Source Camera Identification Using Enhanced Sensor Pattern Noise

Chang-Tsun Li

Department of Computer Science, University of Warwick, Coventry CV4 7AL, UK

c-t.li@warwick.ac.uk

ABSTRACT

Sensor pattern noises (SPNs), extracted from digital images to serve as the *fingerprints* of imaging devices, have been proved as an effective way for digital device identification. However, as we demonstrate in this work, the limitation of the current method of extracting the sensor pattern noise is that the SPNs extracted from images can be severely contaminated by the details from scenes, and as a result, the identification rate is unsatisfactory unless images of a large size are used. In this work, we propose a novel approach for attenuating the influence of the details from scenes on sensor pattern noises so as to improve the device identification rate of the identifier. The hypothesis underlying our SPN enhancement method is that the stronger a signal component in a SPN is, the less trustworthy the component should be, and thus should be attenuated. This hypothesis suggests that an enhanced SPN can be obtained by assigning weighting factors inversely proportional to the magnitude of the SPN components.

Index Terms—*Source device identification, multimedia forensics, digital forensics, digital investigation, sensor pattern noise*

I. Introduction

As the cost of digital imaging devices, such as camcorders, digital cameras, scanners and cameras embedded in mobile phones, falls and the functionalities of these devices increase, digital imaging become increasingly cheaper in our every-day life. While digital imaging devices bring ever-increasing convenience of image acquisition, powerful, yet easy-to-use digital image processing tools also provide effective means for manipulating images that can serve good and malicious purposes. As a result, the use of digital images in forensic investigations becomes more frequent and important. Typical image forensics includes source device identification, source device linking, classification of images taken by unknown cameras, integrity verification, authentication, etc.

Usually the process of acquiring a photo with an ordinary digital camera is similar to the diagram illustrated in Figure 1. The light from the scene enters a set of lenses and passes through an anti-aliasing filter before reaching a colour filter array (CFA) that is intended to admit one of the red (R), green (G) and blue (B) components of the light per pixel for the following semi-conductor sensor to convert the signal into electronic form. A de-mosaicing process is subsequently carried out to get the intensities of the other two colours for each pixel by interpolating the colour information within a neighbourhood. A sequence of image processing operations, such as colour correction, white balancing, Gamma correction, enhancing, JPEG compression, etc. then take place before the photo is saved in the storage medium. The hardware or software used in each stage in the image acquisition pipeline as illustrated in Figure 1 may leave unique traces in images, which can lead to the identification of the imaging device. As such, to help with forensic investigations, researchers have proposed ways of identifying and linking source devices, classifying images and verifying the integrity of images based on the detection of existence or local inconsistencies of device attributes or data processing related characteristics, such as sensor pattern noise (SPN) [1-8], camera response function [9], re-sampling artefacts [10], colour filter array (CFA) interpolation artefacts [11, 12], JPEG compression [13, 14], lens aberration [15, 16], etc. Other device and image attributes such as

binary similarity measures, image quality measures and higher order wavelet statistics have also been exploited to identify and classify source devices [17 - 19].

While many methods [9-12] require that specific assumptions be satisfied, methods based on sensor pattern noise [1-8, 20-22] have drawn much attention due to the relaxation of the similar assumptions. Another advantage of sensor pattern noise is that it can identify not only camera models of the same make, but also individual cameras of the same model [1, 6]. The deterministic component of sensor pattern noise (SPN) is mainly caused by imperfections during the sensor manufacturing process and different sensitivity of pixels to light due to the inhomogeneity of silicon wafers [23, 24]. It is because of the inconsistency and the uniqueness of manufacturing imperfections and the variable sensitivity of each pixel to light that even sensors made from the same silicon wafer would possess uncorrelated pattern noise, which can be extracted from the images produced by the devices. This property makes sensor pattern noise a robust fingerprint for identifying and linking source devices and verifying the integrity of images. The reader is referred to [23] and [24] for more details in relation to sensor pattern noise.

II. Limitation of Existing SPN Extraction Model

Because sensor pattern noise appears as high-frequency signal in images, most image forensic methods based on sensor pattern noise [2-8] adopt the model proposed in [1] or its variant [25] for extracting the SPN, n, from an image I. The model is formulated as

$$n = DWT(I) - F(DWT(I)) \qquad (1)$$

where DWT is the Discrete Wavelet Transform and F is a denoising function, which filters out the sensor pattern noise in the DWT domain. Although various denoising filters can be used as F, the wavelet-based denoising filter described in Appendix A of [1] has been reported as effective in producing good results. We can see from Eq. (1) that the SPN, n, literally covers the high-frequency components of I.

The key limitation of Eq. (1) is that the SPN, n, can be severely contaminated by details from the scene because scene details account for the high-frequency components of I and their magnitude

is far greater than that of sensor pattern noise. For example Figure 2(a), (b) and (c) show a *reference* SPN of a camera, which is the average SPN of 50 images of blue sky taken by a digital camera, the image of a natural scene taken by the same camera, and the SPN extracted from the image of Figure 2(b), respectively. Figure 2(a) is what a "clean" SPN should look like. However, from Figure 2(c) we can see that the SPN contains strong details from the scene, which dominates the real SPN. Note the intensity of Figure 2(a) and (c) has been up scaled 9 and 3 times for visualisation purpose.

In the scenario of SPN-based source device identification, the investigator usually has a collection of devices or a database of *reference* sensor pattern noises, each representing one device, in his/her possession. The reason of creating such a clean reference SPN – the average of a number (say 20 to 50) of SPNs extracted from natural images, as illustrated in Figure 2(a) - is that it can better represent the imaging camera. However, source camera linking, which is about establishing whether or not the images under investigation are taken by the same camera without the camera and its reference SPN in the investigator's possession, is a more challenging problem than source camera identification. The investigation can only be carried out based on one SPN from each image and if one or more SPNs are severely contaminated by the details of the scenes, the chance of reaching a correct conclusion cannot be expected to high. An even more challenging application is blind / unsupervised image classification aiming at classifying a large set of images in the absence of the imaging cameras and reference SPNs. Given a large number of images, classification based on the SPNs extracted from images of their original size (e.g., 3 or 4 mega pixels) is computationally prohibitive. This entails the need for carrying out the classification task based on the SPNs from smaller blocks cropped from the original images. However, cropping reduces the number of SPN components, consequently increasing the intra-class variation of SPNs. To address these issues, the contaminated SPN needs to be cleaned or enhanced in some way. Although enhancing the SPN extracted with Eq. (1) has been attempted by Chen et al. [25], their objective is to attenuate the artefacts due to colour interpolation, row-wise and column-wise operation and JPEG compression, rather than to deal with scene interference. To our best knowledge, SPN enhancing methodology aiming at attenuating the interference from scene details is currently lacking. It is therefore our

intention to propose a method for effectively enhancing sensor pattern noise in Section III and to report in Section IV a sequence of experiments carried out to test the proposed SPN enhancers.

III. Proposed Sensor Pattern Noise Enhancer

Given the fact that the magnitude of scene details tend to be far greater than that of the sensor pattern noise, , as demonstrated in Figure 2(c), the hypothesis underlying our SPN enhancer is that

The stronger a signal component in n is, the more likely that it is associated with strong scene details, and thus the less trustworthy the component should be.

This hypothesis suggests that *an enhanced fingerprint n_e can be obtained by assigning less significant weighting factors to strong components of n in the Digital Wavelet Transform (DWT) domain in order to attenuate the interference of scene details.* There are various mathematical models for realising the afore-mentioned hypothesis. In this work, we propose five models, as formulated in Eq. (2) to (6) to be applied to the unenhanced SPN extracted with Eq. (1) in conjunction with the wavelet-based denoising filter described in Appendix A of [1].

Model 1:
$$n_e(i,j) = \begin{cases} \frac{n(i,j)}{\alpha} & \text{,if } 0 \le n(i,j) \le \alpha \\ e^{-0.5\frac{(n(i,j)-\alpha)^2}{\alpha^2}} & \text{,if } n(i,j) > \alpha \\ \frac{n(i,j)}{\alpha} & \text{,if } -\alpha \le n(i,j) < 0 \\ -e^{-0.5\frac{(n(i,j)+\alpha)^2}{\alpha^2}} & \text{,if } n(i,j) < -\alpha \end{cases} \tag{2}$$

Model 2:
$$n_e(i,j) = \begin{cases} \frac{n(i,j)}{\alpha} & \text{,if } 0 \le n(i,j) \le \alpha \\ e^{\alpha-n(i,j)} & \text{,if } n(i,j) > \alpha \\ \frac{n(i,j)}{\alpha} & \text{,if } -\alpha \le n(i,j) < 0 \\ -e^{\alpha+n(i,j)} & \text{,if } n(i,j) < -\alpha \end{cases} \tag{3}$$

Model 3:
$$n_e(i,j) = \begin{cases} 1 - e^{-n(i,j)} & \text{, if } 0 \le n(i,j) \le \alpha \\ (1 - e^{-\alpha}) \cdot e^{\alpha - n(i,j)} & \text{, if } n(i,j) > \alpha \\ -1 + e^{n(i,j)} & \text{, if } -\alpha \le n(i,j) < 0 \\ (-1 + e^{-\alpha}) \cdot e^{\alpha + n(i,j)} & \text{, if } n(i,j) < -\alpha \end{cases} \tag{4}$$

Model 4:
$$n_e(i,j) = \begin{cases} 1 - \frac{n(i,j)}{\alpha} & \text{, if } 0 \le n(i,j) \le \alpha \\ -1 - \frac{n(i,j)}{\alpha} & \text{, if } -\alpha \le n(i,j) < 0 \\ 0 & \text{, otherwise} \end{cases} \tag{5}$$

Model 5:
$$n_e(i,j) = \begin{cases} e^{-0.5 n^2(i,j)/\alpha^2} & \text{, if } 0 \le n(i,j) \\ -e^{-0.5 n^2(i,j)/\alpha^2} & \text{, otherwise} \end{cases} \tag{6}$$

where $n(i,j)$ and $n_e(i,j)$ are the (i,j)th component of n and n_e, respectively. These five models can also be better presented graphically as demonstrated in Figure 3(a) to (e). Eq. (2) − (4) allow the magnitude of n_e to *grow* monotonically in accordance with the magnitude of n if $|n| \le \alpha$ (a threshold to be decided by the user) and to *decrease* monotonically and rapidly with respect to $|n|$ if $|n| > \alpha$ while Eq. (5) and (6) allow the magnitude of n_e, (i.e., $|n_e|$) to *decrease* monotonically with respect to the magnitude of n. We can see that α of Eq. (2) to (6) determines the performance of each model. These five models are not picked at random, but are motivated by the following considerations.

- Stronger SPN components ($|n| > \alpha$) should be attenuated monotonically and rapidly with respect to $|n|$ to suppress the influence from scene details. This conforms to the falling tails in all five models, starting from the points where $|n|$ becomes greater than α, although the falling rates are different for different models.

- For weaker SPN components (i.e., $|n| \le \alpha$), different considerations as discussed later are reflected in the five models.

  - *Linear* transformation (Models 1 and 2), as Eq. (2) to (3) and Figure 3(a) to (b) suggest: This is to give those weak components the same weight ($1/\alpha$) and is the most

conservative transformation. However, since how scene details can be theoretically modelled is unclear, empirical tuning of the significance of the weaker (more trustworthy) components in some way other than linear transformation should also be studied. As such, the following two types of transformation are also considered.

-   *Non-linear exponential* transformation (Model 3), as formulated in Eq. (4) and illustrated in Figure 3(c): Like the linear transformation, this non-linear exponential transformation is also a moderate operation because the orders of the transformed components remain unchanged. However, by the gradients at various points of the transformation curves, we can see that the model gives greater significance to the SPN components on the lower ends and less significance to those closer to $\pm \alpha$, while Models 1 and 2 indiscriminatively give equal weight to every n in the range $[-\alpha, \alpha]$. It is worth noting that not any non-linear exponential model with a monotonically increasing (decreasing) transformation curve in the range $0 < n < \alpha$ ($0 > n > \alpha$) can produce effective SPN enhancement. For example, a *non-linear exponential* transformation (Model 6), as formulated in Eq. (7) and shown in Figure 3(f), does not make physical sense and should be avoided because, by the gradients at various points of the curves, we can see the model is giving less significance to the weaker but more trustworthy components than the stronger but less trustworthy ones. We will discuss this in Part A of Section IV.

$$\text{Model 6:} \quad n_e(i,j) = \begin{cases} 1 - e^{-n(i,j)} & , \text{if } 0 \leq n(i,j) \leq \alpha \\ (1 - e^{-\alpha}) \cdot e^{-\alpha + n(i,j)} & , \text{if } n(i,j) > \alpha \\ -1 + e^{n(i,j)} & , \text{if } -\alpha \leq n(i,j) < 0 \\ (-1 + e^{-\alpha}) \cdot e^{-\alpha - n(i,j)} & , \text{if } n(i,j) < -\alpha \end{cases} \quad (7)$$

-   *Inversely proportional* transformation (Models 4 and 5), as formulated in Eq. (5) and (6), and illustrated in Figure 3(d) and 3(e): These are the most radical transformations among all models because they reverse the order of the magnitude (e.g., 0 in the unenhanced n is mapped to the maximum value of 1 in enhanced n_e). This is intended to lay even more trust on the components with low magnitude. Therefore, they are still

in consistence with our hypothesis because, throughout the entire spectrum, the weaker components are given greater significance than the stronger ones.

IV. Experiments

In the following experiments, we use 1200 photos of 1536 × 2048 pixels taken in JPEG format (with JPEG quality factor approximately ranging from 93 to 97) by six cameras, each responsible for 200. The six cameras are Canon IXUS 850IS, Canon PowerShot A400, Canon IXY Digital 500, FujiFilm A602, FujiFilm FinePix A902 and Olympus FE210. The photos contain a wide variety of natural indoor and outdoor scenes taken during holidays, around campus and cities, in offices and laboratories, etc. To enhance a SPN, we first perform Discrete Wavelet Transform (DWT), conduct low-pass filtering in the DWT domain, extract the SPN n using Eq. (1) in DWT domain, and finally apply an enhancement model to the unenhanced SNP directly in the DWT domain to get the enhanced version n_e.

Each *reference* SPN, which represents each of the six cameras, is generated by calculating the average of the SPNs extracted from 50 photos of blue sky taken by the digital camera. Note because the photos of the blue sky do not contain significant high-frequency details and 50 SPNs are averaged to generate the reference SPN, therefore we did not apply any enhancing model to enhance those photos of blue sky. The 50 photos for creating the reference SPN are not included in the test set in the following experiments.

Source device identification requires similarity comparisons among SPNs, therefore the feasibility of the chosen similarity metrics is important. As proposed in [22], Fridrich suggested the use of the Peak to Correlation Energy (PCE) measure, which has proved to be a more stable detection statistics than normalised cross-correlation when applied to the scenarios in which the images of interest may have undergone geometrical manipulations, such as rotation or scaling. The purpose of this work is to demonstrate the capability of the proposed SPN enhancers in dealing with the interference of details from the scene, geometrical transformations will not be applied in order to prevent biased evaluation

from happening. Therefore, in the following experiments, normalised cross-correlation will be used to measure the similarity between SPNs. The normalised cross-correlation between signal n_i and n_j is defined as

$$\rho(i,j) = \frac{(n_i - \bar{n}_i) \cdot (n_j - \bar{n}_j)}{\left\| n_i - \bar{n}_i \right\| \cdot \left\| n_j - \bar{n}_j \right\|}, \quad j \in [1, 6] \tag{8}$$

where $\bar{n}_i$ and $\bar{n}_j$ are the means of n_i and n_j, respectively.

A. Selection of Enhancing Model and Parameter

The main theme of this work is the conception of the hypothesis that **the stronger a signal component in n is, the more likely that it is associated with strong scene details, and thus the less trustworthy the component should be,** while the five models (Eq. (2) to (6)) are just to validate the hypothesis. There is no theoretical backing for choosing the optimal model from Eq. (2) to (6) because the theory for modelling sensor pattern noise and scene details is not in existence at present. Feasible models other than these five can certainly be adopted in the future if found.

We have carried out a sequence of source camera identification experiments, based on 1200 image blocks of 128 × 128 pixels cropped from the centre of the afore-mentioned 1200 photos, to evaluate various combinations of the five models (i.e., Eq. (2) – (6)) and 30 different values of α in order to validate our hypothesis. As we will demonstrate in Part B of Section IV and Table 2, the reason of using image blocks of this size is that the performance of the models are not close to 100% when image blocks of this size is used, which leaves room for revealing the real performance of each model. To identify the source camera of an image, the SPN is extracted from the image and the similarity between the SPN and each of the six reference SPNs is calculated using Eq. (8). The image

is deemed as taken by the camera corresponding to the maximum of the six similarity values. The results are listed in Table 1 and plotted in Figure 4. The following observations can be made:

- Models 1 and 2, formulated in Eq. (2) and (3), perform reasonably well with the value of α in the relatively smaller ranges of [3, 4] and [4, 6], respectively, when compared to the performance of the other three models. However, as can be seen in Figure 4, their performance curves drop rapidly as the value of α grows. This indicates that SPN enhancement through *linear* transformation when $|n| < \alpha$ is more sensitive to changes of α. Moreover, the only difference between Models 1 and 2 is that the attenuation rate of Model 2 is greater than Model 1 when $|n| > \alpha$ (See Figure 3(a) and (b)). This factor accounts for the more moderate declining rate of performance of Model 2 than that of Model 1 after their respective performance peaks, as shown in Figure 4 and Table 1, and indicates that *a greater attenuation rate is preferable for strong SPN components*.

- Model 3 applies *non-linear exponential* transformation to SPN components when $|n| < \alpha$. Figure 4 indicates that it performs stably well in a wider range [4, 11] of α, with a peak identification rate of 1039 out 1200 images at $\alpha = 5, 6$ and 9 (see Table 1). Moreover, its performance curve drops more gracefully than Model 1 and 2 as α grows. It is worth noting that, according to Eq. (3) and (4), the transformation employed in Model 2 for $|n| > \alpha$ is basically the same as that employed in Model 3, except that the latter has a factor of $\pm (1 - e^{-\alpha})$ which is $\approx \pm 1$. So we can conclude that the performance difference between the two models is due to the non-linear transformation effect when $|n| < \alpha$, as discussed at the end of Section III. The explanation for this effect is that, as shown in Figure 3(b) and (c), when $|n| < \alpha$, the gradients at various points of the transformation curve of Model 2 remains *constant* while the gradients of Model 3 *decreases monotonically* with respect to $|n|$. This means Model 2 indiscriminatively assigns an equal weight to every component when $|n| < \alpha$ while Model 3 adaptively decreases the weight as $|n|$ grows (i.e., as the influence of scene details gets stronger).

- Model 4 and 5 apply *inversely proportional* transformation to the SPN components when $|n| < \alpha$. Both models have equivalent peak identification rate of 1039 and 1040 out of 1200 images, respectively. Model 4 performs at peak level when $\alpha = 18$, which is far greater then the value of α ($\alpha = 7$) at which Model 5's performance peaks. This is because when α is lower the slope of the straight transformation line of Model 4 is greater, and as a result, the small and trustworthy components get over-attenuated. However, as shown in Figure 4, Model 4's performance appears to be marginally more stable than Model 5's after its performance peaks. This is because Model 4 sets n to 0 when $|n| > \alpha$.

- Although Model 6's peak performance level (1014/1200 when $\alpha = 3$) is only 2.17% lower than the global peak (1040/1200 of Model 5 when $\alpha = 7$), this model is not only counter-intuitive but also inconsistent with our hypothesis. The main difference between Model 3 and Model 6 is that when $|n| < \alpha$, their transformation curves go up towards $\pm \alpha$ with *decreasing* and *increasing* gradients, respectively. This indicates that, within this range, while Model 3 gives greater weight to the small and trustworthy components, Model 6 does the opposite. Consequently, as its corresponding plot in Figure 4 shows, its performance is highly sensitive to the value of α.

From the above discussions, we can conclude that Model 1 to 5 are all feasible models for enhancing SPNs, with Models 3, 4 and 5 being more preferable because they perform more stable within wider ranges of values of α. Stability is important because it gives the user high confidence in their choices. We also observed that the highest performance level (1040/1200) is reached by Model 5 with $\alpha = 7$. However, this does not mean that this is the optimal combination because theoretical approaches for finding the optimal model and its parameters are currently lacking and it is in no way possible to exhaust the infinite numbers of models and parameters to identify the optimal combination.

B. Source Camera Identification

To validate our hypothesis, we have carried out camera identification tests on the 1200 photos using Model 5 with $\alpha = 7$. Instead of testing the enhancer on the full-sized images of 1536×2048 pixels only, we also test it on image blocks of 8 different sizes cropped from the centre of the full-sized images. Moreover, in real applications, identification should be based on *whether the similarity is greater than a feasible threshold*. Table 2 shows the *true positive* rate *with* and *without* applying Model 5 to the sensor pattern noises extracted with Eq. (1) when a correlation threshold of $t = 0.01$ is applied. Note that in this experiment the SPN of each image is only compared to the reference SPN of the camera that actually took the image in question, i.e., the *source* camera. The image is deemed as taken by the source camera if the similarity value is greater than t. It is clear from Table 2 that the larger the image blocks are, the greater the performance becomes. We can also see that, in all cases, enhancing the SPNs always yields greater performance and the performance differences become more significant as the image blocks get bigger.

Another useful measure for demonstrating the performance of the methods is *false positive*. Table 3 shows the *false positive* rates when a correlation threshold of 0.01 is applied. Note that, in this experiment, the SPN of each image is compared to the 5 reference SPNs of the cameras that are *not* the source camera of the image in question. The image is deemed as taken by the cameras that are *not* the source camera if their similarity values are greater than a threshold 0.01. From Table 3, the performance differences are even more prominent when the image block sizes are small. An interesting phenomenon, which can be observed from Table 3, is that for both methods, when scanning from the right hand side of the table, the false positive rates decrease slightly and reach the minimum when the image block size is 1024×1024 pixels. The false rates then increase significantly afterwards. This is particularly clear for the case without enhancement. After applying other threshold values of 0.005, 0.015, 0.02, 0.025 and 0.03, we observed the same phenomenon. We have no explanation for this at present, but it is interesting to look into the reasons in the future.

Table 2 and 3 have validated the hypothesis and demonstrated the superiority of the proposed SPN enhancing model. Figure 2(d) shows the enhanced version of Figure 2(c) after Model 5, with α

= 7, is applied. We can see that the influential details from the scene, that are prominent in Figure 2(c), have been significantly removed from Figure 2(d).

C. Impact of Colour Saturation

In many photos, the upper-left or upper-right corners are homogeneous background, such as the sky or a wall of plain colour, where the sensor pattern noise is less contaminated by details from the scenes than other areas. Therefore, if only a block is to be taken from a photo for forensic analysis, either one of these two corners are good candidates because the probability of getting a low-variation block from these two corners are greater than from other areas. Based on this rationale, we have also carried out the same camera identification experiment on image blocks of 128 × 128 pixels cropped from theses two corners and the centre of the 1200 photos. The results are listed in Table 4. Each number in the "No. saturated blocks" row is the number of saturated blocks out of 1200 blocks cropped from different areas of interest. In our experiment, if over 50% of the pixels of a block have the intensities of all three colour channels equal to 255, the block is deemed as saturated. The "No. saturated blocks" row conforms to our expectation that the two corners at the top of photos are more likely to be saturated than the central area. The "Identification rate (%): Saturation *included*" row of Table 4 shows that when the saturated blocks are included in the identification experiment, the identification rates based on the blocks cropped from different areas of interest are almost the same. Note that conclusion could not be drawn from this row alone, because these three statistics may vary when different dataset is used. However, this row is helpful in demonstrating the impact of colour saturation when comparing the statistics in the "Identification rate (%): Saturation *excluded*" row. This later row indicates that, when those saturated blocks are excluded, the identification rates based on the blocks cropped from the two corners are significantly higher than that based on the blocks cropped from the centre of images. This is not a surprising observation because usually the main objects appear in the centre of photos, where normal imaging and illumination conditions are met, while the two corners at the top of photos are more likely to be saturated due to imaging and

illumination conditions, thus giving rise to the loss of sensor pattern noise. So we suggest that blocks be taken from the centre of photos if the SPNs of small image blocks cropped *automatically* by the system are to be used for forensic applications, such as unsupervised image classification.

V.Conclusions

In this work we have pointed out that sensor pattern noise, as the fingerprint for identifying source imaging devices, extracted with the commonly used model of Eq. (1) proposed in [6] can be severely contaminated by the details from the scene. To circumvent this limitation we envisaged the hypothesis that the stronger a component of the sensor pattern noise is, the less trustworthy the component should be and proposed 5 enhancing models (Model 1 to 5) for realising the hypothesis, with Model 3, 4 and 5 being more preferable. The hypothesis is tested by assigning greater weighting to the smaller SPN components. Experiments on source device identification have confirmed the soundness of our hypothesis.

Another related digital forensics application is that there are circumstances where a forensic investigator has a large set of images taken by an unknown number of unknown digital cameras and wishes to cluster those images into a number of classes, each including the images acquired by the same camera. The main challenges in this scenario are:

- The forensic investigator does not have the cameras that have taken the images to generate reference SNPs for comparison.
- No prior knowledge about the number and types of the imaging devices are available.
- With a large dataset, exhaustive and iterative pair-wise SPN comparison is computationally prohibitive.
- Given the shear number of images, analysing each image in its full size is computationally infeasible.

In the near future, we intend to devise an unsupervised image classifier based on the enhanced sensor pattern noise using our SPN enhancers to address the afore-mentioned issues.

Acknowledgements

References

[1] J. Lukáš, 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, June 2006.

[2] M. Goljan, M. Chen and J. Fridrich, "Identifying Common Source Digital Camera from Image Pairs," in *Proc. IEEE International Conference on Image Processing*, pp. 14-19, San Antonio, USA, September 2007.

[3] Y. Sutcu, S. Batram, H. T. Sencar and N. Memon, "Improvements on Sensor Noise based Source Camera Identification," in *Proc. of IEEE International Conference on Multimedia and Expo*, pp. 24 – 27, Beijin, China, 2 - 5 July 2007.

[4] R. Caldelli, I. Amerini, F. Picchioni and A. De Rosa and F. Uccheddu, "Multimedia Forensic Techniques for Acquisition Device Identification and Digital Image Authentication," in *Handbook of Research on Computational Forensics, Digital Crime and Investigation: Methods and Solutions*, C.-T. Li (Ed.), Hershey, PA: Information Science Reference (IGI Global), Nov. 2009.

[5] R. Caldelli, I. Amerini and F. Picchioni, "Distinguishing between Camera and Scanned Images by Means of Frequency Analysis," *International Journal of Digital Crime and Forensics*, vol. 2, no. 1, Jan – March 2010.

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

[7] C.-T. Li, "Methods for Identifying Imaging Devices and Classifying Images Acquired by Unknown Imaging Devices," *Pending UK Patent*, Application Number 0902406.5.

[8] N. Khanna, G. T.-C. Chiu, J. P. Allebach and E. J. Delp, "Forensic Techniques for Classifying Scanner, Computer Generated and Digital Camera Images." In *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Las Vegas, USA, 30 March - 4 April, 2008.

[9] Y. F. Hsu and S. F. Chang, "Image Splicing Detection Using Camera Response Function Consistency and Automatic Segmentation," in *Proc. IEEE International Conference on Multimedia and Expo*, Beijin, China, 2 - 5 July 2007.

[10] A. C. Popescu and H. Farid, "Exposing Digital Forgeries by Detecting Traces of Resampling." *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 758–767, 2005.

[11] A.C. Popescu and H. Farid, "Exposing Digital Forgeries in Color Filter Array Interpolated Images." *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3948–3959, 2005.

[12] A. Swaminathan, M. 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, March 2007.

[13] M. J. Sorell, "Digital Camera Source Identification through JPEG Quantisation," in *Multimedia Forensics and Security*, C.-T. Li (Ed.), Hershey, PA: Information Science Reference (IGI Global), 2008.

[14] M. J. Sorell, "Conditions for Effective Detection and Identification of Primary Quantisation of Re-Quantized JPEG Images," *International Journal of Digital Crime and Forensics*, vol. 1, no. 2, pp.13-27, April - June 2009.

[15] S. Choi, E. Y. Lam and K. K. Y. Wong, "Source Camera Identification Using Footprints from Lens Aberration," in *Proceedings of the SPIE* 2006.

[16] V. T. Lanh, S. Emmanuel, M. S. Kankanhalli, "Identifying Source Cell Phone Using Chromatic Aberration," in *Proc. IEEE Conference on Multimedia and Expo*, Beijin, China, 2 - 5 July 2007.

[17] B. Sankur, O. Celiktutan and I. Avcibas, "Blind Identification of Cell Phone Cameras," in *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX*, vol. 6505, January 29–February 1, San Jose, CA, pp. 1H–1I, 2007.

[18] G. Xu, S. Gao, Y. Q. Shi, W. Su and R. Hu, "Camera-Model Identification Using Markovian Transition Probability Matrix," in *Proc. International Workshop on Digital Watermarking*, pp. 294-307, Guildford, UK, 24-26, August, 2009.

[19] P. Sutthiwan, J. Ye and Y. Q. Shi, "An Enhanced Statistical Approach to Identifying Photorealistic Images," in *Proc. International Workshop on Digital Watermarking*, pp. 323-335, Guildford, UK, 24-26, August, 2009.

[20] S. Bayram, H. T. Sencar and N. Memon, "Video Copy Detection Based on Source Device Characteristics: a Complementary Approach to Content-Based Methods," in *Proc. of the First ACM International Conference on Multimedia Information Retrieval*, pp. 435 - 442, Vancouver, Canada, 30 - 31 October, 2008.

[21] H. Gou, A. Swaminathan and M. Wu, "Intrinsic Sensor Noise Features for Forensic Analysis on Scanners and Scanned Images," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 3, pp. 476 - 491 Sept. 2009

[22] J. Fridrich, "Digital Image Forensic Using Sensor Noise," *IEEE Signal Processing Magazine*, vol. 26, no. 2, pp. 26-37, March 2009.

[23] J. R. Janesick, *Scientific Charge-Coupled Devices*, Bellingham, WA: SPIE, vol. PM83, 2001.

[24] T. Yamada, "CCD Image Sensors," in *Image Sensors and Signal processing for Digital Still Cameras*, ed. by J. Nakamura, Taylor & Francis Group, 2006.

[25] M. Chen, J. Fridrich and M. Goljan, "Digital Imaging Sensor Identification (Further Study)," in *Proc. of SPIE Electronic Imaging*, Photonics West, January 2007.
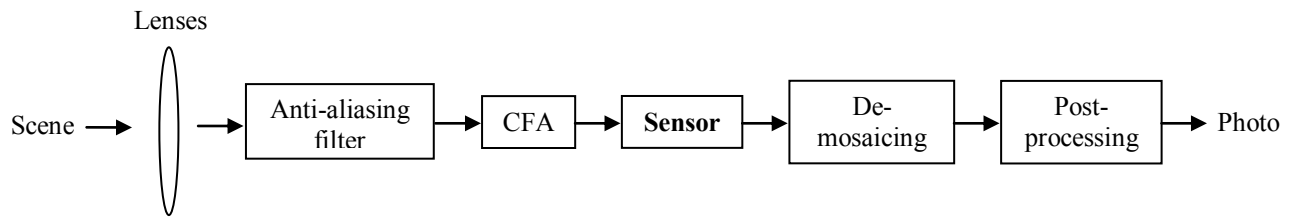
Figure 1. The image acquisition process of an ordinary digital camera.
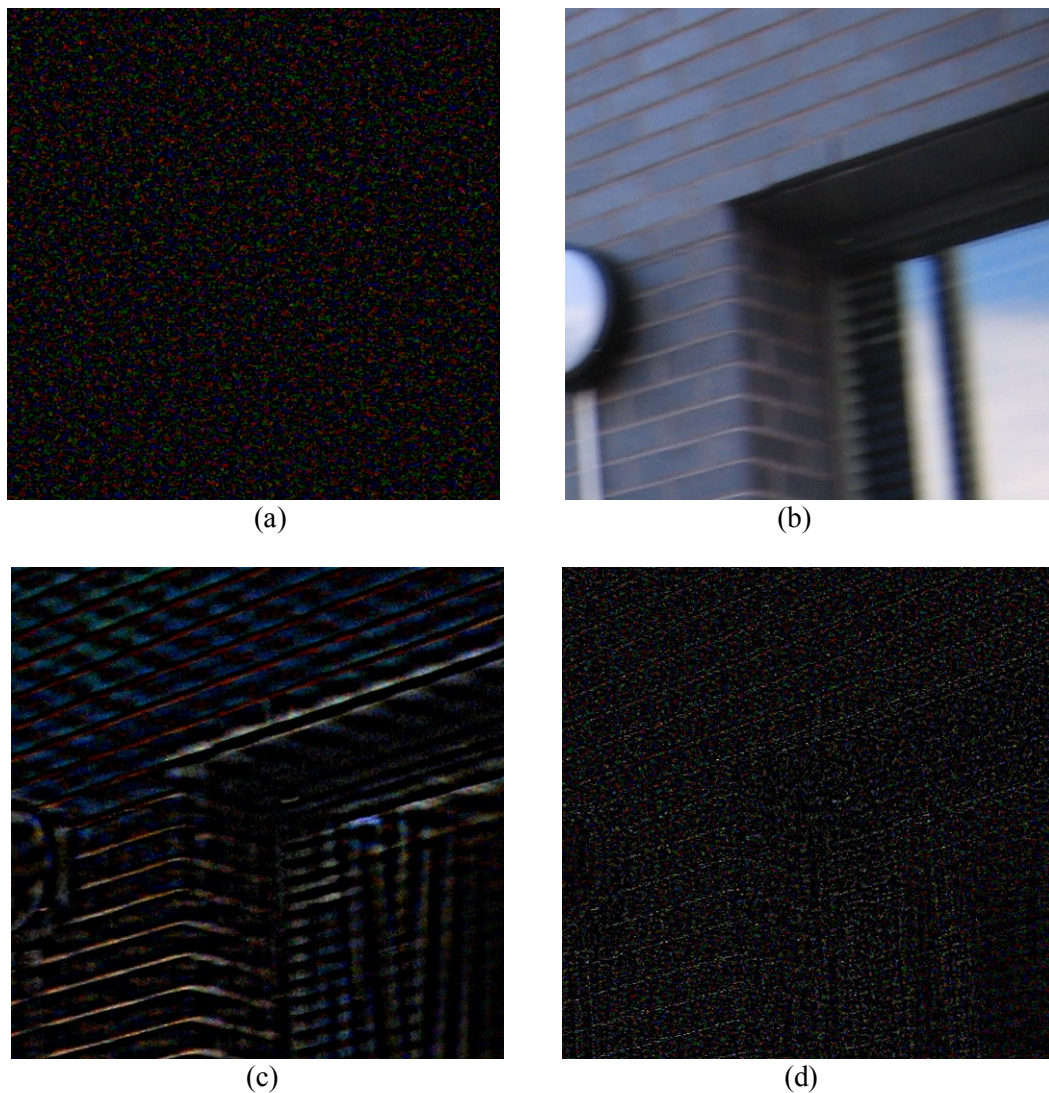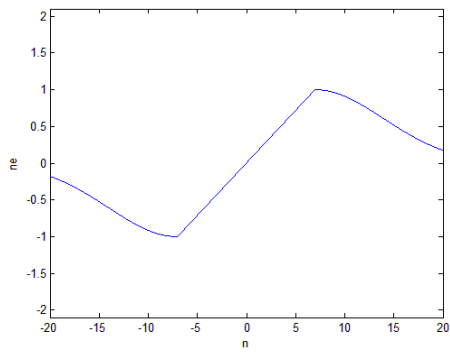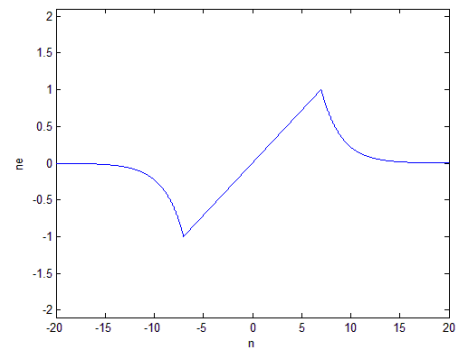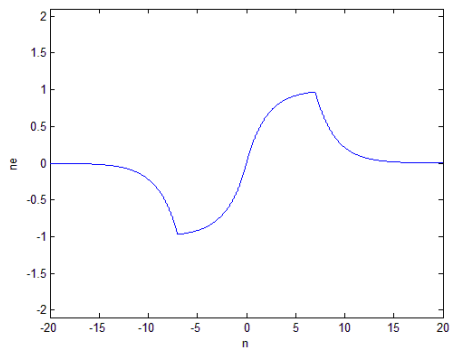

(a)

(b)

(c)

(d)

Figure 2. (a) A clean *reference* SPN taken from blue sky images, (b) An image of natural scene, (c) The SPN extracted from Figure 2(b) that is contaminated by the details from the scene. (d) The enhanced version of Figure 2(c) using Model 5 (i.e., Eq. (6)) with $\alpha = 7$. Note the intensity of Figure 2(a) and (c) has been up scaled 9 and 3 times, respectively, for visualisation purpose.
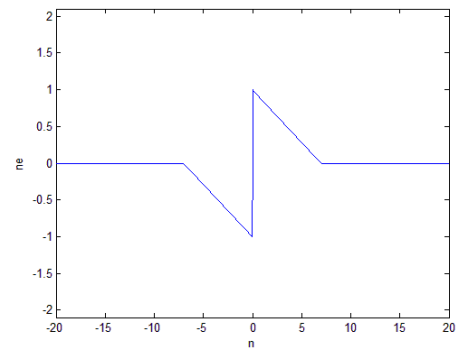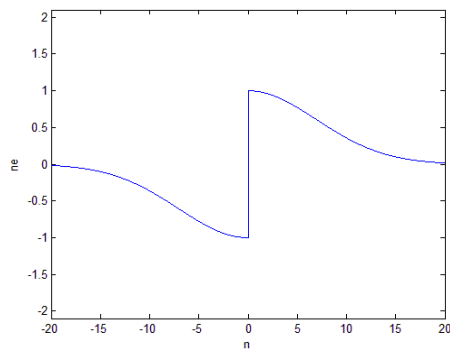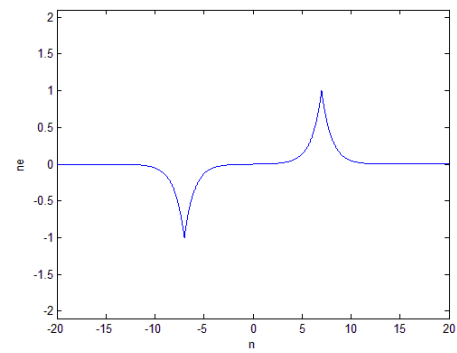
Figure 3. Six models for digital fingerprint enhancement. (a) –(f) correspond to Eq. (2) – (7), respectively.
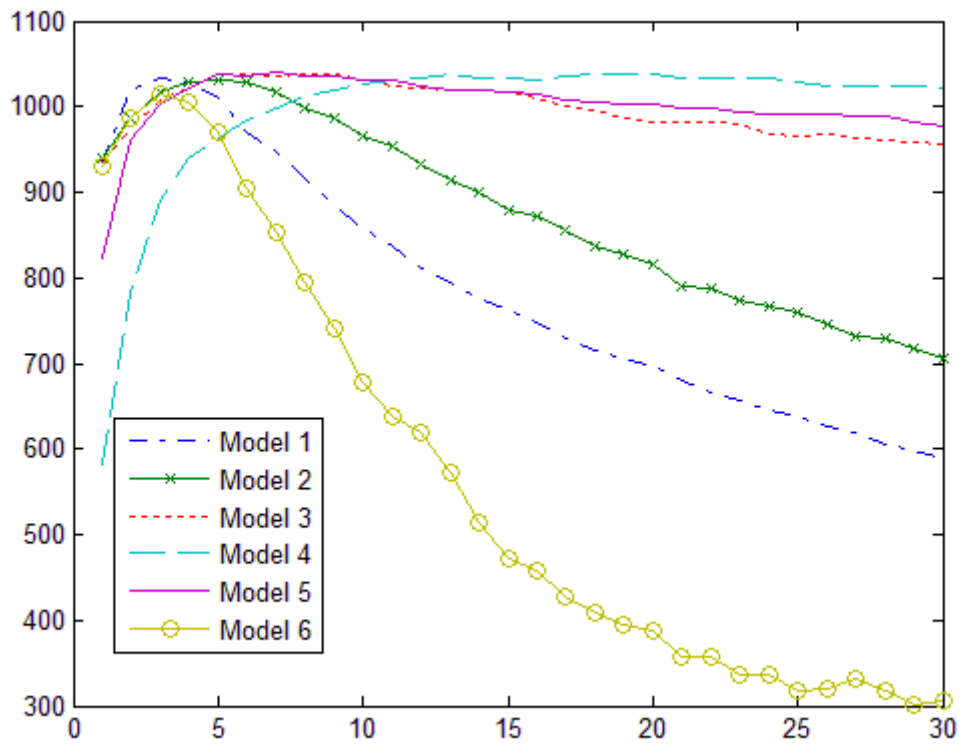
Figure 4. Performance, in terms of *number of correct source camera identifications out of 1200 images*, of various SPN enhancing models when applied in conjunction with different values of α.

Table 1. Performance, in terms of *number of correct source camera identifications out of 1200 images*, of various SPN enhancing models when applied in conjunction with different values of α.

Model	α														
	1	2	3	4	5	6	7	8	9	10	11	1	13	14	15
1	934	1020	**1033**	1029	1010	971	947	916	883	859	837	811	794	776	762
2	940	986	1017	1029	**1032**	1029	1018	999	987	966	954	932	914	899	880
3	936	976	1008	1021	1039	**1039**	1036	1038	1039	1032	1024	1021	1020	1019	1016
4	582	783	890	940	964	985	998	1012	1020	1027	1031	1033	1037	1034	1033
5	823	960	1003	1021	1039	1035	**1040**	1036	1036	1031	1030	1024	1019	1020	1017
6	931	987	**1014**	1006	970	904	853	795	741	678	637	619	573	514	473

Model	α														
	1	2	3	4	5	6	7	8	9	10	11	1	13	14	15
1	934	1020	**1033**	1029	1010	971	947	916	883	859	837	811	794	776	762
2	940	986	1017	1029	**1032**	1029	1018	999	987	966	954	932	914	899	880
3	936	976	1008	1021	1039	**1039**	1036	1038	1039	1032	1024	1021	1020	1019	1016
4	582	783	890	940	964	985	998	1012	1020	1027	1031	1033	1037	1034	1033
5	823	960	1003	1021	1039	1035	**1040**	1036	1036	1031	1030	1024	1019	1020	1017
6	931	987	**1014**	1006	970	904	853	795	741	678	637	619	573	514	473

Table 2. True positive rates *with* and *without* applying Model 5 to the sensor pattern noises with α = 7. Note that in this experiment the image is deemed as taken by the *source* camera if the similarity value is greater than a threshold 0.01.

	True positive rate (%) at different photo sizes								
	128 ×128	128 × 256	256 × 256	256 × 512	512 × 512	512 ×1024	1024 ×1024	1024 ×2048	1536 ×2048
without enhancement	61.68	67.5	71.42	77.92	82.33	87.12	93.25	96.75	97.42
with enhancement	79.75	85.58	91.00	93.17	94.75	96.33	97.95	98.25	98.25

Table 3. False positive rates *with* and *without* applying Model 5 to the sensor pattern noises with α = 7. Note that in this experiment the image is deemed as taken by the cameras that are *not* the source camera if their similarity values are greater than a threshold 0.01.

	False positive rate (%) at different photo sizes								
	128 ×128	128 × 256	256 × 256	256 × 512	512 × 512	512 ×1024	1024 ×1024	1024 ×2048	1536 ×2048
without enhancement	41.68	38.68	32.60	25.71	16.28	6.75	1.90	2.40	12.03
with enhancement	8.33	3.22	0.95	0.15	0.03	0	0	0.03	0.4

Table 4. Identification rates with colour saturation taken into account.

	Area of interest		
	upper-left corner	centre	upper-right corner
No. saturated blocks	113	7	110
Identification rate (%): Saturation *included*	86.83	86.67	85.25
Identification rate (%): Saturation *excluded*	92.27	86.83	90.65