https://doi.org/10.1016/j.diin.2017.07.005

A Local Variance Based Approach to Alleviate the Scene Content Interference for Source Camera Identification

Chao Shi, Ngai-Fong Law, Frank H. F. Leung, and Wan-Chi Siu

Abstract—**Identifying the source camera of images is becoming increasingly important nowadays. A popular approach is to use a type of pattern noise called photo-response non-uniformity (PRNU). The noise of image contains the patterns which can be used as a fingerprint. Despite that, the PRNU-based approach is sensitive towards scene content and image intensity. The identification is poor in areas having low or saturated intensity, or in areas with complicated texture. The reliability of different regions is difficult to model in that it depends on the interaction of scene content and the characteristics of the denoisng filter used to extract the noise. In this paper, we showed that the local variance of the noise residual can measure the reliability of the pixel for PRNU-based source camera identification. Hence, we proposed to use local variance to characterize the severeness of the scene content artifacts. The local variance is then incorporated to the general matched filter and peak to correlation energy (PCE) detector to provide an optimal framework for signal detection. The proposed method is tested against several state-of-art methods. The experimental results show that the local variance based approach outperformed other state-of-the-art methods in terms of identification accuracy.**

I. Introduction

DUE to the popularity of image capturing devices, digital images could be presented as evidence in the court. Reliable identification of digital image origins will be very useful for law enforcement. For example, in the cases involving illegal images, identification of the possessor of the source camera taking these illegal images would be important in the investigation. Clues from different stages of the image acquisition process can be used to identify the source camera of the query image. The simplest approach is to examine the electronic file itself and look for clues in headers or other associated information. One example is the Exchangeable Image File (EXIF) header that contains information about the digital camera type and the conditions under which the image was taken. However, these data may not be available if the image is resaved in a different format or recompressed. Besides, the header information can be easily modified which undermines the credibility of this method. Another approach of source identification is to embed digital watermark in the image that contains information about the digital camera (Blythe and Fridrich 2004, 11-13)(Wong 1998, 374-379). However, adding watermarking feature in cameras incurs additional cost. Most cameras also do not support this feature (Sencar and Memon 2008, 325-348).

A feature based approach which utilizes the artifacts produced by image processing operations for source identification was considered in (Mehdi, Sencar, and Memon 2004, 709-712 Vol. 1). Features were extracted from digital images which were then used to train the support vector machine (SVM) for source identification. However, the identification accuracy was limited. Pixel defects can also be employed to identify the camera source. This method, however, has limitation because defective pixels may not exist or may be eliminated by post processing. Recently, a powerful approach based on the photo-response non-uniformity (PRNU) has been proposed (Lukas, Fridrich, and Goljan 2006, 205-214). The PRNU provides unique characteristics of a specific digital camera. It arises from the imperfection during the manufacturing of the silicon wafer in the imaging sensor. The imperfection makes each pixel respond differently to the illumination of light (Li and Li 2010, 3052-3055). Artifacts resulting from the pixel by pixel difference will then exist in all images taken by that particular camera. This unique artifact called PRNU can be used as the fingerprint of a specific camera. Since the PRNU is different from one camera to another, it can identify not only camera models but also individual cameras of the same model. Moreover, the PRNU signal remains more or less unchanged after operations like lossy compression, cropping, printing, downsizing etc. (Fridrich 2009, 26-37) (Rosenfeld and Sencar 2009, 72540M-72540M-7) (Goljan, Fridrich, and Lukáš 2008, 68190I-68190I-12) (Alles, Geradts, and Veenman 2008, 557-567).

To obtain the PRNU signal, denoising is applied to those images under test and the resultant noise residual forms the PRNU signal. One problem of using PRNU for source camera identification is that the scene content can severely contaminate the extracted PRNU. If images contain a lot of textures, have low or saturated intensities, the accuracy of identification will drop. Several methods have been proposed to suppress the influence of scene content. In (Chen, Fridrich, and Goljan 2007, 65050P-65050P-13), a maximum likelihood method is proposed to estimate the camera reference PRNU. Kang et al. (Kang et al. 2012, 393-402) proposed to whiten the PRNU in the frequency domain to estimate the reference PRNU. The approaches which make use of the idea of reliable regions have been proposed in (Li 2010, 280-287) (McCloskey 2008, 1-6) (Chan, Law, and Siu 2013, 215-225) (Shi et al. 2014, 1-7). In (Chan, Law, and Siu 2013, 215-225), (Shi et al. 2014, 1-7),

learning based methods are adopted, in which a training phase is required before camera identification. In (Li 2010, 280-287), Li made a hypothesis that the stronger a signal component is, the more likely that it is associated with strong scene details, and thus the less trustworthy the component should be. Based on this hypothesis, Li proposed five different models to shrink the noise residuals with high magnitude. However, there is no proof showing that the five models give the optimal weightings.

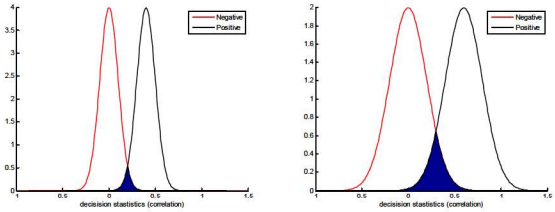

Fig. 1. Illustration of aggressive weighting allocation problem. The left figure shows the distribution of the positive and negative samples when no weighting is applied. The right figure shows distribution when the over aggressive weightings are applied to the test.

Although methods that reduced the scene content influence has been proposed, it is very important to allocate appropriate weightings to different regions of the image. If the weightings are assigned too aggressively, for example assigning a weighting of 1 to the smoothest regions and 0 to other regions, the decision statistics obtained will be highly unstable because only small portion out of the whole image is utilized for the detection. On the contrary, if the weightings are assigned too conservatively, the improvement of accuracy will not be significant. This problem is illustrated in Figure 1. Figure 1 shows the distribution of the decision statistics for positive and negative images. Figure 1 (a) is the distributions when no weighting is applied and (b) is the distributions when an over aggressive weighting is adopted. In the case of (b), although the mean correlation for matching case becomes larger, the chance of false detection is higher because the standard deviation of correlation becomes larger which results in a larger overlapping between the positive and negative situations. This example illustrates the intuitive idea that assigning larger weights to the reliable regions may fail to improve the detection accuracy. Hence, determining the optimal weights that give the highest detection accuracy is not a straight forward task.

In this paper, on one hand, we study how the noise residual strength affects the source camera detection accuracy. On the other hand, we proposed a method to obtain the weighting for each pixel based on the general matched filter which has been proved to be the optimal detector. The rest of the paper is organized as follows. In section II, the use of PRNU for source camera identification is reviewed. Then the motivation and details of the proposed algorithm is presented in section III. Experimental results and analysis are given in section IV. Finally, section V concludes this paper.

II. Camera Identification with PRNU

The general PRNU-based camera identification approach includes two phases: camera fingerprint construction and testing. In the first phase, the PRNU signal for each source camera was constructed. In the testing phase, the noise residue signal from the testing image was extracted as in the first phase. Correlation between the noise residue from the testing image and the camera PRNU was then calculated to decide whether the testing image was taken from that particular camera.

A. PRNU extraction from images

The PRNU-based approach was proposed by Lukas and Fridrich in (Lukas, Fridrich, and Goljan 2006, 205-214). To decide whether an image was taken by the reference camera or not, a set of images from the reference camera was obtained. These images were then denoised using a denoising filter to generate a set of noise residues. The PRNU of a camera C (called the reference PRNU), denoted as $\hat{\mathbf{K}}_C$, was obtained by taking the average of these noise residues, i.e.,

$$\hat{\mathbf{K}}_C = 1/N \sum_{k=1}^{N} \mathbf{W}_k, \tag{1}$$

where $\mathbf{W}_k$ is the noise residue from the kth image and N is the number of images used to estimate the PRNU. Instead of using a simple average, a maximum likelihood estimator has been proposed to estimate the PRNU (Chen, Fridrich, and Goljan 2007, 65050P-65050P-13) (Chen et al. 2007, 342-358). Under the maximum likelihood framework, the reference PRNU was obtained as,

$$\hat{\mathbf{K}}_C = \sum_{k=1}^{N} (\mathbf{I}_k \mathbf{W}_k) / \sum_{k=1}^{N} (\mathbf{I}_k)^2, \tag{2}$$

where $\mathbf{W}_k$ and $\mathbf{I}_k$ are the noise residue and intensity of the kth image respectively. The product operation among vectors is elementwise in this paper. The camera reference PRNU for a testing image p is calculated as $\mathbf{I}_p \hat{\mathbf{K}}_C$ where $\mathbf{I}_p$ is the intensity of the testing image p. To remove the periodical structure artifacts due to color interpolation, sensor design and on-sensor signal transfer, Chen and Fridrich et al. (Chen et al. 2008, 74-90) proposed to perform a zero-mean operation to the camera reference PRNU.

B. Correlation detector

To determine whether a testing image p was obtained from a particular reference camera C, the correlation detector is used. It is defined as,

$$\rho_C(p) = corr(\mathbf{W}_p, \hat{\mathbf{K}}_C) = \frac{(\mathbf{W}_p - \overline{\mathbf{W}_p}) \cdot (\mathbf{I}_p \hat{\mathbf{K}}_C - \overline{\hat{\mathbf{K}}_C})}{\left\| \mathbf{W}_p - \overline{\mathbf{W}_p} \right\| \left\| \mathbf{I}_p \hat{\mathbf{K}}_C - \overline{\hat{\mathbf{K}}_C} \right\|} \tag{3}$$

where $\mathbf{W}_p$ is the noise residue of the testing image p, the bar above the symbol denotes its mean value, $\cdot$ denotes the dot product and $\|.\|$ denotes the L_2 norm. The correlation value measures the similarity between the noise residual extracted from the testing image and the camera reference PRNU. A high correlation value implies a high chance that the testing image p is taken by the reference camera C. If the reference PRNU is

estimated from equation (2), the correlation detector will become,

$$\rho_C(p) = corr(\mathbf{W}_p, \mathbf{I}_p\hat{\mathbf{K}}_C) = \frac{(\mathbf{W}_p - \overline{\mathbf{W}_p})\cdot(\mathbf{I}_p\hat{\mathbf{K}}_C - \overline{\mathbf{I}_p\hat{\mathbf{K}}_C})}{\left\|\mathbf{W}_p - \overline{\mathbf{W}_p}\right\|\left\|\mathbf{I}_p\hat{\mathbf{K}}_C - \overline{\mathbf{I}_p\hat{\mathbf{K}}_C}\right\|} \quad (4)$$

III. The Proposed Method

A. Motivation

According to Chen and Fridrich et al. (Chen et al. 2008, 74-90), the noise residual can be expressed as,

$$\mathbf{W} = \mathbf{I} - F(\mathbf{I}) = \mathbf{TIK} + \Xi. \quad (5)$$

where Ξ contains a combination of other independent noise and scene content artifact due to the imperfection of the denoising filter $F(.)$ and $\mathbf{T}$ is a attenuation factor which indicates how much the PRNU signal is retained in $\mathbf{W}$ because the denoising filter may remove part of the PRNU signal . In the source camera identification problem, we may refer the term $\mathbf{IK}$ as the signal of interest and Ξ as the undesired noise. Because of the scene content artifacts, Ξ is not stationary within images. In areas with large texture, the noise Ξ will likely have large variance.

To resolve the scene content problem, Li (Li 2010, 280-287) has proposed five models to suppress the noise residual whose magnitude is large for the reason that the noise with larger magnitude is more likely to be contaminated by the scene content artifact. From equation (5), it can be seen that Li's (Li 2010, 280-287) assumption is plausible as the scene content artifact may lead to a large value for Ξ and if $\mathbf{IK}$ is much smaller than Ξ, the noise residual $\mathbf{W}$ will be large as well. If we examine the correlation detector (3) or (4) carefully, we can find that the element with large magnitude in $\mathbf{W}$ contributes more to the resultant correlation value but from previous analysis large noise residual may not be reliable. Therefore attenuating large values in noise residual should be an effective method to compensate for the scene content problem. However, from equation (5), the large magnitude of $\mathbf{W}$ may also arise from the strong PRNU signal. That means a large $\mathbf{IK}$ may also produce a large value of $\mathbf{W}$ even when it does not contain any scene content.

Motivated by this, for each pixel, instead of using its own value, we propose to utilize its neighbor pixels to estimate the amount of scene content artifacts. Let I_{ij} be the intensity value of the pixel (i,j) in an image. The local variance and the mean of the noise residual can be defined as

$$\sigma^2_{\text{local}ij} = \frac{1}{|A|}\sum_{(m,n)\in A}\left(W_{mn} - \overline{W}_{\text{local}ij}\right)^2, \quad (6)$$

$$\text{where} \quad \overline{W}_{\text{local}ij} = \frac{1}{|A|}\sum_{mn\in A} W_n$$

where A is the neighbor of the pixel I_{ij} and $\overline{W}_{\text{local}}$ is the mean of the noise residual within A. As scene content artifact usually occurs in the complicated texture areas, the noise residual with scene content artifact is likely to have large variance. Therefore, for each pixel, the variance of its neighbor pixels is calculated. The variance value can be used as a measure of the scene content artifacts. Since the local variance is estimated from a set of pixels with similar textures, it is a more reliable estimate of the amount of scene content artifact than the noise residual magnitude of a single pixel.

Experiments were conducted to reveal the effectiveness of the local variance on characterizing the scene content artifact. A total of 300 natural images taken by six cameras from the Dresden image database (Gloe and Böhme 2010, 150-159) are selected. The camera detail is summarized in Table 1. Each image is cropped into a size of 256×256. Another 50 images are used to estimate the reference PRNU $\mathbf{K}$ of each camera using equation (2). For each pixel, the product of the noise residual and the PRNU is calculated, i.e. $C_{ij} = W_{ij} \bullet I_{ij}K_{ij}$ which can be seen as the covariance between the noise residual and PRNU for that pixel because the mean of W_{ij} and K_{ij} over an image will approach 0 if the image size is large. To study the relationship between C_{ij} and the noise residual magnitude $|W_{ij}|$, the domain of $|W_{ij}|$ is divided into 20 equal intervals in which the mean value μ and standard deviation σ of C_{ij} in each interval is calculated. The result is shown in Figure 2(a), where the red dots represent the mean value μ and the red bars are of the length of 0.1σ. The standard deviation is scaled in the figure because it is too large compared with μ. A large mean value indicates that the strength of the PRNU signal is strong for the pixels in that bin. On the other hand, if the standard deviation of the covariance is large, the undesired noise like the scene artifact associated with that pixel should be large since the undesired noise is independent to the PRNU signal. Hence, a reliable pixel should have a large mean covariance and small standard deviation. In Figure 2(a) it can be observed that both

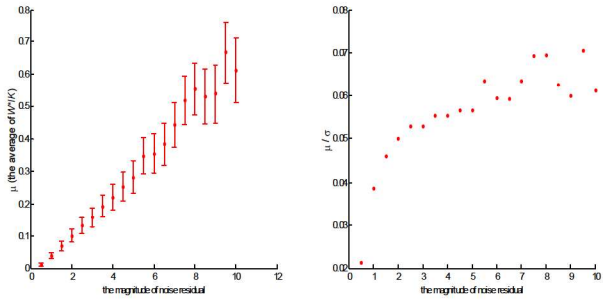

Fig. 2. The relation between the covariance and the magnitude of noise residual

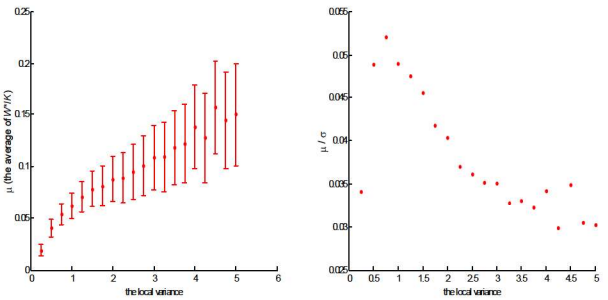

Fig. 3. The relation between the covariance and the local variance

the mean value μ and the standard deviation σ increase with the magnitude of noise residual $|W_{ij}|$. Hence the reliability of pixels can be hardly decided from Figure 2(a). Therefore, Figure 2(b) plots the signal to noise ratio μ/σ which can better measure the reliability of the data. From Figure 2(b), it can be seen that μ/σ will not increase with $|W_{ij}|$ when $|W_{ij}|$ becomes large. Similarly, the relation between C_{ij} and the local standard deviation $\sigma_{\text{local},ij}$ are plotted in Figure 3. Although Figure 3 (a) is similar to Figure 2 (a) in that both the mean value μ and standard deviation σ increase with the local variance, it is noticeable that the value of μ/σ drops with σ_{local} increasing which indicates that when σ_{local} gets larger, the corresponding pixel is less reliable. Therefore $\sigma_{\text{local},ij}$ should be a more sensitive measure for the reliability of the pixel. Hence the local variance can be used as a measure of the severeness of the scene content artifacts. In the next section we will demonstrate how to weight the pixels according to the local variance.

TABLE 1
CAMERA DETAILS FOR THE EXPERIMENT.

Camera model	Number	Sensor	Resolution	Format
Canon Ixus70	2	1/2.5"	3072 ×2304	JPEG
Nikon Coolpix S710	2	1/1.72"	4352×3264	JPEG
Sony DSC-H50	2	1/2.3"	3456×2592	JPEG

B. The Proposed Method

The source identification problem is formulated in the framework of hypothesis testing. In the general signal detection problem, the two hypothesis are defined as,

$$\mathrm{H}_0: x_j = w_j$$

and

$$\mathrm{H}_1: x_j = s_j + w_j$$

where x_j is the noisy signal observed, s_j is the noise free signal to be detected, w_j is the white Gaussian noise (WGN) and j is the signal index. For such signal detection problem, the optimal detector will be the general matched filter (Turin 1976, 1092-1112) (Kay 1998) in that it gives the largest True Positive Rate for any given False Positive Rate. The general matched filter can be expressed as,

$$y = \sum_{j=1}^{N} s_j x_j / \sigma_j^2 \tag{7}$$

where σ_j^2 is the expected variance of w_j, N is the length of the signal and y is the output of the filter. Under the problem of source camera identification, the noisy data observed is noise residual W_{ij} of pixel I_{ij}, the noise free signal to be detected is $T_{ij}I_{ij}K_{ij}$ and the noise of the observed data is $\Xi_{i,j}$ (Chen et al. 2008, 74-90). Hence, the detection problem will become,

$$\mathrm{H}_0: x_{ij} = \Xi_{ij}$$
$$\mathrm{H}_1: x_{ij} = T_{ij}I_{ij}K_{ij} + \Xi_{ij}$$

Therefore the optimal detector will be

$$\begin{aligned} y &= \sum_{i=1,j=1}^{i=M,j=N} T_{ij}I_{ij}K_{ij}W_{ij} / \sigma_{ij}^2 \\ &= \sum_{i=1,j=1}^{i=M,j=N} (T_{ij} / \sigma_{ij}^2) I_{ij}K_{ij}W_{ij} \end{aligned} \tag{8}$$

From (8), it can be seen that to achieve the optimal detection performance, the value of attenuation factor T_{ij} and Ξ_{ij} must be known. However, their values cannot be estimated easily, due to the fact that they depend on the scene content and the characteristic of the denoising filter. Therefore, usually the cross-correlation or the peak to correlation energy (PCE) is used as detection filter instead of general matched filter. Nevertheless, neither cross-correlation nor PCE is a good detector for the reason that they do not take the distinct distribution of each pixel into consideration. Hence, the distribution of the test statistics is difficult to be modeled due to the fact that the expected variance of PCE and cross correlation vary from one image to another.

The value of T_{ij} and Ξ_{ij} can be calculated by learning based approached as in (Chen et al. 2008, 74-90). However, the learning based approach requires a large data set for training and the parameters value will be camera model dependent. Therefore, we propose a simple method from which we can estimate the distribution parameter from the single image and take advantage of the optimal detector. Since Ξ_{ij} contains the scene content artifacts, σ_{ij}^2 is scene content dependent. Although $\Xi_{i,j}$ is not stationary due to the variation of scene content artifact, we may assume Ξ_{ij} to be locally stationary because of the local similarity property of images. Thereby, for each pixel I_{ij}, we calculate the local variance $\sigma_{\text{local},ij}^2$ from equation (6). Given Ξ_{ij} is stationary around pixel I_{ij} and Ξ_{ij} is independent to the PRNU signal, we will have,

$$\sigma_{\text{local},ij}^2 = \text{var}(\mathbf{W}_{\text{local}}) \approx \sigma_{ij}^2 + \sigma_{PRNU,ij}^2 \tag{9}$$

where σ_{ij}^2 is the variance of Ξ_{ij} and $\sigma_{PRNU,ij}^2$ is the variance due to the PRNU.

$$\sigma_{PRNU,ij}^2 = I_{ij}^2 T_{ij}^2 \text{var}(\mathbf{K}) \tag{10}$$

where $\text{var}(\mathbf{K})$ is the variance of the PRNU factor **K.** Combining equation (9) and (10), the variance of Ξ_{ij} can be calculated as,

$$\sigma_{ij}^2 = \sigma_{\text{local},ij}^2 - I_{ij}^2 T_{ij}^2 \text{var}(\mathbf{K}) \tag{11}$$

The PRNU **K** can be considered stationary over the image and $\text{var}(\mathbf{K})$ should be constant in the image. Though the value of $\text{var}(\mathbf{K})$ varies from one sensor to another, it is small as compared with σ_{ij}^2. Therefore, the value of $\text{var}(\mathbf{K})$ does not have significant influence on the estimation of $\sigma_{\text{local},ij}^2$ and an accurate estimate of $\text{var}(\mathbf{K})$ is not necessary. To calculate the optimal weighting we still need to know the shaping factor **T**. However, since the noise residual is obtained from a very complex process, it is difficult to obtain the value of **T.** In this paper, to simplify the problem, we assume the shaping factor to

be equal to 1 over the image. Whereas, it is worth to mention that any model for estimating **T** can be incorporated into our model. Note σ_{ij}^2 cannot be negative and it cannot be 0. Even at smooth regions, Ξ_{ij} contains the random noise like the quantization noise, σ_{ij}^2 cannot be zero and must be positive. Hence a lower bound for σ_{ij}^2 is set. Therefore,

$$\sigma_{ij}^2 = \max\left(\sigma_{\text{local},ij}^2 - I_{ij}^2 T_{ij}^2 \text{var}(\mathbf{K}), b_l\right) \tag{12}$$

where b_l is the lower bound for σ_k^2. Assuming that σ_k^2 is Gaussian distributed, the following can be calculated,

$$\text{E}(\text{y} : \text{H}_0) = 0 \tag{13}$$

$$\text{E}(\text{y} : \text{H}_1) = \sum_{i=1,j=1}^{i=M,j=N} \frac{(T_{ij} I_{ij} K_{ij})^2}{\sigma_{ij}^2} \tag{14}$$

$$\text{var}(\text{y} : \text{H}_0) = \text{var}(\text{y}, : \text{H}_1) = \sum_{i=1,j=1}^{i=M,j=N} \frac{(T_{ij} I_{ij} K_{ij})^2}{\sigma_{ij}^2} \tag{15}$$

where $\text{E}(\text{y} : \text{H}_0)$ and $\text{E}(\text{y} : \text{H}_1)$is the expectation of y under H_0 and H_1 respectively and $\text{var}(\text{y} : \text{H}_0)$ and $\text{var}(\text{y} : \text{H}_1)$ are the variances. The general matched filter is normalized by $\text{std}(\text{y}, : \text{H}_0)$ such that all the detection statistics will have variance equal to 1. Therefore, the normalized general matched filter will be

$$y = \sum_{j=1}^{N} (T_j / \sigma_j^2) I_j K_j W_j \Big/ \sqrt{\sum_{j=1}^{N} \left(T_j I_j K_j\right)^2 / \sigma_j^2} \tag{16}$$

However, in practice, there are some weak correlations between the PRNU fingerprints estimated from different cameras due to the JPEG compression, color interpolation etc. The cameras from the same manufacturer usually have higher correlations than those from different manufacturers. This correlation is likely to increase the False Acceptance Rate since such inter camera correlation might be detected as positive samples. The general matched filter does not take this situation into consideration. The Peak to Correlation Energy (PCE) does well in reducing the False Acceptance Rate caused by inter camera correlation. The PCE is defined as,

$$PCE(\mathbf{x},\mathbf{y}) = \frac{\rho(\mathbf{s}_{peak} = 0, \mathbf{x}, \mathbf{y})^2}{\frac{1}{MN - |A|} \sum_{s \notin A} \rho(s, \mathbf{x}, \mathbf{y})^2} \tag{17}$$

where $\rho(\mathbf{s}, \boldsymbol{x}, \boldsymbol{y})$ is the dot product between $\boldsymbol{x} - \bar{\boldsymbol{x}}$ and $\boldsymbol{y}(\text{s}) - \bar{\boldsymbol{y}}$, $\boldsymbol{y}(\text{s})$ is obtained by circularly shift $\boldsymbol{y}$ by a two dimension vector **s**, A is a small neighbor around the peak and MN is the width and height of the image. Since the image is assumed not to be scaled or translated, the peak should be at the location 0. The nominator of the PCE is the square of the correlation between signals $\boldsymbol{x}$ and $\boldsymbol{y}$ without normalization. The correlation can then be normalized by the denominator which is the square of the correlation value for all the possible shift vector **s** except for the shift close to the original position. Owing to the fact that the patterns that are shared among different cameras usually have a periodic nature, the correlations between cameras sensor will still exist even though the images are shifted. Hence, if there is a strong pattern that is shared between the two cameras tested, the denominator of PCE is likely to be high and the PCE value will be lowered. Thereby, the chance of False Acceptance can be reduced. To utilize this great feature of PCE, we propose to incorporate the general match filter into the PCE detector. The unnormalized general matched filter in equation (8) can be rewritten as,

$$\begin{aligned} y &= \sum_{i=1,j=1}^{i=M,j=N} \frac{T_{ij}}{\sigma_{ij}^2} I_{ij} K_{ij} W_{ij} \\ &= \sum_{i=1,j=1}^{i=M,j=N} \left(\frac{1}{\sigma_{ij}} W_{ij} \cdot \frac{T_j}{\sigma_{ij}} I_{ij} K_{ij} \right) \end{aligned} \tag{18}$$

The first term in the equation is the noise residual term which is normalized by the standard deviation of the undesired noise. The second term is the reference PRNU signal which is also normalized by the standard deviation of the undesired noise. We proposed to use the normalized signal as the input of the PCE detector. Hence, the final detector can be expressed as,

Fig. 4. Local variance estimated using square window with equal weight. Left: The original photo. Right: The local variance map computed using a squared window

$$PCE(\mathbf{x},\mathbf{y}) = \frac{\rho(\mathbf{s}_{peak} = 0, \mathbf{x}, \mathbf{y})^2}{\frac{1}{MN - |A|} \sum_{s \notin A} \rho(s, \mathbf{x}, \mathbf{y})^2}, \quad \text{where } \mathbf{x} = \frac{\mathbf{W}}{\boldsymbol{\sigma}} \text{ and } \mathbf{y} = \frac{\mathbf{TIK}}{\boldsymbol{\sigma}} \tag{19}$$

Note all the variables are written in matrix from.

C. *Local variance Estimation*

Since the weighting of each pixel is determined by the local variance in the proposed model, it is very important to estimate the local variance precisely. An accurate estimation of the local variance will improve the overall performance of the source identification. A square window with equal weight is used to estimate the local variance from the noise residual. The resultant local variance is likely to be discontinuous with many blocks as shown in Figure 4. A better option is to use a

Gaussian kernel as weighting to estimate the local variance. The local variance estimated with a Gaussian kernel is expressed as,

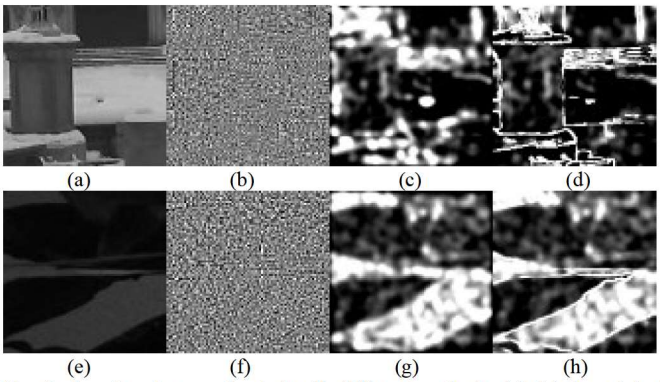

Fig. 5. Local variance estimated with different methods. (a), (e) the original image. (b), (f) the noise residual. (c), (g) the local variance estimated with the Gaussian kernel. (d), (h) the local variance estimated with bilateral kernel.

$$\sigma^2_{\text{local},i,j} = \frac{1}{|A|}\sum_{(m,n)\in A} h_G(m,n,i,j)(W_{n,m} - \overline{W}_{\text{local},i,j})^2 \quad (20)$$

$$h_G(m,n,i,j) = \frac{1}{\sqrt{2\pi}\sigma_G} e^{\frac{-\mathrm{D}(m,n,i,j)^2}{2\sigma_G^2}} \quad (21)$$

where function D(*m,n,i,j*) is the distance between pixel (*m,n*) and (*i,j*) and σ_G is the standard deviation parameter for the kernel. With the Gaussian kernel, pixels that are far from the center will be assigned small weightings for estimating the local variance. Figure 5 (c) and (g) show the local variance estimated from the Gaussian kernel for two images. However, it is found that the large variance of the noise residual often appeared along the edges of the original image. The Gaussian kernel tends to underestimate this kind of local variance because the local variance along the edge is averaged by its nearby pixels. In order to precisely estimate the local variance, the bilateral kernel can be used. The bilateral kernel (Tomasi and Manduchi 1998, 839-846) penalizes the pixels that have large variation compared with the center pixel such that they will have smaller weightings. By using the standard deviation map of the original image as guide image, the joint bilateral filtering can be used to estimate the local variance of the noise residual. The joint bilateral kernel (Caraffa, Tarel, and Charbonnier 2015, 1199-1208) is defined as,

$$\sigma^2_{\text{local},i,j} = \frac{1}{|A|}\sum_{(m,n)\in A} h_{\mathbf{B}}(m,n,i,j)(W_{n,m} - \overline{W}_{\text{local},i,j})^2 \quad (22)$$

$$h_B(m,n,i,j) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-\mathrm{D}(m,n,i,j)^2}{2\sigma_s^2}} e^{\frac{-(g(m,n)-g(i,j))^2}{2\sigma_r^2}} \quad (23)$$

where σ_s and σ_r are the spatial and range parameters of the bilateral filter which determines how sensitive the filter is towards the change of distance and value difference to the center pixel respectively and $g(.)$ is the local standard deviation of the original image for a given position. Since the variance for the edge region is large while that for the smooth region is small, with the guided bilateral filter, the local variance will be estimated along the edges if they are presented in the image. Figure 5(d) and (h) show the local variance estimated with the guided bilateral filter. It can be observed that, compared with the estimation of Gaussian filter, the edge regions have a larger local variance which is closer to the expected value. The joint bilateral filter can be implemented efficiently using the method described in (Yang, Tan, and Ahuja 2009, 557-564).

In summary, the proposed algorithm for source camera identification will be as follows,

1. Estimate the reference PRNU with equation (2).
2. Extract the Noise residual **W** with equation (5).
3. Preprocess the reference PRNU and noise residual with zero mean operation and Wiener Filter in frequency domain.
4. Compute the local variance $\sigma^2_{\text{local},ij}$ with equation (22) and (23) from the extracted noise residual.
5. Estimate the variance of the undesired noise σ^2_{ij} with equation (12).
6. Compute the detection statistics with equation (19).
7. Make decision based on the detection statistics obtained.

IV. Experimental Result

A. Experimental setup

Images from the Dresden Image Database (Gloe and Böhme 2010, 150-159) are used as testing data for the experiments. Nineteen different cameras are randomly selected from the database for testing. The cameras covered a wide range of brands and models. For most of the camera models selected, two cameras are used to test the distinguishing ability within the same model. For the cameras of the same make or model, similar post processing operations like JPEG compression and color interpolation are applied. Therefore, there should be certain correlations among the noise residual of these devices and it is more difficult to distinguish photos from them. Table 2 summarizes the details about the cameras under test, including the camera model and the number of cameras in each model. The images cover a large variety of scenes and they are taken with different camera settings such as different ISO and different focal length which makes the source camera identification more difficult. Example photos are shown in Figure 6. For each camera, 50 images are selected randomly to estimate the reference PRNU, and another 50 images are used as testing image. Totally 950 images are tested. The color images are converted to gray level image before PRNU extraction.

As the identification accuracy is usually very high for large images, it is difficult to compare the performance of different algorithms. To obtain a clear idea about the performance of different algorithms, the size of the image is cropped to 256 × 256 and 128 × 128. The experiment is conducted only on the two image sizes respectively.

TABLE 2 CAMERA DETAILS FOR THE EXPERIMENT

Device ID	Camera model	Sensor	Resolution	Format	No. of Devices
1	Canon Ixus 55	1/2.5"	2592 × 1944	JPEG	1
2, 3	Canon Ixus 70	1/2.5"	3072 × 2304	JPEG	2
4, 5	Casio_EX-Z150	1/2.5"	3264 × 2448	JPEG	2
6, 7	Nikon CoolPix S710	1/1.72"	4352 × 3264	JPEG	2

Fig. 6. Examples of Test Photos from Dresden Image Database

8, 9	Nikon_D200	23.6×15.8	3872 × 2592	JPEG	2
10, 11	Olympus 1050 SW	1/2.33"	3648 × 2736	JPEG	2
12, 13	Panasonic DMC FZ50	1/2.5"	3648 × 2736	JPEG	2
14, 15	Samsung NV15	1/2.5"	3648 × 2736	JPEG	2
16, 17	Sony DSC T77	1/2.5"	3648 × 2736	JPEG	2
18, 19	Sony DSC W170	1/2.5"	3648 × 2736	JPEG	2

B. Experiment Methodology

The performance of our proposed method is compared with the basic algorithm (Lukas, Fridrich, and Goljan 2006, 205-214), the MLE method (Chen et al. 2008, 74-90), the phase pattern noise method (Kang et al. 2012, 393-402) and Li's enhancement pattern noise method (Li 2010, 280-287). For Li's method, model 3 is selected for comparison because of their superior performance as compared with other models in (Li 2010, 280-287). To extract the noise residual **W**, the wavelet denoising filter described in [3] is used since it has been reported as an effective method. For the basic algorithm (Lukas, Fridrich, and Goljan 2006, 205-214), the reference PRNU for each camera is estimated from 50 images with equation (2). For our method, Li's method (Li 2010, 280-287) and the MLE method (Chen et al. 2008, 74-90), the reference PRNU is estimated with equation (3) and the correlation for a particular image will be calculated by equation (5). For the phase pattern method (Kang et al. 2012, 393-402), before estimating the reference PRNU, the images are whitened in the frequency domain and transformed back to the spatial domain. To remove the periodical patterns, zero-mean operation and Wiener filter in frequency domain are used to preprocess the image as described in (Chen et al. 2008, 74-90).

For the proposed method, the detection statistics is calculated by the proposed detector as equation (19). For other methods, both the cross correlation and the peak to correlation energy (PCE) (Goljan, Fridrich, and Filler 2009, 72540I-72540I-12) (Goljan 2009, 454-468) are used as detection statistics as equation (17). The PCE has been reported more suitable for camera fingerprint detection because the presence of hidden periodic signal will lower PCE and reduce the possibility of false alarm (Goljan 2009, 454-468).

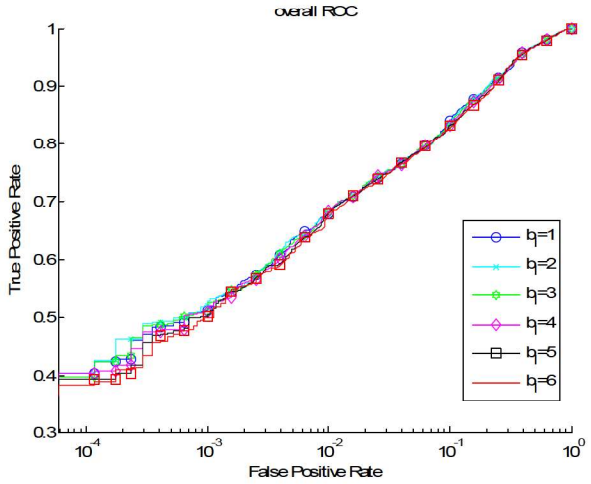

Fig. 7. ROC curves using different b_l

The choice of the lower bound b_l is important in that it decides the largest possible weight for the image pixels. The weight of a pixel cannot be infinitely large since at least there will be some unpredictable random noise for each pixel. To study the influence of b_l, we have tested the proposed algorithm with different settings of b_l i.e. b_l = 1, 2, 3, 4, 5 and 6 for 128 ×128 images. The Receiver Operating Curve (ROC) and the detection rate for given false acceptance rates are shown in Figure 7 and Table 3 respectively. We can observe that the proposed algorithm is not sensitive to the choice of b_l. The algorithm performs slightly better when b_l=2, 3 or 4. The value of b_l is then set to 4 in the following experiment.

The guided bilateral filter is used to estimate the local variance of the noise image. The spatial parameter and the range parameter determine the characteristics of the bilateral filter. Since the parameter of the bilateral filter will only influence the estimation of the local variance, it does not directly affect the performance of the algorithm. The performance of the algorithm is not sensitive to the parameter of bilateral filter. Similar experiments are carried with a set of different parameter settings and the spatial and range parameter

are determined as $\sigma_s^2 = 5$ and $\sigma_r^2 = 3$ which achieve good detection accuracy. The size of the window for estimating the local variance is set to 21. Since the pixels far from the center of the window contribute little to the estimation of the local variance due to the property of the bilateral filter, it is not necessary to make the window too large.

TABLE 3 THE TPR FOR GIVEN FPR USING DIFFERENT b_l

b_l	FPR=10^{-3}	FPR=10^{-2}
1	0.51579	0.68000
2	0.52842	0.68105
3	0.52526	0.68211
4	0.51684	0.68421
5	0.50526	0.68000
6	0.49895	0.68211

C. *Experimental result and analysis*

The ROC curve illustrates the performance of the source camera identification system by plotting the True Positive Rate (TPR) versus the False Positive Rate (FPR). Figure 8 and Figure 9 show the overall ROC performance over the 19 cameras for different methods at different settings. In practice, a low False Positive Rate is desired to ensure a low probability of False Alarm. Hence we set the thresholds of detection such that the False Positive Rate equals to 10^{-3} and 10^{-2}. The True Positive Rates at the given False Positive Rate are shown in Table 4 and Table 5 for 128 ×128 and 256 × 256 respectively.

It can be seen that for all the methods tested, the accuracy with the PCE as the test statistics is higher than that using the cross correlation detector. For the same method, using the PCE as the detector can always improve the detection accuracy. Another observation is that, for the phase method, the accuracy of detection using the cross correlation and PCE is very close. Using the correlation as the detector, the detection accuracy is better than other methods except the proposed method. Possible explanation is that the phase method can remove the common patterns caused by image post processing operations. Since the PCE can also reduce the effect of common patterns in the sensor noise, the two detectors will make little difference for the phase method.

From the experimental results we can see that the proposed method outperforms other state-of-the-art methods tested. The ROC curve obtained with the proposed methods is above those produced by other methods for both image size tested. The proposed method also gives the highest TPR at different levels of FPR among all the methods tested for both image sizes. For 128 × 128 image size, the TPR improvement is 3.58% - 9.96% at FPR = 10^{-3} and 3.58% - 9.16% at FPR = 10^{-2} as compared with other methods. The second best method is the phase method which performs well at FPR = 10^{-3} but it does not perform so well at the FPR = 10^{-2}. On the contrary, the proposed method performs constantly well for both FPR levels. For 256 × 256 images, the TPR improvement is 2.42% - 7.15% at FPR = 10^{-3} and 1.26% - 4.10% at FPR = 10^{-2}. It can be seen that the improvement is larger for smaller images. The reason for which the proposed method outperforms other methods is that the proposed method can better handle the scene content artifacts problems. The local variance estimated from the guided bilateral filter provides an accurate measure of the reliability of each pixel. The scene content artifacts make the detection statistics unreliable while the general matched filter can allocate optimal weightings to all the pixels such that the resultant detection statistics becomes more reliable. The combination of general matched filter and PCE detector can also resolve the problem of correlation among different sensors.

For the source camera classification problem, all the images will be classified according to their source camera. To further evaluate the performance of the proposed method, the experiment for source camera classification is also conducted over the same set of data. In this experiment, each image will be attributed to the camera whose fingerprint produces the largest test statistics. The classification accuracy is tabulated in Table 6 which shows that the proposed method performs the best among all the methods for the source camera classification problem. For 128 × 128 image, the improvement in accuracy is 3.69% - 4.43% and for 256 × 256 images, the improvement of

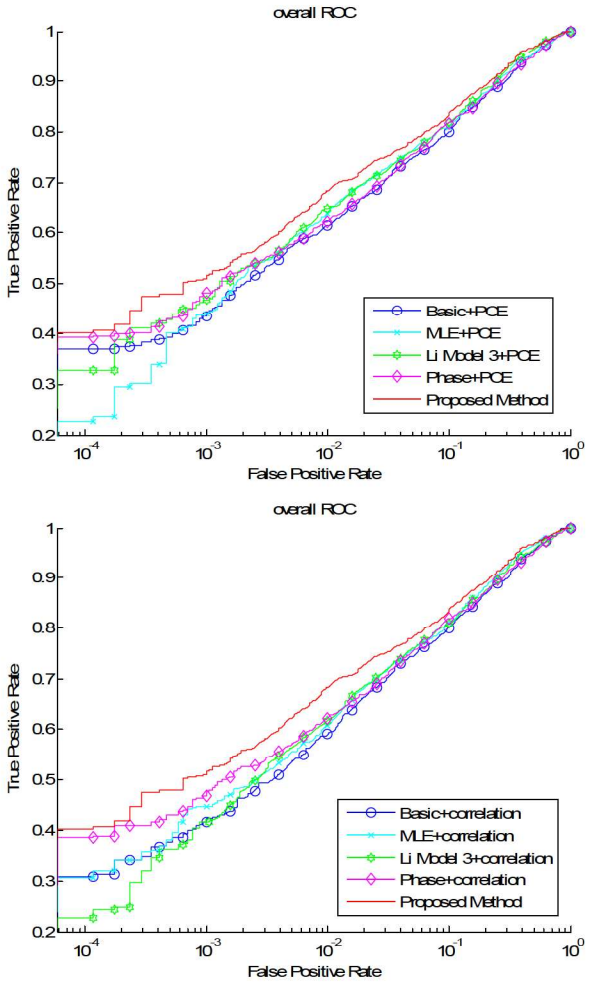

Fig. 8. The overall ROC curve for different methods for image size of 128 × 128. Top figure shows the results comparison with other methods using PCE as detection metric and bottom figure shows the comparison with other methods using the Normalized Cross Correlation (NCC) as detection metric.

accuracy is 1.47% - 2.63%. The experimental results also indicate that the proposed method improves the accuracy more

significantly when the image size become smaller. Table 7 and 8 show the classification result for each camera where the row number is the device ID of real source camera and the column number is the device ID that the image is classified to. It is noticeable that the classification accuracy varies from camera to camera. The classification accuracy is low for devices 5 and 11. It can also be observed that the images are misclassified to different cameras and the misclassification within the same model is low. It indicates that the correlation within the same model has limited influence to the classification accuracy. Table 9 and 10 shows the classification result of each camera

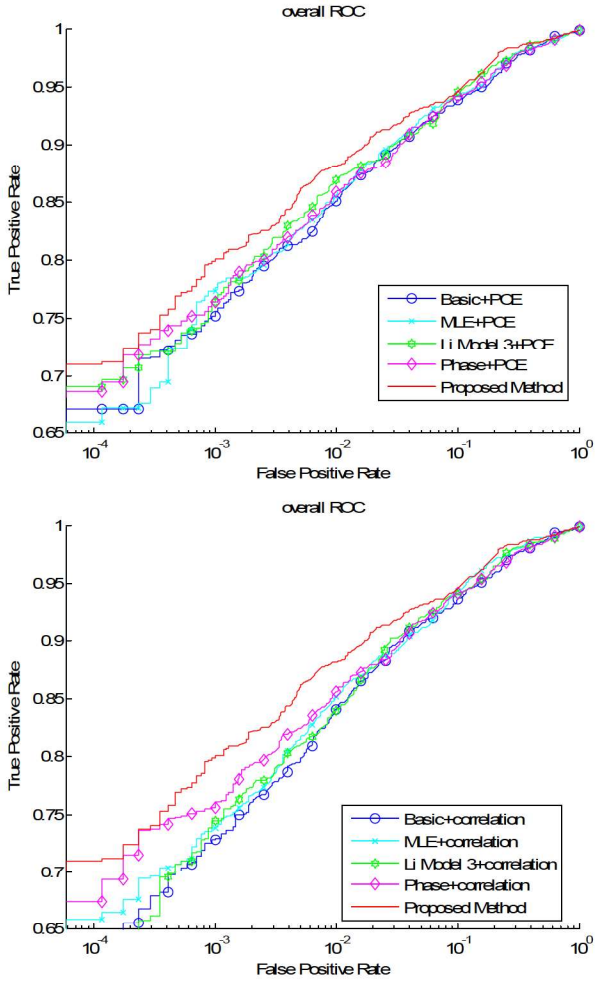

Fig. 9. The overall ROC curve for different methods for image size of 256 × 256

for all the methods tested from which we can see that the improvement of overall accuracy is due to correct detection from different cameras. The accuracy of the proposed method is the highest for 12 cameras and 13 for 128 × 128 images and 256×256 images respectively and for other cameras the proposed method also has accuracy close to that of the highest method. In summary, the proposed method has the best performance for both the source camera identification and source camera classification problem among all the methods tested.

TABLE 4 THE TRUE POSITIVE RATE FOR GIVEN FALSE POSITIVE RATE FOR IMAGE SIZE 128 × 128

	FPR=10^{-3}		FPR=10^{-2}	
	TPR	Improvement of proposed method	TPR	Improvement of proposed method
Basic + correlation	0.42421	+9.26316%	0.59263	+9.15789%
Basic + PCE	0.44316	+7.36842%	0.61474	+6.94737%
MLE + correlation	0.44784	+6.90023%	0.60906	+7.51484%
MLE + PCE	0.44000	+7.68421%	0.63895	+4.52632%
Model3 + correlation	0.41728	+9.95608%	0.61855	+6.56647%
Model3 + PCE	0.46842	+4.84211%	0.64842	+3.57895%
Phase + correlation	0.47684	+4.00000%	0.62316	+6.10526%
Phase + PCE	0.48105	+3.57895%	0.62211	+6.21053%
Proposed Method	**0.51684**	-	**0.68421**	-

TABLE 5 THE TRUE POSITIVE RATE FOR GIVEN FALSE POSITIVE RATE FOR IMAGE SIZE 256 × 256

	FPR=10^{-3}		FPR=10^{-2}	
	TPR	Improvement of proposed method	TPR	Improvement of proposed method
Basic + correlation	0.72947	+7.15789%	0.84105	+4.10526%
Basic + PCE	0.75579	+4.52632%	0.85158	+3.05263%
MLE + correlation	0.73973	+6.13266%	0.85248	+2.96290%
MLE + PCE	0.77684	+2.42105%	0.85684	+2.52632%
Model3 + correlation	0.74499	+5.60579%	0.83983	+4.22739%
Model3 + PCE	0.76737	+3.36842%	0.86947	+1.26316%
Phase + correlation	0.76105	+4.00000%	0.85579	+2.63158%
Phase + PCE	0.76421	+3.68421%	0.85895	+2.31579%
Proposed Method	**0.80105**	-	**0.88211**	-

TABLE 6 THE SOURCE CAMERA CLASSIFICATION ACCURACY

	128 × 128		256 × 256	
	Accuracy	Improvement of proposed method	Accuracy	Improvement of proposed method
Basic + correlation	0.7189	+4.43%	0.9042	+2.21%
Basic + PCE	0.7232	+4.00%	0.9011	+2.52%
MLE + correlation	0.7263	+3.69%	0.9116	+1.47%
MLE + PCE	0.7284	+3.48%	0.9032	+2.31%
Model3 + correlation	0.7263	+3.69%	0.9063	+2.00%
Model3 + PCE	0.7326	+3.06%	0.9116	+1.47%
Phase + correlation	0.7263	+3.69%	0.9000	+2.63%
Phase + PCE	0.7263	+3.69%	0.9000	+2.63%
Proposed Method	**0.7632**	-	0.9263	-

V. CONCLUSION

Though PRNU has been proved as an effective means for source camera identification, the scene content artifact can severely deteriorate the performance of PRNU-based camera identification. The detection accuracy will be low if the image contains dark areas and complicated texture, the image size is small and the quality of image is bad. In this paper, we have studied the relation between the reliability of detection and the local variance. We use the local variance of each pixel to estimate the distribution of the undesired noise signal. With the estimated distribution, the general matched filter which is an optimal detector theoretically is used. The proposed method is compared with several state-of-art methods. The experiments show that the proposed method outperformed other state-of-the-art methods in terms of the ROC curve and detection accuracy for both source camera identification and classification problems.

ACKNOWLEDGMENT

The work was supported by the GRF Grant PolyU 152080/14E (project code: BQ44N) of the Hong Kong SAR Government and the Center for Signal Processing, the Hong Kong Polytechnic University (Project GYN20). Shi Chao acknowledges the research studentships provided by the University.

TABLE 7 SOURCE CLASSIFICATION MATRIX FOR IMAGE SIZE 128 × 128

ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	38	2	0	1	0	0	1	0	2	1	0	1	1	0	2	0	1	0	0
2	0	40	2	1	1	1	2	1	0	1	1	0	0	0	0	0	0	0	0
3	0	1	37	0	1	1	0	3	2	1	0	1	0	1	1	0	0	0	1
4	0	3	3	33	0	1	0	0	0	2	1	1	0	0	2	1	1	2	0
5	3	0	1	2	28	0	1	1	1	1	4	1	1	1	1	2	1	1	0
6	4	1	2	1	4	31	3	2	2	3	1	3	1	1	3	2	2	1	1
7	0	2	2	1	1	2	36	0	1	1	1	0	1	1	0	1	0	0	0
8	0	0	0	0	0	0	0	50	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	50	0	0	0	0	0	0	0	0	0	0
10	1	0	0	0	2	4	1	1	2	28	1	2	1	1	4	1	0	0	1
11	0	1	3	1	1	2	1	2	3	2	26	0	0	1	4	2	1	0	0
12	0	0	0	1	0	2	0	0	1	0	0	34	2	1	0	3	0	6	0
13	0	1	2	0	0	0	0	1	0	0	0	2	38	0	0	3	2	1	0
14	2	0	1	0	0	0	0	1	0	2	0	0	0	42	1	0	0	0	1
15	0	2	0	0	0	0	0	0	0	1	1	0	0	1	44	0	0	1	0
16	1	1	0	0	1	0	0	0	0	0	0	1	0	0	0	42	1	2	1
17	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	1	43	1	2
18	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	2	2	42	2
19	0	0	0	1	0	0	0	0	0	0	0	2	1	0	0	1	0	1	44

TABLE 8 SOURCE CLASSIFICATION MATRIX FOR IMAGE SIZE 128 × 128

ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	46	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
2	1	47	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0
3	0	0	48	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
4	1	2	0	43	0	1	0	0	1	1	0	0	0	0	0	0	0	0	1
5	0	0	1	1	41	0	1	0	0	0	1	1	1	0	1	0	1	1	0

6	1	2	2	1	1	46	1	1	1	1	1	1	1	1	3	1	1	1	1
7	0	0	0	0	0	0	48	0	0	0	0	0	0	0	1	0	0	0	1
8	0	0	0	0	0	0	0	50	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	50	0	0	0	0	0	0	0	0	0	0
10	3	0	1	0	1	1	0	2	1	34	0	2	0	1	2	1	1	0	0
11	0	1	0	0	1	1	0	3	2	1	39	0	0	1	0	0	1	0	0
12	0	0	0	0	0	0	0	0	0	1	0	47	1	0	0	0	0	1	0
13	0	0	0	0	0	0	0	0	0	0	0	0	47	0	0	1	0	1	1
14	0	0	0	0	0	0	0	0	0	0	0	0	0	50	0	0	0	0	0
15	0	0	0	1	0	0	0	0	0	0	0	0	0	0	48	1	0	0	0
16	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	49	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	50	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	49	1
19	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	48

TABLE 9 NUMBER OF CORRECT CLASSIFICATION FOR EACH CAMERA FOR IMAGE SIZE 128 × 128

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Basic + correlation	36	41	34	28	23	27	35	**50**	**50**	17	24	33	35	38	**44**	44	**45**	40	39
Basic + PCE	37	41	34	29	23	28	33	**50**	**50**	17	25	**35**	37	38	43	44	44	40	39
MLE + correlation	34	36	32	31	25	27	**36**	**50**	49	25	23	33	34	**42**	**44**	42	42	40	**45**
MLE + PCE	35	41	35	30	23	28	34	**50**	49	22	23	32	36	40	**44**	**43**	44	41	43
Model3 + correlation	35	41	34	30	22	28	34	**50**	49	21	24	32	33	41	**44**	**43**	44	41	44
Model3 + PCE	34	38	35	31	25	28	**36**	**50**	**50**	26	23	34	36	41	**44**	42	41	40	43
Phase + correlation	35	42	32	29	26	**33**	**36**	**50**	**50**	19	**28**	29	31	36	43	41	**45**	**42**	43
Phase + PCE	35	**43**	32	29	26	**33**	**36**	**50**	**50**	19	27	29	31	36	43	41	45	**42**	43
Proposed Method	**38**	40	**36**	**33**	**28**	31	**36**	**50**	**50**	**28**	26	34	**38**	**42**	**44**	42	43	**42**	44

TABLE 10 NUMBER OF CORRECT CLASSIFICATION FOR EACH CAMERA FOR IMAGE SIZE 256 × 256

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Basic + correlation	**48**	47	47	40	38	43	47	**50**	**50**	30	30	**47**	47	48	47	49	**50**	49	**49**
Basic + PCE	46	**48**	46	**43**	40	42	48	**50**	**50**	31	33	45	47	48	**48**	49	**50**	49	46
MLE + correlation	47	47	46	**43**	40	40	46	**50**	**50**	35	38	45	44	49	49	**50**	**50**	**50**	48
MLE + PCE	**48**	47	47	40	39	41	46	**50**	**50**	**36**	33	**47**	45	47	46	49	**50**	49	45
Model3 + correlation	**48**	47	47	40	38	43	47	**50**	**50**	32	32	46	48	47	47	49	**50**	49	**49**
Model3 + PCE	47	47	46	42	**41**	39	46	**50**	**50**	35	38	45	44	49	49	**50**	**50**	**50**	48
Phase + correlation	46	**48**	46	43	**41**	41	**48**	**50**	**50**	33	35	45	45	48	**48**	49	**50**	49	46
Phase + PCE	**48**	47	47	40	39	41	46	**50**	**50**	36	33	**47**	45	47	46	49	**50**	49	45
Proposed Method	46	47	**48**	**43**	**41**	**46**	**48**	**50**	**50**	34	**39**	**47**	**47**	**50**	**48**	49	**50**	49	48

References

Alles, Erwin J., Zeno JMH Geradts, and Cor J. Veenman. 2008. Source camera identification for low resolution heavily compressed images. Paper presented at Computational Sciences and Its Applications, 2008. ICCSA'08. International Conference on, .

Blythe, P., and J. Fridrich. 2004. Secure digital camera. Paper presented at Digital Forensic Research Workshop, .

Caraffa, Laurent, Jean-Philippe Tarel, and Pierre Charbonnier. 2015. The guided bilateral filter: When the Joint/Cross bilateral filter becomes robust. *Image Processing, IEEE Transactions on* 24 (4): 1199-208.

Chan, Lit-HIung, Ngai-Fong Law, and Wan-Chi Siu. 2013. A confidence map and pixel-based weighted correlation for PRNU-based camera identification. *Digital Investigation* 10 (3): 215-25.

Chen, M., J. Fridrich, M. Goljan, and J. Lukás. 2008. Determining image origin and integrity using sensor noise. *Information Forensics and Security, IEEE Transactions on* 3 (1): 74-90.

Chen, Mo, Jessica Fridrich, and Miroslav Goljan. 2007. Digital imaging sensor identification (further study). Paper presented at Electronic Imaging 2007, .

Chen, Mo, Jessica Fridrich, Jan Lukáš, and Miroslav Goljan. 2007. Imaging sensor noise as digital x-ray for revealing forgeries. Paper presented at Information hiding, .

Fridrich, Jessica. 2009. Digital image forensics. *Signal Processing Magazine, IEEE* 26 (2): 26-37.

Gloe, Thomas, and Rainer Böhme. 2010. The dresden image database for benchmarking digital image forensics. *Journal of Digital Forensic Practice* 3 (2-4): 150-9.

Goljan, Miroslav. 2009. Digital camera identification from images–Estimating false acceptance probability. In *Digital watermarking.*, 454-468Springer.

Goljan, Miroslav, Jessica Fridrich, and Tomáš Filler. 2009. Large scale test of sensor fingerprint camera identification. Paper presented at IS&T/SPIE Electronic Imaging, .

Goljan, Miroslav, Jessica Fridrich, and Jan Lukáš. 2008. Camera identification from printed images. Paper presented at Electronic Imaging 2008, .

Kang, Xiangui, Yinxiang Li, Zhenhua Qu, and Jiwu Huang. 2012. Enhancing source camera identification performance with a camera reference phase sensor pattern noise. *Information Forensics and Security, IEEE Transactions on* 7 (2): 393-402.

Kay, Steven M. 1998. Fundamentals of statistical signal processing, vol. II: Detection theory. *Signal Processing.Upper Saddle River, NJ: Prentice Hall*.

Li, Chang-Tsun. 2010. Source camera identification using enhanced sensor pattern noise. *Information Forensics and Security, IEEE Transactions on* 5 (2): 280-7.

Li, Chang-Tsun, and Yue Li. 2010. Digital camera identification using colour-decoupled photo response non-uniformity noise pattern. Paper presented at Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on, .

Lukas, Jan, Jessica Fridrich, and Miroslav Goljan. 2006. Digital camera identification from sensor pattern noise. *Information Forensics and Security, IEEE Transactions on* 1 (2): 205-14.

McCloskey, Scott. 2008. Confidence weighting for sensor fingerprinting. Paper presented at Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on, .

Mehdi, KL, H. T. Sencar, and N. Memon. 2004. Blind source camera identification. Paper presented at Image Processing, 2004. ICIP'04. 2004 International Conference on, .

Rosenfeld, Kurt, and Husrev T. Sencar. 2009. A study of the robustness of prnu-based camera identification. Paper presented at IS&T/SPIE Electronic Imaging, .

Sencar, Husrev T., and Nasir Memon. 2008. Overview of state-of-the-art in digital image forensics. *Algorithms, Architectures and Information Systems Security* 3 : 325-48.

Shi, Chao, Ngai-Fong Law, Hung-Fat Leung, and Wan-Chi Siu. 2014. Weighting optimization with neural network for photo-response-non-uniformity-based source camera identification. Paper presented at Asia-Pacific Signal and Information Processing Association, 2014 Annual Summit and Conference (APSIPA), .

Tomasi, Carlo, and Roberto Manduchi. 1998. Bilateral filtering for gray and color images. Paper presented at Computer Vision, 1998. Sixth International Conference on, .

Turin, George L. 1976. An introduction to digital matched filters. *Proceedings of the IEEE* 64 (7): 1092-112.

Wong, Ping Wah. 1998. A watermark for image integrity and ownership verification. Paper presented at IS AND TS PICS CONFERENCE, .

Yang, Qingxiong, Kar-Han Tan, and Narendra Ahuja. 2009. Real-time O (1) bilateral filtering. Paper presented at Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, .