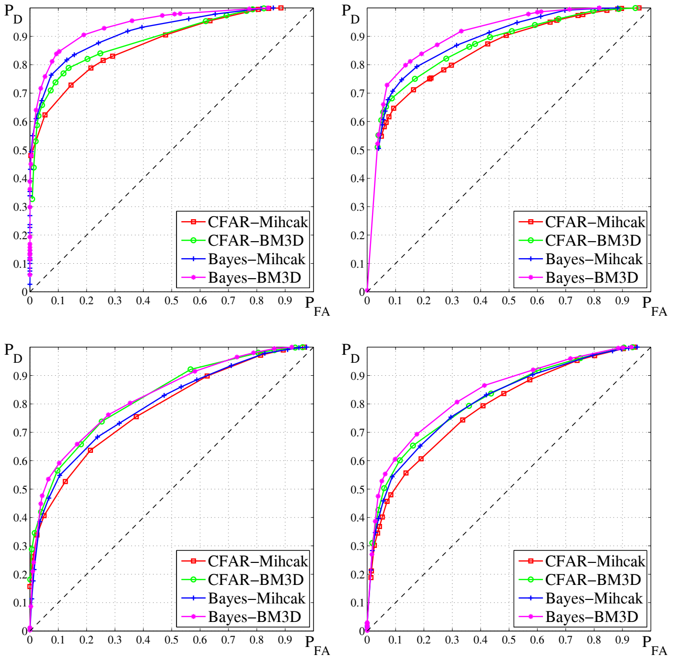

Fig. 5. Image-level ROCs for all cameras under test: Canon EOS and IXUS (top), Nikon D200 and Coolpix (bottom).

the false alarm probability, as the ratio between the number of pixels declared forged (in red) and the image size. The second mask, obtained for the forged image, is used to compute the detection probability, as the ratio between the number of correctly identified forged pixels (in green) and the forgery size (white + green).

A. *Pixel-level analysis*

In Fig. 3 we show, for each of the four cameras, the ROCs (receiver operating characteristics) of the four algorithms, computed on the complete test set (600 images with forgeries of various sizes). To improve readability, we show only a close-up of the most relevant top-left region. Each ROC is computed by varying the algorithm main parameters, γ_1 and γ_2 for CFAR, α and β for Bayes, in a wide range, and then taking the upper envelope of the resulting (P_{FA}, P_D) points. Although CFAR-Mihcak can be considered one of the best forgery detection techniques known to date, Bayes-BM3D technique improves clearly upon it, showing for all cameras a uniformly better ROC. Both the improved denoising filter and the global Bayesian formulation provide significant improvements over CFAR-Mihcak. Their joint use, however, improves performance still further, showing that better data do not solve all problems by themselves, nor is sufficient to adopt a more clever decision strategy irrespective of data quality. Of course, results change slightly from camera to camera but the general behavior is always the same. For the IXUS the ROCs are much closer to one another. For the two Nikon cameras the performance is generally worse than with the Canon cameras. This is not surprising since the performance depends strongly on the average intensity of the PRNU noise, which varies significantly for different manufacturers and camera models, and is somewhat smaller for the Nikon. This fact explains also why, for the Nikon cameras, unlike the Canon, CFAR-BM3D outperforms Bayes-Mihcak. In fact, given the weaker PRNU, improving data reliability is more rewarding than using a better optimization strategy on unreliable data.

To gain insight into how the performance depends on forgery size, Fig. 4 shows in distinct graphs the ROCs (same close-up as before) computed only on large (384×384 pixels), medium (256×256) and small size (128×128) forgeries. To save space results are shown only for the Canon EOS camera, but the same general behavior is observed in all cases. The performance is always very good (remember that we are zooming on the upper-left part of the (P_{FA}, P_D) domain) but, as expected, gets worse and worse as the forgery size decreases, because some small forgeries might be missed altogether. It is worth noting that the gap between the Bayesian and CFAR approach grows very wide in the case of large forgeries, where Bayesian techniques exhibit a near-optimal behavior. In fact, the Bayesian approach allows one to detect large forgeries even in unfavourable conditions, such as dark and textured regions, where CFAR techniques can be fooled.

Although the *pixel-level* false alarm and detection probabilities reported in the above figures are widespread performance indicators in this field, they are not fully appropriate to assess forgery detection performance. A typical user is mostly interested, in order of decreasing importance, in *i)* establishing whether an image is genuine or it has been tampered with, *ii)* finding the approximate location of detected forgeries, *iii)* knowing their approximate size and shape. In fact, once a forgery has been detected, together with its approximate location, one can resort to many other tools or even just visual inspection to obtain more detailed information. Moreover, automatic forgery detection can be used to pre-screen a large number of images, in order to select those more likely to

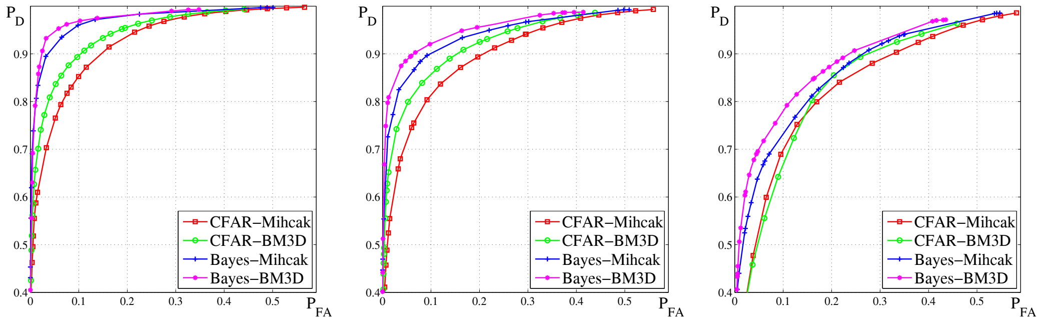

Fig. 4. Pixel-level ROCs (close-up) for the Canon EOS camera, for large (left), medium (middle), and small size (right) forgeries.

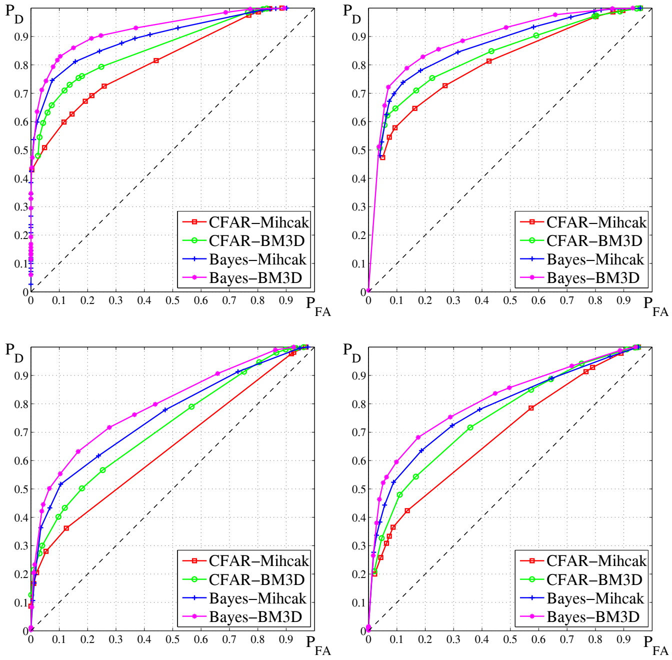

Fig. 6. Object-level ROCs ($\omega = 0.2$) for all cameras under test: Canon EOS and IXUS (top), Nikon D200 and Coolpix (bottom).

have been tampered with for visual analysis by expert photo-interpreters. Therefore, in the following we present two more sets of curves, the customary *image-level* results, and what we call *object-level* results, concerning the probability of correctly detecting forged objects within an image.

B. Image-level analysis

To compute image-level results we consider only the global decision made on the whole image. More precisely, the image under test is declared forged if any of its pixels is (remember that small regions are erased right away from the map) and genuine otherwise. False alarm and detection probabilities are then computed as the fraction of genuine and, respectively, forged images that are declared forged by the algorithm. In Fig. 5 we show results for the four cameras under test, computed over the complete test set. The general behavior is the same as in the case of pixel-level ROCs, with the only difference that all curves are now further from the top-left corner (ideal performance). However, this is obvious considering that a wrong decision on just a small number of pixels may cause a wrong image-level decision on the whole image.

C. Object-level analysis

Image-level ROCs provide information on the ability of an algorithm to classify images as forged or genuine. A more ambitious goal, however, is to localize forgeries with a reasonable accuracy. The PRNU-based approach, unlike many others, has the potential to address this task, therefore it is certainly worth carrying out a specific experimental analysis. To this end we compute object-level ROCs, by modifying the definition of correct detection w.r.t. the image-level case. More precisely, we declare a correct detection only when the binary mask output by the algorithm covers a significant fraction, $\omega \in [0, 1]$, of the actual forged region, namely, going back to the example of Fig. 2, if the green area is large enough w.r.t. forgery. Therefore, all situations in which the output map covers just a tiny part of the actual forgery, or is even disjoint with it, providing little or no hints for visual inspection, are regarded as missing detections.

Fig. 6 reports, for all cameras, the object-level performance computed on the complete test set with $\omega = 0.2$. As expected, all ROCs drift downward w.r.t. the corresponding image-level curves. The impairment, however, is almost negligible for the Bayesian algorithms, with either Mihcak or BM3D denoising, while it is dramatic for the CFAR techniques, especially for the Nikon cameras. In hindsight, this could be expected, since the MRF prior drives the optimization towards a compact and well localized output mask, while no guarantees in this sense is given by the CFAR-based thresholding. It is also interesting that ROCs are now in the same order of performance for all cameras, Bayes-BM3D > Bayes-Mihcak > CFAR-BM3D > CFAR-Mihcak, confirming that at object-level the most important improvement comes from making decisions globally as opposed to locally. Fig. 7 shows, with reference only to the Canon EOS, that the Bayes techniques have an object-level performance almost independent of ω, while for the CFAR techniques a clear dependence is observed, indicating a less accurate forgery localization.

We complete this analysis by reporting, in Fig. 8, again for the Canon EOS camera, the object-level ROCs computed separately on large (384×384 pixels), medium (256×256) and small (128×128) forgeries. As expected, performance drops for all algorithms as the forgery size reduces and is quite

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2014.2302078, IEEE Transactions on Information Forensics and Security

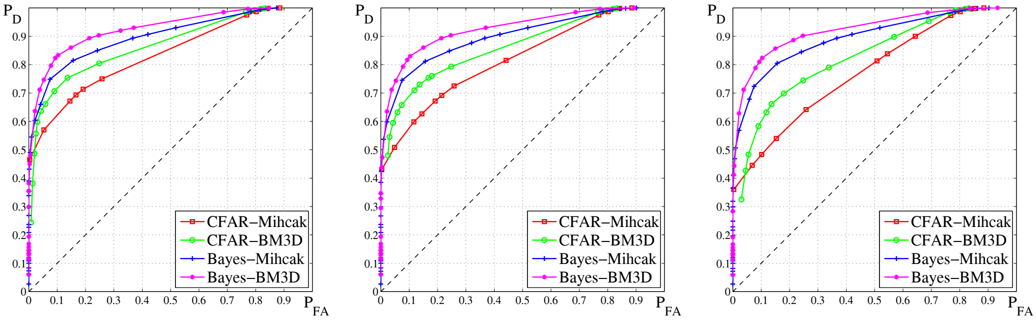

Fig. 7. Object-level ROCs for Canon EOS camera, for ω=0.1 (left), ω=0.2 (middle), and ω=0.4 (right).

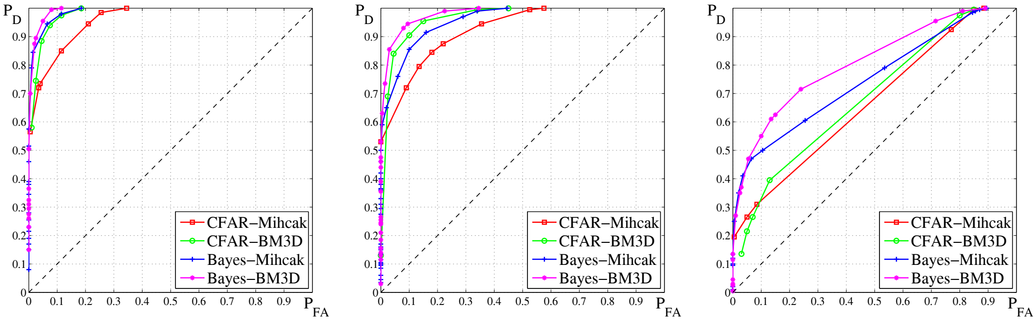

Fig. 8. Object-level ROCs (ω=0.2) for Canon EOS camera, for large (left), medium (middle), and small size (right) forgeries.

poor for the smallest size, which coincides with the size of the window used to compute the decision statistics. In this condition, the statistics used to detect forged pixels are always (except for a single point) computed on mixed data thus impairing their diagnostic power. This is certainly one of the major limitations of the PRNU approach, and probably the main topic to deal with in future research. Barring this limiting case, performance is generally good, especially for the proposed version of the technique, which provides always a significant gain w.r.t. the original one.

In summary, this performance analysis shows that the proposed technique is not only able to reliably classify a test image as genuine or forged, but also to localize with good accuracy the detected forgery, provided it is large enough w.r.t. the analysis window. Under all these respects, it improves significantly over the reference technique.

D. Sample results on realistic cases

To gain a better insight into the behavior of proposed and reference algorithms we now consider some sample results obtained by running the algorithms on a set of real-world tampered images taken by our four cameras. We consider both the insertion of a new object into the scene (splicing) and the cancellation of an object from the scene, realized by copy-moving part of the background to hide it. For both CFAR-Mihcak and Bayes-BM3D we use the set of parameters corresponding to the best operating points in their respective object-level ROCs with $\omega = 0.20$. As an example, for the Canon EOS-450 camera, we use $\alpha = \log 3, \beta = 48$ for Bayes-BM3D, and $\gamma_1 = 0.026, \gamma_2 = 0.013$ for CFAR-Mihcak. Here we focus on purpose on some situations where clear differences arise in the algorithms' behavior and the reader is alerted not to give statistical significance to this short review of cases.

Fig. 9 shows some selected examples, one for each row, reporting from left to right the original (a) and tampered (b) versions of the image, the correlation index fields predicted (c), computed after Mihcak filtering (d), and computed after BM3D filtering (e), displayed with the same colorbar used in Fig. 1, and finally the color-coded detection masks provided by the CFAR-Mihcak (f) and the Bayes-BM3D (g) algorithms.

The first example (Camel) corresponds to a simple case, with a large, bright and smooth forgery which, in fact, is successfully and accurately detected by both algorithms. Notice that the index fields computed after Mihcak and BM3D filtering follow closely the predicted field except in the region of the forgery, where the original PRNU is missing and a large deviation is observed. Notice also that no obvious visual hint suggests the presence of a forgery, making the detection virtually impossible without proper tools. The second example (Road Sign) is similar to the previous one except for the two forgeries that are much smaller than before, about the size of the 128×128 pixel analysis window. As a consequence, the CFAR-Mihcak algorithm misses both of them, while Bayes-BM3D detects them correctly. Notice, in the center-bottom of the computed index correlation fields, the reddish square region with sharp edges caused by a single defective sensor, a feature that could be exploited to improve performance. The third image (Wading Bird) is much darker than the previous ones and also more textured. As a consequence, the computed correlation indexes are quite small (blueish), especially after Mihcak filtering, which might easily induce false alarms. The optimum thresholds of the original algorithm, set so as to limit false alarms, discard the forgery region as well, causing

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2014.2302078, IEEE Transactions on Information Forensics and Security

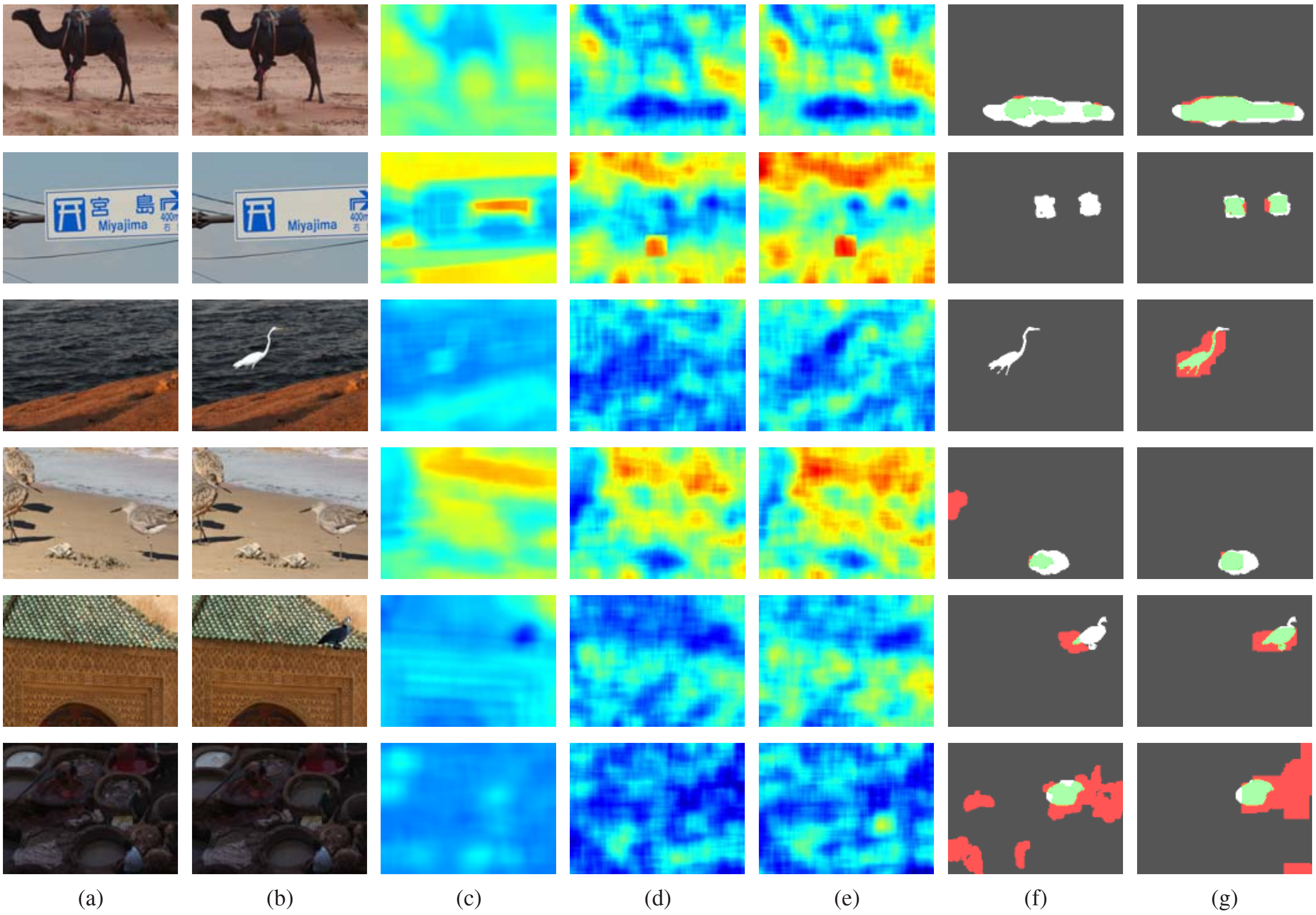

Fig. 9. Forgery detection results for some selected examples: Camel, Road Sign, Wading Bird, Beach, Roof, Market. For each image we show from left to right: original image, tampered version, predicted correlation index field $\widehat{\rho}$, correlation index field ρ computed after Mihcak filtering, correlation index field ρ computed after BM3D filtering, color-coded detection mask provided by CFAR-Mihcak, color-coded detection mask provided by Bayes-BM3D.

a missed detection. The proposed technique, instead, detects accurately the spliced object, thanks to both the better filtering and the more sophisticated decision strategy. The problem with textures is even more evident in the fourth example (Beach), where both algorithms detect easily the copy-move forgery, but CFAR-Mihcak presents also a false alarm caused by the bird feathers, highly textured, and imperfectly filtered by the wavelet-based denoiser. Also in the subsequent example (Roof) texture causes some problems to CFAR-Mihcak. The algorithm detects a forged area which only partially (less than 10%) overlaps the actual spliced object and is therefore classified as a false alarm. This decision appears to be correct in this case, since the algorithm seems to follow the textured roof rather than the dark bird. Finally, we show the case of a dark and textured image (Market) where both algorithms fail, detecting the forgery but generating multiple false alarms as well. With such a low quality image, on the other hand, it is difficult to envision successful strategies. We do not even show the case of forgeries much smaller than the analysis window, since they are obviously missed by definition by both algorithms. Notice, however, going back to the Wading Bird example, that the spliced object is indeed much thinner than 128 pixels, although longer than that, but still detected by the Bayesian strategy.

E. Comparison of running times

Tab.II reports some data on computational complexity measured by the average CPU times spent on a 3.40Ghz, 64bit desktop computer with 8GB memory for a 768 × 1024-pixel image. The proposed technique is somewhat more complex than the original, but not dramatically so. Most of the gap is due to the costly nonlocal denoising, which requires a large number of expensive block matching operations. The decision phase, instead, takes much less than expected. Thanks to the efficient convex optimization method, it requires less than 12 seconds without subsampling, which become just 0.31 seconds with subsampling, less than 5% of the overall CPU time. After the final upsampling, some inexpensive morphological filtering is used to slightly enlarge and smooth all map contours.

F. Comparison with other reference techniques

To complete our analysis, we show here the results of an experiment designed to compare performance with some state-of-the-art techniques, A-DJPG [13], based on JPEG artifacts, and CFA-Loc [19], based on the analysis of the CFA, with code available for both on the Authors' web-site [59]. In order to fairly compare techniques based on very different principles and hypothesis, a pretty realistic setting was designed. A test

	CFAR-Mihcak		Bayes-BM3D	
	mean	st.dev.	mean	st.dev.
Denoising	0.45	0.03	5.45	0.12
Index Field Computation	1.30	0.01	1.30	0.01
Decision/Optimization	0.01	0.00	0.31	0.14
Morphological Filtering	0.03	0.00	0.04	0.00
Overall	1.79	0.03	7.10	0.21

TABLE I
MEAN CPU TIME (S) AND STANDARD DEVIATION FOR THE ORIGINAL AND PROPOSED TECHNIQUES.

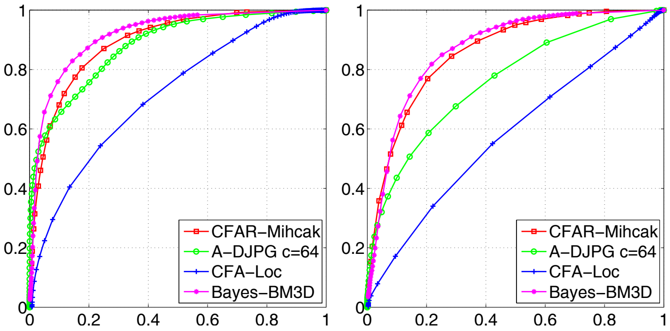

Fig. 10. Pixel-level ROCs for PRNU-based and reference techniques with uncompressed (left) and compressed (right) images.

set of 300 768×1024-pixel images, 75 for each of our test cameras, was created, with realistic forgeries, some of them drawn from the web [60], and others designed by ourselves using the dataset of uncompressed color images available at [61]. We included an equal number of small, medium, and large forgeries, classified based on the largest circle included in the forgery area, with radius $r \leq 64$, $r \in [65, 128]$ and $r \in [129, 192]$, respectively. Note that some forgeries are much smaller than the smallest squares of 128×128 pixels used in the previous analysis. Both uncompressed and compressed ($Q > 75$) forgeries were used, half of them resampled, the other half either rescaled or rotated.

Fig. 10 shows ROCs obtained at pixel level on the test set. To evaluate robustness of the algorithms we considered also the case in which all tampered images were JPEG compressed with quality factor 90. In both cases, Bayes-BM3D performs better than CFAR-Mihcak which, in turn, performs slightly better than A-DJPG and much better than CFA-Loc. In A-DJPG we used all 64 DCT coefficients to generate the likelihood map, since 6 coefficients, as suggested in [13] were not enough to obtain a good performance. It is also interesting to note that while all techniques exhibit a worse performance on compressed images, the impairment is much stronger for A-DJPG and CFA-Loc than for the PRNU-based techniques.

Although a thorough comparative analysis would be certainly interesting, it is out of the scope of this work. It is worth pointing out, however, that the proposed technique improves upon the well-known and extensively tested technique of [21], and can be therefore expected to inherit its good and robust performance.

V. Conclusion and Future Research

Image forgery detection is becoming more challenging by the day, due to the unrelenting advances in image processing. PRNU-based forgery detection is one of the most promising approaches for this task. Provided one has the opportunity to estimate the camera "fingerprint", that is, its PRNU noise pattern, all kinds of forgeries can be dealt with in a uniform manner and with a consistent good performance. Here, we improve upon the seminal PRNU-based forgery detection technique proposed in [21] by recasting the problem in a Bayesian framework, and by modeling the decision variables as a Markov random field, thus accounting for their spatial dependencies. In addition, we resort to state-of-the-art signal and image processing tools: nonlocal denoising to improve estimation of noise residual, and convex optimization to reach a globally optimal solution in a limited time. As a result, the proposed technique provides a significant and consistent performance gain over the original, especially in terms of object-level detection ability, the main parameter of interest for the applications. A modified version of the proposed algorithm together with some other forgery detection tools [62], allowed the GRIP team, led by one of the Authors, to win phase-2 of the First IEEE IFS-TC Image Forensics Challenge [63].

Despite the present advances, there is still much room for improvements. As an example, we are working to design a better and more robust predictor. Our major goal for future research, however, is to improve spatial resolution, allowing for the detection of smaller forgeries. Prior work on this topic [34] showed that image-level segmentation can help increasing resolution in some suitable cases, but segmentation itself is a very challenging and unreliable process. We are currently working towards a new version of this algorithm based on soft segmentation [64].

Acknowledgment

The Authors are grateful to Mr. Davide Cozzolino for his precious help with the experiments, and to the Anonymous Reviewers for the many comments and suggestions that helped improve the paper.

References

[1] "Photo tampering throughout history," http://www.fourandsix.com/photo-tampering-history/

[2] G. Zhou, and D. Lv, "An Overview of Digital Watermarking in Image Forensics," *International Joint Conference on Computational Sciences and Optimization (CSO)*, pp. 332–335, Apr. 2011.

[3] S. Battiato, G.M. Farinella, G.M. Messina and G. Puglisi, "Robust Image Alignment for Tampering Detection," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 4, pp. 1105–1117, 2012.

[4] Y. Zhao, S. Wang, X. Zhang, and H. Yao, "Robust Hashing for Image Authentication Using Zernike Moments and Local Features," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 1, pp. 55–63, Jan. 2013.

[5] I. Amerini, L. Ballan, R. Caldelli, A. Del Bimbo, and G. Serra, "A SIFT-Based Forensic Method for Copy and Move Attack Detection and Transformation Recovery," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 1099–1110, 2011.

[6] P. Kakar, and N. Sudha, "Exposing Postprocessed Copy-Paste Forgeries Through Transform-Invariant Features," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 3, pp. 1018–1028, June 2012.

[7] V. Christlein, C. Riess, J. Jordan, C. Riess, and E. Angelopoulou, "An Evaluation of Popular Copy-Move Forgery Detection Approaches," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 6, pp. 1841–1854, Dec. 2012.

[8] M.K. Johnson, and H. Farid, "Metric measurements on a plane from a single image," *Dept. Comput. Sci., Dartmouth College, Tech. Rep. TR2006-579*, 2006.

[9] M.K. Johnson, and H. Farid, "Exposing Digital Forgeries in Complex Lighting Environments," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 3, pp. 450–461, Sep. 2007.

[10] Q. Liu, X. Cao, C. Deng and X. Guo, "Identifying image composites through shadow matte consistency," *IEEE Transactions on Information Forensics and Security*, vol. 6, pp. 1111–1122, 2011.

[11] Z. Lin, J. He, X. Tang, and C.-K. Tang, "Fast, automatic and fine-grained tampered JPEG image detection via DCT coefficient analysis," *Pattern Recognition*, vol. 42, pp. 2492–2501, 2009.

[12] Y.-L. Chen, and C.-T. Hsu, "Detecting Recompression of JPEG Images via Periodicity Analysis of Compression Artifacts for Tampering Detection," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 2, pp. 396–406, Jun. 2011.

[13] T. Bianchi, and A. Piva, "Image Forgery Localization via Block-Grained Analysis of JPEG Artifacts," *IEEE Transactions on Information Forensics and Security*, vol. 7, pp. 1003–1017, Jun. 2012.

[14] F. Zach, C. Riess, and E. Angelopoulou, "Automated Image Forgery Detection through Classification of JPEG Ghosts," *Pattern Recognition*, vol. 7476, pp. 185–194, 2012.

[15] I. Yerushalmy, and H. Hel-Or, "Digital Image Forgery Detection Based on Lens and Sensor Aberration," *International Journal of Computer Vision*, vol. 92, no. 1, pp. 71–91, Nov. 2011.

[16] H. Fu, and X. Cao, "Forgery Authentication in Extreme Wide-Angle Lens Using Distortion Cue and Fake Saliency Map," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 4, pp. 1301–1314, Aug. 2012.

[17] A.C. Popescu, and H. Farid, "Exposing digital forgeries in color filter array interpolated images," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3948–3959, Oct. 2005.

[18] A.C. Kot, "Accurate Detection of Demosaicing Regularity for Digital Image Forensics," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 4, pp. 899–910, Dec. 2009.

[19] P. Ferrara, T. Bianchi, A. De Rosa, and A. Piva, "Image Forgery Localization via Fine-Grained Analysis of CFA Artifacts," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 5, pp. 1566–1577, 2012.

[20] J. Lukas, J. Fridrich, and M. Goljan, "Detecting digital image forgeries using sensor pattern noise," *Proceedings of the SPIE*, vol. 6072, pp. 362–372, 2006.

[21] M. Chen, J. Fridrich, M. Goljan, and J. Lukas, "Determining Image Origin and Integrity Using Sensor Noise" *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 1, pp. 74–90, 2008.

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

[23] Y. Chen and V.L.L. Thing, "A Study on the Photo Response Non-Uniformity Noise Pattern based Image Forensics in Real-World Applications," in *The 2012 World Congress in Computer Science Computer Engineering and Applied Computing*, 2012.

[24] C.-T. Li and R. Satta, "An Empirical Investigation into the Correlation between Vignetting Effect and the Quality of Sensor Pattern Noise," *IET Computer Vision*, vol. 6, no. 6 pp. 560-566, Nov. 2012.

[25] G. Chierchia, S. Parrilli, G. Poggi, C. Sansone, and L. Verdoliva, "On the influence of denoising in PRNU based forgery detection," in *Proceedings of the 2nd ACM workshop on Multimedia in Forensics, Security and Intelligence* pp. 117–122, 2010.

[26] C.T. Li, "Source Camera Identification Using Enhanced Sensor Pattern Noise," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 2, pp. 280–287, June 2010.

[27] J. Fridrich, "Sensor Defects in Digital Image Forensic," *Digital Image Forensics*, pp. 179–218, 2012.

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

[29] M. Goljan and J. Fridrich, "Sensor-fingerprint based identification of images corrected for lens distortion," in *Proc. of SPIE: Media Watermarking, Security, and Forensics*, vol. 8303, 2012.

[30] T. Gloe, S. Pfennig, and M. Kirchner, "Unexpected Artefacts in PRNU-Based Camera Identification : A Dresden Image Database Case-Study," in *The 14th ACM Workshop on Multimedia and Security*, pp. 109–114, 2012.

[31] M. Goljan, and J. Fridrich, "Digital camera identification from images – Estimating false acceptance probability," in *Proc. 8th Int. Workshop Digital Watermarking*, 2008.

[32] X. Kang, Y. Li, Z. Qu, J. Huang, "Enhancing Source Camera Identification Performance With a Camera Reference Phase Sensor Pattern Noise," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 393–402, 2012.

[33] C. Zhang and H. Zhang, "Exposing Digital Image Forgeries by Using Canonical Correlation Analysis," *20th International Conference on Pattern Recognition*, pp. 838–841, 2010.

[34] G. Chierchia, S. Parrilli, G. Poggi, L. Verdoliva, and C. Sansone, "PRNU-based detection of small-size image forgeries," *International Conference on Digital Signal Processing (DSP)*, pp. 1–6, 2011.

[35] B.B. Liu, Y. Hu, and H.K. Lee, "Source camera identification from significant noise residual regions," in *IEEE International Conference on Image Processing*, pp. 1749–1752, 2010.

[36] S. Bayram, H.T. Sencar, and N. Memon, "Efficient Sensor Fingerprint Matching Through Fingerprint Binarization," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 4, pp. 1404–1413, Aug. 2012.

[37] M. Goljan, J. Fridrich, and T. Filler, "Managing a large database of camera fingerprints," in *Proc. SPIE Conf. Electronic Imaging: Media Forensics and Security XII*, vol. 7541, 2010.

[38] Y. Hu, C. Li, Z. Lai, and S. Zhang, "Fast Camera Fingerprint Search Algorithm for Source Camera," in *5th International Symposium on Communications, Control and Signal Processing*, pp. 2–4, 2012.

[39] G. Chierchia, G. Poggi, C. Sansone, and L. Verdoliva, "PRNU-based forgery detection with regularity constraints and global optimization," *IEEE 15th International Workshop on Multimedia Signal Processing (MMSP)*, pp. 236–241, Pula (I), Oct. 2013.

[40] S. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*, Prentice Hall, 1993.

[41] S. Geman, D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, Nov.1984.

[42] J. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society*, Series B 48, pp. 259–302, 1986.

[43] S.Z. Li, *Markov random field modeling in image analysis*, Springer-Verlag, 2001.

[44] C. D'Elia, G. Poggi, G. Scarpa, "A tree-structured Markov random field model for Bayesian image segmentation," *IEEE Transactions on Image Processing,* vol. 12, no. 10, pp. 1259–1273, Oct. 2003.

[45] S. Kirkpatrick, C. Gelatt, M. Vecchi, "Optimization by simulated annealing," *Science* vol. 220, pp. 671–680, 1983.

[46] Y. Boykov, O. Veksler, R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 1222–1239, 2001.

[47] J.S. Yedidia, W.T. Freeman, Y. Weiss, "Generalized belief propagation," *NIPS*, pp. 689–695, 2000.

[48] T.F. Chan, S. Esedoglu, M. Nikolova, "Algorithms for finding global minimizers of image segmentation and denoising models," *SIAM Journal on Applied Mathematics*, vol. 66, no. 5, pp. 1632–1648, 2006.

[49] T. Pock, D. Cremers, H. Bischof, A. Chambolle, "Global solutions of variational models with convex regularization," *SIAM J. Img. Sci.*, vol. 3, pp. 1122–1145, Dec. 2010.

[50] P.L. Combettes, J.-C. Pesquet, "Proximal splitting methods in signal processing," *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, pp. 185–212, Springer-Verlag, 2011.

[51] G. Chierchia, N. Pustelnik, J.-C. Pesquet, B. Pesquet-Popescu, "Epigraphical projection and proximal tools for solving constrained convex optimization problems – Part I," 2012, http://arxiv.org/abs/1210.5844.

[52] A. Chambolle, T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging", *J. Math. Imag. Vis.*, vol. 40, no. 1, pp. 120–145, 2011.

[53] B.C. Vũ, "A splitting algorithm for dual monotone inclusions involving cocoercive operators", *Adv. Comput. Math.*, vol. 38, no. 3, pp. 667–681, April 2013.

[54] P.L. Combettes, J.-C. Pesquet, "Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators," *Set-Valued Var. Anal.*, vol. 20, no. 2, pp. 307–330, June 2012.

[55] L. Condat, "A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms," *Journal of Optimization Theory and Applications*, in press, 2013.

[56] A. Buades, B. Coll, J.M. Morel, “A Review of image denoising algorithms, with a new one,” *Multiscale Modeling and Simulation*, vol. 4, no. 2, pp. 490–530, 2005.

[57] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, “Image denoising by sparse 3-D transform-domain collaborative filtering,” *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[58] S. Parrilli, M. Poderico, C.V. Angelino, L. Verdoliva, “A nonlocal SAR image denoising algorithm based on LLMMSE wavelet shrinkage,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, pp. 606–616, Feb. 2012.

[59] http://iapp.det.unifi.it/index.php/english/materials_en/source-code_en

[60] http://staff.science.uva.nl/ aloi/

[61] http://homepages.lboro.ac.uk/ cogs/datasets/ucid/ucid.html

[62] D. Cozzolino, D. Gragnaniello, and L. Verdoliva, “A novel framework for image forgery localization,” *Technical report of 1st IEEE IFS-TC Image Forensics Challenge - (Phase 2)*, Nov. 2013. http://arxiv.org/pdf/1311.6932v1.pdf

[63] http://ifc.recod.ic.unicamp.br/fc.website/

[64] G. Chierchia, D. Cozzolino, G. Poggi, C. Sansone, L. Verdoliva, “Guided filtering for PRNU-based localization of small-size image forgeries,” *2014 IEEE International Conference on Acoustics, Speech, and Signal Processing*, submitted.

Giovanni Chierchia received the engineering degree in computer science from University of Naples Federico II, Italy, in 2010. From 2010 to 2011, he worked as research engineer at Institut Mines-Télécom, Télécom ParisTech, CNRS LTCI, Paris, France, and is currently a Ph.D. student in the same institution. His research interests are in convex optimization, focusing on the solution of constrained convex formulation of inverse problems using proximal algorithms, and in image forgery detection based on the analysis of the sensor noise pattern.

Giovanni Poggi received the Laurea degree in electronic engineering from the University Federico II of Naples, Italy, in 1988. He is currently a Professor of telecommunications with the same institution, and Coordinator of the Telecommunication Engineering School. His research interests are in statistical image processing, including compression, restoration, segmentation, and classification, with application to the area of remote-sensing, both optical and SAR, and digital forensics. Prof. Poggi has been an Associate Editor for the IEEE Transactions on Image Processing and Elsevier Signal Processing

Carlo Sansone is currently Full Professor of Computer Science with the Department of Electrical Engineering and Information Technology of the University University of Naples Federico II, Naples, Italy. His research interests cover the areas of image analysis and recognition, pattern recognition, graph matching and information fusion. From an applicative point of view, his main contributions were in the fields of biomedical image analysis, biometrics, intrusion detection in computer networks and image forensics. He has authored more than 150 research papers in international journals and conference proceedings. He is Associate editor of Electronic Letters on Computer Vision and Image Analysis and Elsevier Information Fusion. He was also co-editor of two special issues on International Journals, and three books. Prof. Sansone is Vice-President of the GIRPR the Italian Chapter of the International Association for Pattern Recognition (IAPR) and is member of the IEEE.

Luisa Verdoliva received the Laurea degree in telecommunications engineering and the Ph.D. degree in information engineering from the University Federico II of Naples, Italy, in 1998 and 2002, respectively. She is currently Assistant Professor of telecommunications with the same institution. Her research is on image processing, in particular compression and restoration of remote-sensing images, both optical and SAR, and digital forensics. Dr. Verdoliva lead the GRIP team of the Federico II University who ranked first in both the forgery detection and localization phases of the First Image Forensics Challenge organized in 2013 by the IEEE Information Forensics and Security Technical Committee (IFS-TC).