Effective Source Camera Identification based on MSEPLL Denoising Applied to Small Image Patches

Wen-Na Zhang[1, 2], Yun-Xia Liu[1, 2, 3,*], Ze-Yu Zou[1, 2], Yun-Li Zang[4], Yang Yang[5] and Bonnie Ngai-Fong LAW[6]
[1] School of Information Science and Engineering, University of Jinan, Jinan 250022, China
[2]Shandong Provincial Key Laboratory of Network Based Intelligent Computing, University of Jinan, Jinan 250022, China
[3] School of Control Science and Engineering Shandong University, Jinan 250061, China
E-mail: ujn_zhangwn@qq.com, ise_liuyx@ujn.edu.cn, zeyuzou@foxmail.com, Tel: +86-531-82767500
[4] Integrated Electronic Systems Lab Co., Ltd, Jinan 250100, China
E-mail: zangyunli@ieslab.cn, Tel: +86-531-88018000
[5] School of Information Science and Engineering Shandong University, Qingdao 266237, China
E-mail: yyang@sdu.edu.cn, Tel: +86-532-58630701
[6] Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong, China
E-mail: ngai.fong.law@polyu.edu.hk, Tel: +852-27664746

***Abstract*— Sensor Pattern Noise (SPN) has proven to be an effective fingerprint for source camera identification, while its estimation accuracy heavily relies on denoising algorithm. In this paper, an effective source camera identification scheme based on Multi-Scale Expected Patch Log Likelihood (MSEPLL) denoising algorithm is proposed, firstly. With enhanced prior modeling across multiple scales, MSEPLL can accurately restore the original image. As a consequence, estimated SPN is less influenced by image content. Secondly, the source camera identification problem is formulated by hypothesis testing, where normalized correlation coefficient is adopted for SPN detection. Finally, the effectiveness of the proposed method is verified by abundant experiments in terms of identification accuracy as well as receiver operating characteristic. Performance improvement is more prominent for small image patches, which is more conducive to real forensics applications.**

I. INTRODUCTION

With the rapid development of imaging devices such as mobile phones and digital cameras, digital images have gradually become a popular manner for people to record and share their daily lives. People can arbitrarily modify digital images on mobile phones or computer terminals, which bring hidden serious information security problems despite of convenience. It is often necessary to verify the authenticity and integrity of images through technological means in applications such as judicial forensics and news documentary. Therefore, digital forensics technology has received increasing attention.

Given an image to be analyzed, the task for source camera identification is to identify the source camera device that shot it. Sensor pattern noise (SPN) is proposed in [1] for source camera identification for its uniqueness and stability for the first time. Lukas et al. adopted a wavelet domain adaptive denoising filter [2] to obtain denoised images, where the SPN is then constructed by averaging. The maximum likelihood estimation method (MLE) utilized in [3] is proven to bring performance improvement in SPN estimation accuracy. Li modeled the SPN estimation problem in wavelet domain and proposed six models for SPN enhancement by assigning different weights to SPNs in the wavelet domain [4]. It is revealed in [5] that the utilization of more advanced image denoising algorithm will lead to stable camera identification performance improvement. However, due to the imperfection of denoising methods in edge and texture regions, quality degradation is usually observed in SPN obtained from the above methods that there are certain scene content related structures left in estimated SPN.

To suppress the effects of scene content, Kang et al. proposed an effective SPN predictor based on eight-neighbor context-adaptive interpolation algorithm [6]. The guided filter [7] was first adopted in [8] to identify the source camera, where the filter radius is adaptively determined according to local texture intensity. Ref [9] proposed a block weighted averaging module to further suppress the effects of scene content. A more flexible, pixel-wise confidence map is constructed in [10] with kernel principal component analysis based on local intensity and variance features, where higher weights are assigned to more reliable pixels. However, experimental results reported in these works are mainly based on large patch size settings, such as such as 512×512, 1024×1024 and original image sizes. The SPN identification accuracy drops dramatically with the decease of patch size. Improving the quality of SPNs extracted from small-sized patches has great significance for SPN-based applications [11] [12].

In this paper, an effective source camera identification method based on multi-scale expected patch log likelihood (MSEPLL) denoising algorithm [13] is proposed. With the enhanced representation capability of MSEPLL, estimated SPN is less contaminated by image content, thus enables more accurate camera identification. The effectiveness of the proposed method is verified by experimental results conducted on six cameras selected from the Dresden database. Performance improvement is more prominent for small patch sizes, which favors the demanding source identification application.

The remainder of this paper is organized as follows. Part 2 introduces the MSEPLL denoising algorithm. The proposed

method is presented in Part 3 in detail. Part 4 discusses experimental results, while Part 5 concludes the paper.

II. MSEPLL DENOISING ALGORITHM

In this section, we briefly introduce the multi-scale expected patch Log likelihood (MSEPLL) denoising algorithm, as we rely on this method for accurate SPN estimation.

A. Multi-Scale Expected Patch Log Likelihood

Under an image restoration setting, assuming the clean image X is contaminated by the linear operator $\boldsymbol{A}$ and the additive Gaussian white noise N with the standard deviation σ_0. Given the observed image $Y = \boldsymbol{A}X + N$, the maximum a posterior (MAP) method for recovering X is

$$\begin{aligned}\max_X P(X\mid Y) &= \max_X P(Y\mid X)P(X)\\ &= \min_X -\log P(Y\mid X) - \log P(X),\end{aligned} \tag{1}$$

where $P(X)$ denotes a global prior of the image. The MSEPLL algorithm formulates the global prior by accumulating local patch ingredients:

$$P_{MSEPLL}(X) = w_1\sum_i \log P_1(\boldsymbol{R}_i X) + w_2\sum_i\sum_j \log P_2\left(\hat{\boldsymbol{R}}_i \boldsymbol{D}_j \boldsymbol{H} X\right), \tag{2}$$

where $\boldsymbol{R}_i$ and $\hat{\boldsymbol{R}}_i$ denotes the operator extracting the i^{th} patch from the image and the decimated signal $\boldsymbol{D}_j\boldsymbol{H}X$, $\boldsymbol{D}_j$ corresponds to the j^{th} down-sampling grid, $\boldsymbol{H}$ is a low-pass filter, w_1 and w_2 are the weights representing the importance of the different scales. Note that different local priors (P_1 and P_2) are adopted to model the two successive scales, respectively.

By placing the MSEPLL prior into (1), we have:

$$\min_X \frac{\lambda}{2}\|\boldsymbol{A}X - Y\|_2^2 - w_1\sum_i \log P(\boldsymbol{R}_i X) - w_2\sum_i \log P\left(\hat{\boldsymbol{R}}_i \boldsymbol{S} X\right), \tag{3}$$

where $\lambda = \frac{p}{\sigma_0^2}$, p is the patch size and $\boldsymbol{S} = \boldsymbol{DH}$. Using Half Quadratic Splitting [14] by introducing a set of auxiliary variables, the close form solution could be obtaiend as

$$\begin{aligned}X = &\left(\lambda \boldsymbol{A}^T A + w_1\beta\sum_i \boldsymbol{R}_i^T\boldsymbol{R}_i + w_2\hat{\beta}\sum_i \boldsymbol{S}^T\hat{\boldsymbol{R}}_i^T\hat{\boldsymbol{R}}_i\boldsymbol{S}\right)^{-1}\\ &\left(\lambda \boldsymbol{A}^T Y + w_1\beta\sum_i \boldsymbol{R}_i^T z_i + w_2\hat{\beta}\sum_i \boldsymbol{S}^T\hat{\boldsymbol{R}}_i^T\hat{z}_i\right).\end{aligned} \tag{4}$$

B. Stein's Unbiased Risk Estimate (SURE) and Scale Invariacne

The Stein Unbiased Risk Estimator (SURE) is utilized to find the optimal weights for different scales in (4) in an image adaptive manner. Let $h(Y, \theta)$ denotes the denoising of the noisy image Y with parameter-vector θ, the optimal parameter can be obtained by minimization of the expected error

$$MSE = \|h(Y,\theta)\|_2^2 - 2h(Y,\theta)^T Y + 2\sigma_0^2\nabla\cdot h(Y,\theta), \tag{5}$$

where $h(Y, \theta)$ represents the denoising of the noisy image Y and θ is a parameter-vector the algorithm depends on. Divergence calculation of $\nabla\cdot h(Y,\theta)$ is estimated by numerical approximation:

$$\nabla\cdot h(Y,\theta) = \frac{1}{\varepsilon}B^T\left(h(Y+\varepsilon B,\theta) - h(Y,\theta)\right), \tag{6}$$

where ε is a small constant.

Furthermore, to obtain the desired scale invariance property, the upper-bound of the Kullback-Leibler divergence between the empirical distribution and local model is adopted to tune the filter $\boldsymbol{H}$. Readers may refer to Ref [13] for detailed formula derivation when the Gaussian mixture model (GMM) model is employed for prior modeling of the original scale. For other scales, the filter is adjusted to keep the model unchanged. In this way, the MSEPLL algorithm manages to narrow the gap between global modeling while preserving the local treatment.

III. SOURCE CAMERA IDENTIFICATION BASED ON MSEPLL

In this section, we will first present the sensor pattern noise estimation method based on MSEPLL denoising algorithm, and then introduce the framework for source camera identification.

A. SPN estimation based on MSEPLL

The prominent denoising performance is expected to result in accurate sensor pattern noise estimation. Based on MSEPLL denoiser, the proposed SPN estimation method is consisted of three steps as follows:

Step1. Residual image estimation with MSEPLL denoiser. To suppress the content of images, the residual image R is obtained by

$$R = I - F(I), \tag{7}$$

where $F(\bullet)$ represents the denoising algorithm and in this case the MSEPLL. We followed the common setting in [3] that the standard deviation of noise $\sigma_0 = 3$ while denoising.

Step2. SPN estimation of single image with wiener filtering. Given an image, the SPN S is obtained by pixel-wise adaptive Wiener filtering is applied to residual image that:

$$S(i,j) = R(i,j)\frac{\sigma_0^2}{\hat{\sigma}^2(i,j) + \sigma_0^2}. \tag{8}$$

The local variance $\hat{\sigma}^2(i,j)$ can be estimated by the maximum likelihood estimation [2] as:

$$\hat{\sigma}^2(i,j) = \max\left(0, \frac{1}{M}\sum_{(p,q)\in N_M} R^2(p,q) - \sigma_0^2\right), \tag{9}$$

where M is the cardinality of in neighborhood N_M.

Step3. Camera SPN estimation by aggregation and SPN enhancement. Given a set of N images from the same camera instance c, the estimated SPN is obtained by averaging:

$$\overline{S} = \frac{1}{N}\sum_{k=1}^{N} S_k. \tag{10}$$

As revealed by systematic experimental evaluation, the simple averaging strategy leads to comparable performance to the MLE aggregation method.

In order to suppress artifacts caused by in camera operations inside such as CFA interpolation, the final SPN for camera c is estimated as

$$S_c = WF\left(ZM\left(\overline{S}\right)\right), \tag{11}$$

where ZM (•) represents the zero-mean operation of each row and column, and WF (•) denotes the 3×3 wiener filtering of in the Fourier domain.

B. Source camera identification

To determine if image under investigation is taken from a specific camera, we first compute correlation between noise residue and camera reference pattern, the SPN S_t is estimated by steps 1 and 2 as presented in Section 3.A for a given test image t. The normalized correlation coefficient (NCC) is adopted to measure the similarity between the test SPN S_t and reference SPN S_c of candidate cameras:

$$\rho_c = corr\left(S_t, S_c\right) = \frac{\left(S_t \cdot S_c\right)}{\left\|S_t\right\| \cdot \left\|S_c\right\|}. \tag{12}$$

By assigning test image t to the candidate camera that yields the largest NCC values, we can evaluate the performance in terms of accuracy.

Furthermore, the camera identification task can be formulated as a binary hypothesis problem that

$$\begin{aligned} &H_0\text{:Image is taken by camera } c \\ &H_1\text{: Image is not taken by camera } c \end{aligned}. \tag{13}$$

Let ρ denotes NCC values when image t is taken from camera c and $\overline{\rho}$ denotes those when t is not taken from camera c. If ρ is greater than the varying threshold, the null hypothesis is accepted, conversely, it is rejected. Based on comprehensive evaluation of NCC values, the overall receiver operating characteristic (ROC) curves can be studied. We shall see the detailed experimental results in both settings in the following section.

IV. EXPERIMENTAL RESULTS

A. Experimental setup

To verify the effectiveness of the proposed algorithm, we followed the experimental settings of [9] that 1200 natural images (200 images randomly selected from six cameras) are selected from the Dresden image database [15]. Detailed database information is given in Table I.

For each of the carmer, 25 randomly selected natural images are adopted for camera reference SPN estimation to mimic the actual forensic application. Experimental results are reported based on testing results from the remaining 175 images. In both SPN estimation and identification, only patches form center area is studied as they are more likely suffer from forgery modifications according to human photographing habits. The patch sizes are limited to 64×64, 128×128 and 256×256 throughout the paper. The proposed method is compared with the Basic method [1], the MLE method [3], the Model3 [4], and the GBWA method [9].

TABLE I DETAIL INFORMATION OF CAMERAS

Camera No.	Resolution	Device	No. of images in Dresden
No.1	3264×2448	Canon_Ixus55_0	224
No.2	3872×2592	Nikon_D200_1	380
No.3	3648×2736	Olympus_mju_1050SW_4	202
No.4	3648×2736	Panasonic_DMC-FZ50_1	415
No.5	3072×2304	Samsung_L74wide_0	232
No.6	3456×2592	Sony_DSC-H50_0	284

Two scales are utilized in MSEPLL denoising algorithm, where the filter in the first scale is a standard non-extracted one. The second scale adopts an 11×11 Gaussian filter with standard deviation and down-sampling factor of 2. The weight for second scale is fixed to 0.33. Neighborhood size M is set to 9 in (9) for local variance estimation.

B. Comparison of SPN Estimation Algorithms

Accurate estimated SPN less influenced by scene content is the essence for success camera identification. In order to clearly examine the quality of the camera fingerprint, SPN extracted by different methods are visualized in Fig.1.

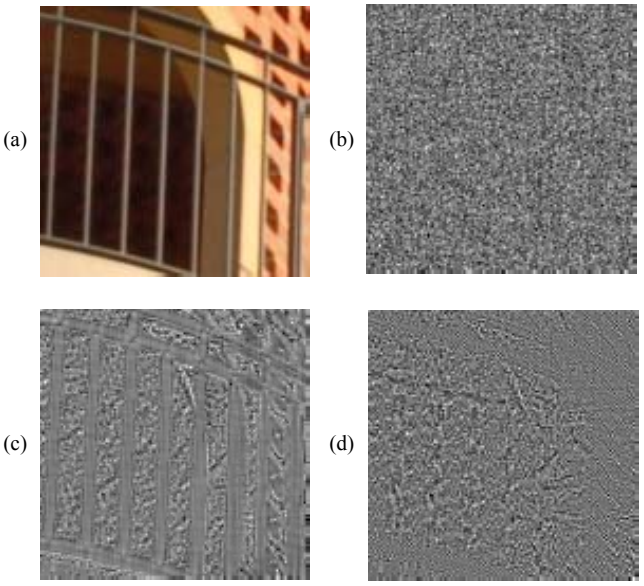

Fig.1 Comparison of SPN estimation methods. (a) Original image. (b) Reference SPN from flat-field images[15]. (c) SPN estimated by GBWA[9]. (d) SPN estimated by proposed method.

For the original image patch in Fig.1(a), the reference SPN is estimated by MLE method from 50 flat-field images[15] shown in Fig.1 (b) demonstrate good randomness like white Gaussian noise, which is hard to obtain in real camera identification application. The SPN estimated by using guided image estimation and block weighted average (GBWA)

method [9] given in Fig.1(c) is heavily contaminated that many scene content related details can be clearly observed, especially in texture and edge regions due to the inferior denoising ability of the guided filter. By imposing a multi-scale prior, method noise in SPN estimated from the proposed algorithm shown in Fig.1(d) has been significantly reduced. It is more similar to the 'ground truth' SPN in Fig.1(b) that is less influenced by scene content.

C. Accuracy Comparison of dfferent Methods

By counting the number of correctly judgments (True) and wrong judgments (False) for each test patch, the accuracy can be obtained as

$$Accuracy = \frac{True}{True + False}. \quad (14)$$

Experimental resutls for patch sizes of 64×64, 128×128 and 256×256 are given in Table II, III and IV, where the proposed method is compared with the state-of-the-art Basic method [1], the MLE method [3], the Model3 in [4], and the GBWA method [9]. Italic figures indicate the highest identification accuracy among the five methods.

TABLE II ACCURACY COMPARISON OF DIFFERENT METHODS FOR 64×64 PATCH SIZE

Methods	No.1 (%)	No.2 (%)	No.3 (%)	No.4 (%)	No.5 (%)	No.6 (%)	Average (%)
Basic[1]	48.57	69.14	44.00	80.57	43.43	88.00	62.29
MLE[3]	47.43	65.71	38.29	69.14	41.71	86.29	58.10
Model3[4]	36.57	66.86	35.43	76.00	36.00	77.71	54.76
GBWA[9]	41.14	83.43	**48.00**	78.29	41.14	89.14	63.52
Proposed	**52.00**	**98.86**	47.43	**96.57**	**52.57**	**98.29**	**74.29**

TABLE III ACCURACY COMPARISON OF DIFFERENT METHODS FOR 128×128 PATCH SIZE

Methods	No.1 (%)	No.2 (%)	No.3 (%)	No.4 (%)	No.5 (%)	No.6 (%)	Average (%)
Basic[1]	73.71	90.86	65.14	94.29	66.29	98.86	81.52
MLE[3]	73.71	96.00	52.57	91.43	58.29	**100**	78.67
Model3[4]	74.29	93.71	68.00	95.43	58.86	95.43	80.95
GBWA[9]	78.29	98.86	**76.57**	96.57	71.43	**100**	86.95
Proposed	**88.00**	**99.43**	68.57	**98.29**	**85.14**	**100**	**89.91**

TABLE IV ACCURACY COMPARISON OF DIFFERENT METHODS FOR 256×256 PATCH SIZE

Methods	No.1 (%)	No.2 (%)	No.3 (%)	No.4 (%)	No.5 (%)	No.6 (%)	Average (%)
Basic[1]	98.29	**100**	82.86	98.29	91.43	**100**	95.14
MLE[3]	93.71	**100**	80.00	98.86	86.86	**100**	93.24
Model3[4]	93.71	**100**	82.86	96.57	90.86	**100**	94.00
GBWA[9]	98.86	**100**	**85.71**	99.43	92.00	**100**	96.00
Proposed	**100**	**100**	84.00	**100**	**99.43**	**100**	**97.24**

We can observe consistent performance improvement over other methods for different patch sizes, besides the No. 3 camera Olympus_mju_1050SW_4 whose identification accuracy is slightly lower than the GBWA [9] method. Further examination reveals that there are more saturated regions in the Olympus images in Dresden database, where the block weighting mechanism in the GBWA [9] method achieves slightly better performance. However, averaged accuracy improvement over the sub-optimal GBWA method on the whole database is 10.77%, 2.96% and 1.24%, respectively. Furthermore, accuracy gain is more obvious for small patch size of 64 × 64, which is a very appealing property of the proposed method.

D. ROC Comparison of dfferent Methods

In order to provide comprehensive evaluation of the statistical performance of the proposed algorithm, ROC curves are investigated. Based on ρ and $\bar{\rho}$, number of true positive (TP), false positive (FP), true negative (TN) and false negative (FN) can be obtained at series of threshold values. Finally, the overall ROC curve is plotted of true positive rate (TPR) with respect to the false positive rate (FPR) defined as:

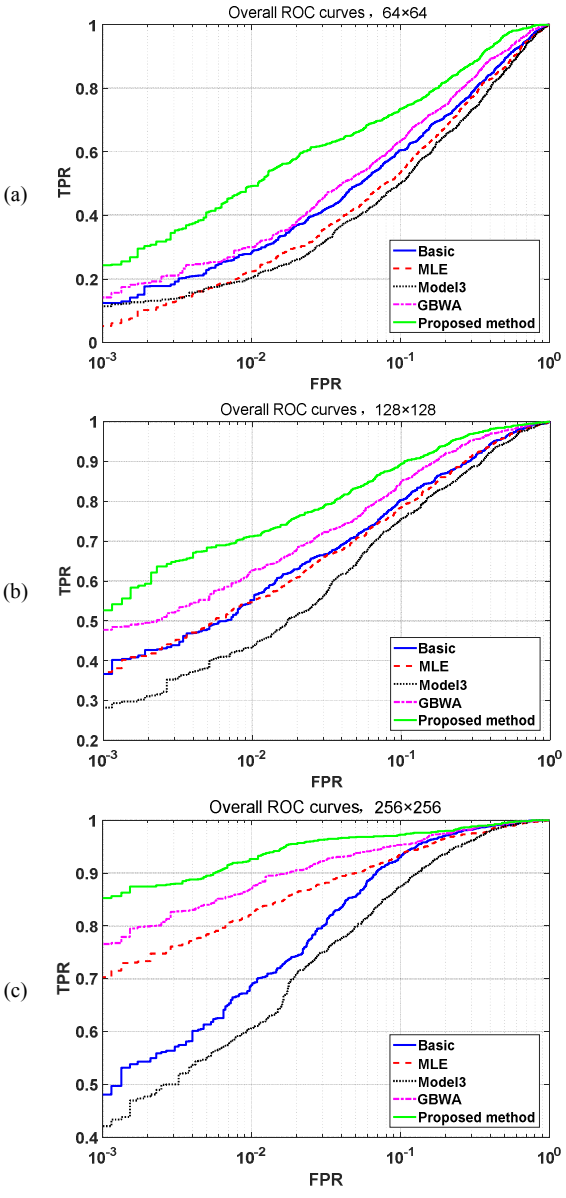

Fig. 2 Comparison of ROC curves for different patch sizes.

$$TPR = \frac{TP}{TP + FN}$$
$$FPR = \frac{FP}{FP + TN} \quad (15)$$

As shown in Fig. 2 that the performance improvement over other methods is obvious for all patch sizes considered. In addition, Table V depicts the TPR comparison results at small FPR (10^{-3}). Average TPR improvement over GBWA method is about 9.43% in comprehensive consideration of all patch sizes.

TABLE V TPR AT FPR = 10-3 FOR DIFFERENT CAMERA IDENTIFICATION METHOD.

Patch Size	Basic[1]	MLE[3]	Model3[4]	GBWA[9]	Proposed
64×64	12.38	5.14	11.43	14.19	**24.29**
128×128	36.67	34.95	33.62	47.71	**52.67**
256×256	48.10	72.19	42.10	72.00	**85.24**

V. CONCLUSIONS

In this paper, an effective sensor pattern noise estimation method is proposed based on MSEPLL denoising algorithm. With accurate prior modeling across different scales, denoised images are less influenced by image content, resulting more accurate SPN estimation. Effectiveness of the proposed method is verified with consistent accuracy increase in abundant source camera identification experiments. Performance improvement is more prominent for small patch sizes, which is an appealing property for future forensic applications [16] [17].

ACKNOWLEDGMENT

This work was supported by National Key Research and Development Program (No. 2018YFC0831105 and 2018YFC0831100), the National Nature Science Foundation of China (No. 61305015, No. 61203269), the National Natural Science Foundation of Shandong Province (No. ZR201702160307), and Shandong Province Key Research and Development Program, China (No. 2016GGX101022).

REFERENCES

[1] Lukáš, Jan, Jessica Fridrich, and Miroslav Goljan. "Digital camera identification from sensor pattern noise." IEEE Transactions on Information Forensics and Security 1.2 (2006): 205-214.

[2] Mihcak, M. Kivanc, Igor Kozintsev, and Kannan Ramchandran. "Spatially adaptive statistical modeling of wavelet image coefficients and its application to denoising." 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No. 99CH36258). Vol. 6. IEEE, 1999.

[3] Chen, Mo, et al. "Determining image origin and integrity using sensor noise." IEEE Transactions on information forensics and security 3.1 (2008): 74-90.

[4] Li, Chang-Tsun. "Source camera identification using enhanced sensor pattern noise." IEEE Transactions on Information Forensics and Security 5.2 (2010): 280-287.

[5] Al-Ani, Mustafa, and Fouad Khelifi. "On the SPN estimation in image forensics: a systematic empirical evaluation." IEEE Transactions on Information Forensics and Security 12.5 (2016): 1067-1081.

[6] Kang, Xiangui, et al. "A context-adaptive SPN predictor for trustworthy source camera identification." EURASIP Journal on Image and video Processing 2014.1 (2014): 19.

[7] He, Kaiming, Jian Sun, and Xiaoou Tang. "Guided image filtering." IEEE transactions on pattern analysis and machine intelligence 35.6 (2012): 1397-1409.

[8] Zeng, Hui, and Xiangui Kang. "Fast source camera identification using content adaptive guided image filter." Journal of forensic sciences 61.2 (2016): 520-526.

[9] Zhang, Le-Bing, Fei Peng, and Min Long. "Identifying source camera using guided image estimation and block weighted average." Journal of Visual Communication and Image Representation 48 (2017): 471-479.

[10] Chan, Lit-Hung, Ngai-Fong Law, and Wan-Chi Siu. "A confidence map and pixel-based weighted correlation for PRNU-based camera identification." Digital Investigation 10.3 (2013): 215-225.

[11] Lin, Xufeng, and Chang-Tsun Li. "Preprocessing reference sensor pattern noise via spectrum equalization." IEEE Transactions on Information Forensics and Security 11.1 (2015): 126-140.

[12] Marra, Francesco, et al. "Blind PRNU-based image clustering for source identification." IEEE Transactions on Information Forensics and Security 12.9 (2017): 2197-2211.

[13] Papyan, Vardan, and Michael Elad. "Multi-scale patch-based image restoration." IEEE Transactions on image processing 25.1 (2015): 249-261.

[14] Geman, Donald, and Chengda Yang. "Nonlinear image recovery with half-quadratic regularization." IEEE transactions on Image Processing 4.7 (1995): 932-946.

[15] Gloe, Thomas, and Rainer Böhme. "The'Dresden Image Database'for benchmarking digital image forensics." Proceedings of the 2010 ACM Symposium on Applied Computing. Acm, 2010.

[16] Valsesia, Diego, et al. "Compressed fingerprint matching and camera identification via random projections." IEEE Transactions on Information Forensics and Security 10.7 (2015): 1472-1485.

[17] Mohanty, Manoranjan, et al. "e-PRNU: Encrypted Domain PRNU-Based Camera Attribution for Preserving Privacy." IEEE Transactions on Dependable and Secure Computing (2019).