International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:10, No:12, 2016

Evaluation of Sensor Pattern Noise Estimators for Source Camera Identification

Benjamin Anderson-Sackaney, Amr Abdel-Dayem

***Abstract*—**This paper presents a comprehensive survey of recent source camera identification (SCI) systems. Then, the performance of various sensor pattern noise (SPN) estimators was experimentally assessed, under common photo response non-uniformity (PRNU) frameworks. The experiments used 1350 natural and 900 flat-field images, captured by 18 individual cameras. 12 different experiments, grouped into three sets, were conducted. The results were analyzed using the receiver operator characteristic (ROC) curves. The experimental results demonstrated that combining the basic SPN estimator with a wavelet-based filtering scheme provides promising results. However, the phase SPN estimator fits better with both patch-based (BM3D) and anisotropic diffusion (AD) filtering schemes.

Keywords—Sensor pattern noise, source camera identification, photo response non-uniformity, anisotropic diffusion, peak to correlation energy ratio.

I. Introduction

SCI has become an important topic for digital forensics due to the increase in use of digital media. Its importance is reflected within forensics on the need to confidently verify whether an image came from a particular camera or not. With the accessibility of digital image technology, the potential of forgery or falsely linking a picture to a camera are sever threats, which need to be reliably and consistently resolved. Examples of such cases include: Using images as evidence of a crime, child pornography, or movie piracy.

The use of PRNU for identifying images to cameras has become a rich field of research for establishing SCI. It makes use of sensor defects inherent to any device used for capturing digital images, which have been shown to be unique to any particular device. This technique is analogous to the use of human fingerprints for identification purposes. In practice, it showed reliable and consistent results for images across similar brands and models and varying conditions, like compression and size, showing its potential for study.

The process of SCI can be broken down into four basic stages: PRNU extraction using denoising filters, reference SPN estimation, enhancement of the PRNU and SPN signals using pre-processing operations, and finally a classification stage, Fig. 1. Estimation of the reference SPN (which is in essence the camera's fingerprint) is accomplished by using the PRNU residuals that are extracted from a set of images from the camera. In practice, they are normally taken as flat-field or blue sky images to obtain the strongest possible PRNU signals, as suggested by Fridrich [1]. It is then compared to the SPN extracted from the image in question to determine its correlation with the given camera.

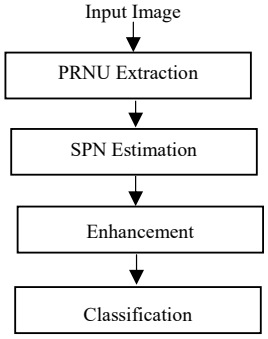

Fig. 1 The block diagram of the SCI system

PRNU extraction is generally performed by using a denoising filtering scheme to smooth away the noise from an image. Afterwards, the denoised image is subtracted from the original, thus obtaining the noise residual. The goal is to ensure that the noise residual contains as much of the PRNU noise, and at the same time, as little other signals from other sources as possible. As studied by Fridrich [1], scene details, such as textures and dark regions, corrupt the PRNU signal since it is modelled as a multiplicative factor with light intensity. The majority of work, found in literature, focuses on creating better filters from which the noise can be extracted [1].

The PRNU extraction stage is followed in one of two ways; either the residual is enhanced and compared to the reference SPN or they are used to construct the reference SPN. Reference SPN estimation is accomplished by performing an estimation procedure on a set of noise residuals gathered from the camera. Using a set of images maximizes the likelihood that the true fingerprint is “averaged” out of the residuals, thus providing a close estimation of the camera's fingerprint noise [1].

The accuracy of extracting the PRNU signals and estimating the SPN is dependent upon the capability of the used technique to filter out the non-unique patterns imparted onto the images by their respective cameras. A specific method can obtain good results under two conditions. First, it should be capable of separating the unique SPN from the image without imparting artifacts of its own onto the image. Second, it should shape the SPN to be as Gaussian as possible.

Benjamin Anderson-Sackaney is with the Department of Mathematics and Computer Science, Laurentian University, Sudbury, Ontario, Canada (e-mail: bandersonsackaney@laurentian.ca).

Amr Abdel-Dayem is with the Department of Mathematics and Computer Science, Laurentian University, Sudbury, Ontario, Canada (phone: +1 705-675-1151 extension (2396); fax: +1–705–673-6591; e-mail: aabdeldayem@lcs.laurentian.ca).

Due to the underlying assumption that the PRNU follows a Gaussian distribution, the closer the noise approaches this distribution, the more likely it contains only the PRNU signal [1].

Across all stages of the SCI framework described, artifacts imprinted on the noise residuals and on the reference SPN will hinder detection once they reach the final stage. This fact justifies the enhancement stage just before detection. In most image capture devices, there are a host of known artifacts which leave distinct imprints on any given image. These causes can be resulted from the colour filter array (CFA) interpolation [2], [3], JPEG compression [4], [5], or of unknown origin [6]. Since these artifacts, along with the those introduced by filtering, leave recognizable traces, methods can be implemented to specifically target and remove these traces, and thus enhancing the reference SPN or the image PRNU before detection.

Once the reference SPN and the residual of the image in question have been enhanced, the next and final step of the SCI process is the classification stage. This is accomplished by employing a classification scheme that compares the similarity of the pixels within the residual of the image in question to the pixels of the reference SPN. Statistical classifiers are commonly used in SCI systems. A statistical value is obtained, and if this value exceeds a user defined threshold, the image in question is deemed to be captured by the reference camera.

Standing as among the best denoising filters for SCI, [7] developed a wavelet-based scheme, which had, for the first time, become practical for real world applications. It is composed a wavelet transformation combined with a minimum square error (MMSE) and a maximum likelihood estimation procedure to suppress periodic noise in the frequency domain. Under the assumption that the image is corrupted by white Gaussian noise, the goal is to find an estimate of the uncorrupted variance field of the wavelet coefficients. This goal is achieved by employing a maximum likelihood estimator, together with a Bootstrap method to approximate the estimated distribution. Finally, the uncorrupted image is constructed using its variance field and the MMSE procedure.

Lukas et al. [8] proposed a basic SPN estimation procedure by simply taking the average of the noise residuals. Their work showed proof of concept, providing a framework that showed a successful SCI system. Along with the basic SPN estimator, the correlation coefficient between the reference SPN and the noise residual was employed during the classification stage. This provides the ability to measure the similarity between the images, under investigations, and a source camera using its SPN. Experimental results show that this framework is robust across different JPEG compression levels and gamma corrections. Moreover, it can distinguish between cameras of the same model. However, its performance deteriorates on images under geometric transformations. This might be a resulted from the desynchronization of the PRNU signal with respect to the camera.

Taking into consideration the possibility of the presence of a geometric transformation of the noise residual with respect to the reference SPN, the normalized cross correlation (NCC) was explored by [9]. It can be considered as a simplification to the energy model derived by [10]. While the computational complexity of the energy model [10] is too high to be used in practical applications, the NCC [9] provides an acceptable approximation with low computational complexity. The idea is to maximize the NCC over all possible spatial shifts between the residual and the SPN. Then, the peak to correlation energy (PCE) ratio is computed, by taking the ratio of the square of the NCC at peak spatial shifts to the average of the square of the NCC values within some neighbourhood around these peak spatial shifts. It acted as a normalization of the NCC statistic that reduces the probability of false alarms during the classification stage. In their experiments, the PCE was shown to be superior to correlation and had become the norm for executing the classification stage.

Chen et al. [11] developed the basic method [8] further by applying the maximum likelihood approach to estimate the noise residuals. The log likelihood estimate was found and differentiated, thus providing an estimate of the reference SPN in the form of a weighted average: the maximum likelihood estimate (MLE). The MLE method was shown to provide better use of data for forgery detection and allowed for better error estimates. However, the experiments conducted by [11], did not contain explicit comparisons to the basic SPN estimation method [8]. Along with the MLE SPN estimator, [11] proposed the enhancement of the residual and the SPN using both the Wiener and the Zero Meaning filtering schemes. They were used to suppress both periodic artifacts and linear patterns from the reference SPN and the image PRNU. While the Wiener filter suppresses the low frequency components, the Zero meaning filter subtracts the means of the rows and columns from each pixel of the noise. This approach improves the source camera detection, particularly when dealing with similar cameras (the same model or manufacturer). It has been shown that combining the Zero Meaning and Wiener Filtering with the MLE provides a reliable forgery detection scheme.

To improve upon the wavelet-based scheme [7], an adaptive spatial (AS) filtering approach was introduced by [12]. It is a spatial filtering scheme that aims at isolating regions with stronger PRNU signals. The algorithm creates an index map by extracting regions, which are deemed to be the least dark and textured. Then, the selected regions are denoised using a median filter followed by an adaptive Wiener filter. Only the selected regions from both the reference SPN and noise residual of the image under investigation are compared. As a result, the proposed approach avoids the use of less than ideal regions in the identification process. Cooper et al. [12] experimentally showed that their proposed approach is robust with respect to variable JPEG compression ratios.

Taking into consideration both the frequency and the spatial domains, [13] used the sparse 3-D transform-domain collaborative block-matching (BM3D) filtering scheme, as introduced by [14], to denoise images for forgery detection.

The idea is to search for similar blocks within the image, based on a given similarity measure. When two blocks are similar, their center pixels are said to be matched. Then, the matching pixels are linearly transformed into a higher dimension in the spatial domain, where shrinkage is applied to these pixels, using hard thresholding. The output from this step is followed by a weighted average of overlapping pixels in the different blocks, producing an estimate of the original image. This process is repeated between the estimated image and the noisy one, where similar blocks are grouped together from both images. It is worth mentioning that during the second iteration, Wiener filter is used instead of hard thresholding to shrink the linearly transformed coefficients. This process is repeated, based on a pre-set convergence criteria. Then, the final estimated image is considered the final filtered output image. As a BM3D filtering approach, the BM3D filter shows superior performance compared to pixel-based filters, particularly when dealing with fine details. Dabov et al. [14] demonstrated the BM3D filter produces higher peak signal-to-noise ratios, when compared to the spatially adaptive Wavelet filter [7]. Experimental studies, conducted by [13], showed that employing the BM3D filter into SCI systems is a promising direction. However, more experiments are still needed to verify its effectiveness in SCI systems under challenging conditions (e.g. variable image sizes, JPEG compression, and various SPN estimators).

Houten et al. [15] used the AD filter for image denoising [16]. Their experimental work shows that the AD filtering scheme provides more accurate results when employed in an SCI system, compared to the wavelet filtering scheme proposed by [7].

Similar to the AD filter, [17] introduced a simplified total variation model (TVM), known as the first step total variation (FSTV). The filtering process is modeled as an energy minimization problem, which is iteratively solved. Similar to the AD filter [15], the FSTV emerged as a simple and computationally efficient approach to removing noise from an image. While it provides minimal denoising, FSTV extracts enough information to successfully complete the SCI process. It was experimentally compared to the AS [12], the AD [15], and the wavelet [7] filtering approaches over variable SPN estimators and JPEG compression ratios. The experiments reported the FSTV filter as the winner in terms of both accuracy and computational time.

In our project, we aim at experimentally checking the performance of the SCI system under varying choices for stages 1, 2 and 3 of Fig. 1. Our goal is to find the best combination among the common existing schemes found in literature.

The rest of this paper is organized as follows. Sections II and III present the experimental setup and the obtained results, respectively. Then, Section IV offers the conclusions of this paper. Finally, Section V highlights the major directions to extend this research in future.

II. Experimental Setup

The images used in these experiments were downloaded from the Dresden Image Database [18]. A group of 75 natural images and 50 flat-field color images were collected from 18 different devices. This resulted in 1350 natural and 900 flat-field images in total. The selected 18 individual cameras spanned four manufacturers and six models as follows:

- Canon Ixus 70: three cameras
- Ixus 55: One camera
- Olympus MJU: five cameras
- Praktica DCZ 5.9: five cameras
- Samsung NV15: three camers
- Samsung SGH-S730M: one camera

MATLAB was used to implement a prototype of the SCI system, described in Fig. 1. In order to check the interaction between the PRNU extraction and the SPN estimation stages alone, we set the third stage of Fig. 1 to null (i.e. there is no enhancement step). The classification stage employs the PCE ratio, as described in [9]. For the PRNU extraction stage, we experimented the wavelet [7], the BM3D [14], and the AD [15] approaches. For the SPN estimation stage, we experimented the basic [8], the MLE [11], the phase [19], and the phase MLE methods. This resulted in twelve different experiments, grouped into three different sets.

For each camera, its 50 flat-field images were used to construct the reference SPN, using the PRNU extraction and SPN estimation procedures (stages 1 and 2 of Fig. 1). Then, for any given camera, the noise residual of each of the 2000 natural images were extracted and compared to the reference SPN using the PCE ratio, generating a “PCE value” for each image. The PCE values were divided into two separate categories; if the natural image originated from the same source as the reference SPN, it was regarded as a true PCE value, and if the natural image originated from a different source from the reference SPN, it was regarded as a false PCE value. This approach sets up a binary hypothesis test, where a threshold was set and compared to the PCE values. If a true PCE value is above the threshold, it is a true detection otherwise it is a false rejection. Similarly, If a false PCE value is above the threshold, it is a false detection otherwise it is a true rejection. The number of true and false positives, as well as the number of true and false negatives, was counted. Then, the true positive (TPR) and the false positive (FPR) rates were calculated as:

$$TPR = \frac{NT}{LT}, \text{ and } FPR = \frac{NF}{LF}$$

where, LT and LF are the total number of true and false images respectively, and NT and NF are the total number of true positives and false positives respectively. These values were used to build ROC curves, where the TPR was plotted against the FPR for each individual case.

III. Results/Discussion

Fig. 2 shows the ROC curves for the first set of

experiments, where the wavelet approach [7] was employed in the first stage of Fig. 1 together with the various SPN estimators described in Section II.

The analysis of Fig. 2 reveals that the basic SPN estimator [8] achieves the highest performance for false positive rates less than 0.001. It shows an increase in the area under its ROC curve over the curves of the MLE [11], the phase [19], and the phase MLE SPN estimators by 14%, 16%, and 34% respectively. However, for false positive rates higher than 0.001, it has the same performance as the MLE SPN estimator. Similarly, the phase SPN estimator yields higher performance compared to the phase MLE SPN estimator for false positive rates lower than 0.002. Finally, this set of ROC curves suggests that for false positive rates higher than 0.002, both the basic and the MLE estimators outperform both the phase and the phase MLE estimators by approximately 28% increase under their corresponding ROC curves.

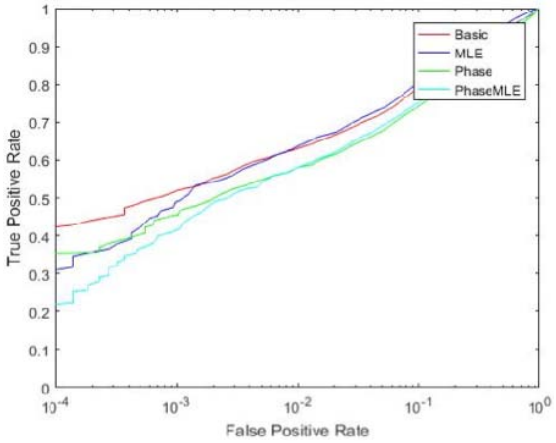

Fig. 2 ROC curves for the first set of experiments: Wavelet approach [7] with various SPN estimators

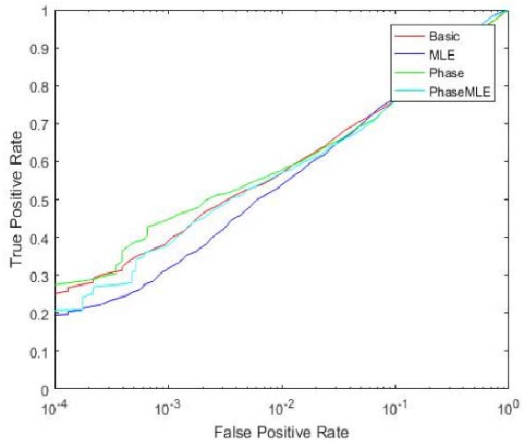

Fig. 3 ROC curves for the second set of experiments: BM3D approach [14] with various SPN estimators.

Fig. 3 shows the ROC curves for the second set of experiments, where the BM3D approach [14] was employed in the first stage of Fig. 1 together with the various SPN estimators described in Section II.

The analysis of Fig. 3 reveals that the there is no significant difference between the curves for false positive rates higher than 0.01. However, the phase SPN estimator outperforms the other three estimators for lower false positive rates. The area under the ROC curve for the phase SPN estimator shows an increase of 34%, 11%, and 9.5% compared to the MLE, the phase MLE, and the basic SPN estimators, respectively. The analysis also revealed that the phase MLE SPN estimator yields the worst performance, compared to the other estimators, when used with the BM3D approach.

Fig. 4 shows the ROC curves for the third set of experiments, where the AD approach [14] was employed in the first stage of Fig. 1 together with the various SPN estimators described in Section II.

The analysis of Fig. 4 shows very little discrimination between the ROC curves for false positive rates higher than 0.002. However, for false positive rates lower than 0.002, the phase SPN estimator outperforms all other estimators. Whereas, the MLE SPN estimator yields the worst performance, compared to the other estimators, when used with the AD approach (the area under its ROC curve is 13% less than the area under the ROC of the phase SPN estimator).

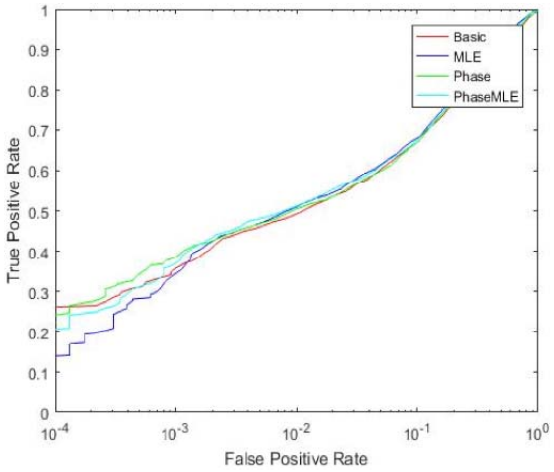

Fig. 4 ROC curves for the third set of experiments: AD approach [15] with various SPN estimators.

IV. Conclusion

In this paper, we experimentally checked the performance of various common SPN estimators under different PRNU extraction approaches. We run three different sets of experiments using a database of 1350 natural and 900 flat-field images, captured by 18 individual cameras. The ROC curves were employed to assess the performance of each SPN estimator. The statistical analysis revealed that the basic SPN estimator outperforms the other candidates when the wavelet approach is used for PRNU extraction. However, the phase SPN estimator outperforms all other estimators when used with either the BM3D or the AD extraction stages.

V. Future Work

In future, we plan to conduct more experiments to check the impact of various enhancement approaches (stage 3 in Fig. 1) on the overall performance of the identification system. Upon completing all experiments, a comprehensive statistical analysis will be employed to recommend the optimum configuration of a SCI system.

References

[1] J. Fridrich "Sensor Defects In Digital Image Forensic", *Digital Image Forensics*, Vol. 2, No. 2, pp. 179-218, July 2012.
[2] S. Bayram, H. Sencar, N. Memon, and I. Avcibas, "Source camera identification based on CFA interpolation," *IEEE International Conference on Image Processing.*, vol. 3, pp. III-69–III-72, Sep. 2005.
[3] M. Swaminathan, Wu, and K. J. R. Liu, "Nonintrusive component forensics of visual sensors using output images," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 1, pp. 91–106, Mar. 2007.
[4] M. J. Sorrell, "Digital camera source identification through JPEG quantisation," *Multimedia Forensics and Security,* pp. 291–313, 2008.
[5] E. J. Alles, Z. J. M. H. Geradts, and C. J. Veenman, "Source camera identification for heavily JPEG compressed low resolution still images,"*Journal of Forensic Science*, vol. 54, no. 3, pp. 628–638, May 2009.
[6] T. Gloe, S. Pfennig, and M. Kirchner, "Unexpected Artefacts in PRNU Based Camera Identification: A 'Dresden Image Database' Case-Study," in *Proc. ACM Workshop Multimedia Security*, pp. 109–114, Sep. 2012.
[7] MK Mihcak, I Kozintsev, K Ramchandran, "Spatially adaptive statistical modeling of wavelet image coefficients and its application to denoising." *IEEE international conference on acoustics, speech, and signal processing*, vol. 6, pp. 3253–3256, May 1999
[8] J. Lukas, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, pp. 205–214, Jun. 2006.
[9] M. Goljan, J. Fridrich "Camera identification from scaled and cropped images," *Proceedings of SPIE electronic imaging, security, forensics, steganography, and watermarking of multimedia contents* X, vol 6819, pp OE1-OE13, 2008
[10] CR. Holt "Two-channel detectors for arbitrary linear channel distortion," *IEEE Trans Acoust Speech Signal Process* ASSP-35(3):267–273, 1987
[11] M. Chen, J. Fridrich, M. Goljan, and J. Lukás, "Determining image origin and integrity using sensor noise," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 1, pp. 74–90, Mar. 2008.
[12] A. Cooper, "Improved photo response non-uniformity (prnu) based source camera identification," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 2, pp. 260-271, Feb. 2012
[13] G. Chierchia, S. Parrilli, G. Poggi, C. Sansone, and L. Verdoliva, "On the influence of denoising in PRNU based forgery detection," *Image Processing and Computer Vision-General*, pp. 117–122, Oct. 2010
[14] K. Dabov, A. Foi, V. Katkovnik, and 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.
[15] W.V. Houten, Z. Geradts, "Using Anisotropic Diffusion for Efficient Extraction of Sensor Noise in Camera Identification", *Journal of Forensic Sciences.*, vol. 57, no. 2, pp. 521-527, March 2012.
[16] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no. 7, pp. 629–639, July 1990.
[17] F. Gisolf, A. Malgoezar, T. Baar, Z. Geradts: 'Improving source camera identification using a simplified total variation based noise removal algorithm', *Digital Investigation*, vol. 10, no. 3, pp. 207-214, Oct. 2013.
[18] T. Gloe and R. Böhme, "The dresden image database for benchmarking digital image forensics," Journal of Digital Forensic Practice, vol. 3, no. 2–4, pp. 150–159, 2010.
[19] X. Kang, Y. Li, Z. Qu, and 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, Apr. 2012.