A reference estimator based on composite sensor pattern noise for source device identification

Ruizhe Li, Chang-Tsun Li, and Yu Guan

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

ABSTRACT

It has been proved that Sensor Pattern Noise (SPN) can serve as an imaging device fingerprint for source camera identification. Reference SPN estimation is a very important procedure within the framework of this application. Most previous works built reference SPN by averaging the SPNs extracted from 50 images of blue sky. However, this method can be problematic. Firstly, in practice we may face the problem of source camera identification in the absence of the imaging cameras and reference SPNs, which means only natural images with scene details are available for reference SPN estimation rather than blue sky images. It is challenging because the reference SPN can be severely contaminated by image content. Secondly, the number of available reference images sometimes is too few for existing methods to estimate a reliable reference SPN. In fact, existing methods lack consideration of the number of available reference images as they were designed for the datasets with abundant images to estimate the reference SPN. In order to deal with the aforementioned problem, in this work, a novel reference estimator is proposed. Experimental results show that our proposed method achieves better performance than the methods based on the averaged reference SPN, especially when few reference images used.

Keywords: Multimedia forensics, source device identification, sensor pattern noise, photo-response non-uniformity.

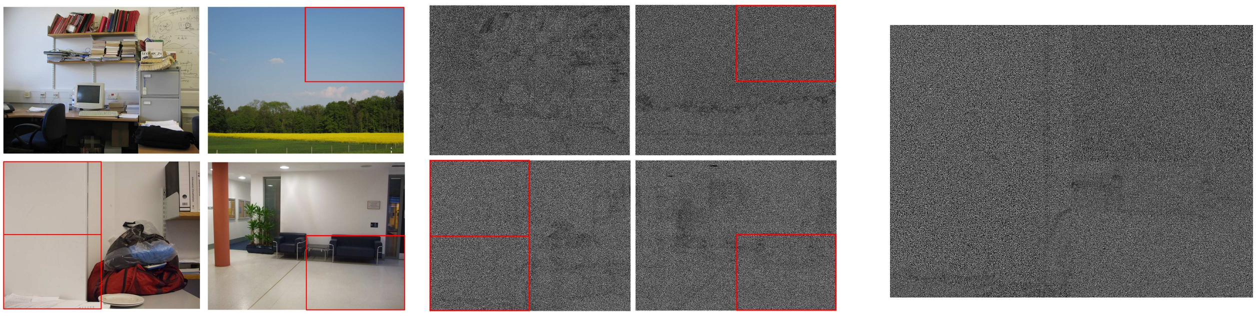
1. INTRODUCTION

With a large amount of digital imaging devices, the use of digital images as key piece of evidences in the fight against crime is emerging, making digital forensic techniques more important. In some cases, forensic investigators need to identify the origin of images in order to link the images to a suspect. Therefore, effective techniques for identifying the origin of digital images are urgently needed.

Sensor pattern noise has been proved as the most efficient forensic techniques for identifying source imaging device[1]. It can be used to identify individual cameras even of the same brand and model because of its uniqueness. Firstly, Lukas *et al.* [1] used a wavelet-based denoising technique [2] to extract the sensor pattern noise as the link between the query image and its origin. Since the main component in SPN is Photo Response Non-Uniformity noise (PRNU) which is a kind of multiplicative noise, Chen *et al.* [3] proposed a maximum likelihood method to estimate the corresponding multiplicative factor from the reference images. Later, Goljan *et al.* [4] introduced the Peak to Correlation Energy ratio (PCE) as a replacement for normalized correlation detector to reduce the false acceptance rate. In [5], Li pointed out that the noise residue contains significant characteristics of the SPN, but it can be easily affected by the image content. To address this issue, Li introduced an enhanced SPN by assigning higher weight to the reliable components. A further investigation into SPN's location-dependent quality is reported by Li and Satta in [6]. In [7], Kang *et al.* applied the context adaptive interpolation to predict the noise-free image for suppressing the impact of image content before SPN feature extraction.

The existing methods generally estimate the camera reference SPN from a number of blue sky images taken by the same digital camera. This method was applied because the smooth and bright images provide pure SPNs and averaging multiple images suppresses the random noises. However, in practice we may face the special case of source device identification, which is lacking the camera and its reference SPN in the investigators possession.

E-mail: ruizhe.li, c-t.li, g.yu@warwick.ac.uk



(a) (b) (c)

Figure 1. (a) Four natural images with scene details. All the four images are taken by the same camera. (b) Four SPNs extracted from the images in Figure (a). (c) A composite SPN consists of blocks collected from the SPNs in Figure 1(b).

Under this circumstance, it is more likely that only ordinary pictures with scene details are available for the reference SPN estimation rather than blue sky images. In addition, the number of available reference images sometimes is too few (< 20) for the existing methods to build a reliable reference SPN. Although there have been several studies dedicated to improving the performance of SPN based source camera identification, an efficient method for estimating the reference SPN from a limited number of images with scene details is still lacking. In this work, we propose a novel approach for estimating reliable reference SPN from a limited number of natural images so as to improve the performance of source device identification. Experimental results show that our proposed method achieves better performance than the schemes based on the averaged reference SPN especially when few reference images used.

The rest of this is organized as follows. In Section 2, we will first describe how to evaluate the quality of SPN blocks for reference SPN estimation. Then a composite reference estimator is proposed. Experimental results are reported in Section 3. The conclusion and future plan are set out in Section 4.

2. PROPOSED METHOD

The SPN extracted from images with strong scene details can be severely contaminated, as a result, the reference SPN estimated by averaging these SPNs is unsatisfactory. However, there should be some relatively smooth and bright region insides these images, which can provide cleaner SPNs than the regions with scene details (because the SPNs extracted from smooth and bright regions are more correlated to ground truth than those from the regions with scene details). Due to the variety of SPN quality within any image with scene details, a more reliable reference SPN can be achieved by assigning higher weight to the SPNs obtained from the smooth and bright regions rather than the averaging method treating the SPNs from all the regions equally. In order to facilitate the following weighting operation, we combine the SPNs with the same quality level (in terms of smoothness and brightness of its source image region) from different locations together to form a new SPN. To estimate the reference SPN, we just assign a certain amount of weighting factors to the composite SPN according to the level of its quality.

For example, Figure 1(a) shows four natural pictures which contain strong details from the scenes, and there are still some smooth and bright regions in each of them. We roughly divided each image into 4 blocks, and marked out the optimal blocks in terms of smoothness and brightness among the blocks in the same location of these 4 images. From Figure 1(b) we can see that the SPNs extracted from these optimal blocks contain less scene details than the SPNs from other blocks. As illustrated in Figure 1(c), a composite SPN, with optimal SPN blocks collected from Figure 1(b), contains cleaner SPN than any individual SPN in Figure 1(b). It suggests that assigning higher weight to the composite SPNs with SPN blocks collected from smoother and brighter image regions, a more reliable reference SPN can be obtained.

2.1 SPN Blocks Quality Measurement

As aforementioned, the quality of each SPN block depends on the smoothness and brightness of its source image region. In order to evaluate the quality of each SPN block for the reference SPN estimation, a measurement Q

is therefore proposed

$$Q(I_{Block}) = \begin{cases} \frac{B(I_{Block})}{E(I_{Block})}, & if 30 < B(I_{Block}) < 253 \\ 0, & else \end{cases} \tag{1}$$

where $B(I_{Block})$ is the average brightness of image block I_{Block}. Notice that SPNs extracted from dark or saturated regions are very weak. So we exclude a region I_{Block} if $B(I_{Block}) \leqslant 30$ or $B(I_{Block}) \geqslant 253$. $E(I_{Block})$ is the entropy of I_{Block}, which is a statistical measure of randomness that can be used to characterize the degree of details in images. The smoother the image region is, the lower the entropy it should be. It is calculated as follow

$$E(I_{Block}) = -\sum_{k=0}^{255} p_k \log_2 p_k \tag{2}$$

where k is a gray level value of a pixel, p_k is the probability of the gray level value k in the image region I_{Block}.

2.2 Methodology

The proposed reference SPN estimator is built in five steps and details for individual steps are shown as follow:
Step 1. Assume there are a set of images $G_i (i = 1, ..., N)$ taken by the same camera. Firstly, we extract raw SPN W_i from each G_i in the discrete wavelet domain, such that

$$W_i = DWT(G_i) - F(DWT(G_i)) \tag{3}$$

where DWT is discrete wavelet transform and F is the wiener filter. Here, we extract SPNs from Red, Green and Blue channel separately and combine them into W_i by using the linear combination as in RGB to grayscale conversion. The obtained W_i is then transformed to the spatial domain and normalized into $H_i (i = 1, ..., N)$.
Step 2. Each image G_i and SPN H_i is then segmented into a set of non-overlapping image blocks $g_{ij} (j = 1, ..., M)$ and SPN blocks $h_{ij} (j = 1, ..., M)$, respectively. So each image G_i and SPN H_i will have M subblocks from M locations. The size of g_{ij} and its corresponding h_{ij} is S. Each obtained image block g_{ij} is then used for evaluating the quality Q_{ij} of its corresponding SPN block h_{ij}. All the obtained SPN blocks are used as candidates for constituting the new SPNs.
Step 3. For each location j, SPN blocks $h_{ij} (i = 1, ..., N)$ from N images are sorted in the descending order according to their quality Q_{ij}. SPN block h_{ij} with better quality will have higher ranking (with 1 as the highest).
Step 4. SPN blocks from different locations with the same ranking are then selected to form a new SPN. By repeating this step N times, a set of N composite SPN $C_d (d = 1, ..., N)$ can be generated, each with the rank value d. Notice that, after this step, not only SPNs constituted clean SPN blocks but also SPNs which are full of scene details can be obtained. In this work, the composite SPNs composited by low-quality SPN blocks are also taken into account for the reference SPN estimation as they still can contribute to remove random noises, especially when the number of available images is limited.
Step 5. Instead of calculating the reference SPN by averaging the SPNs extracted from multiple images, we assign higher weight to the composite SPN C_d with higher ranking. Weighting factor ω_d for each composite SPN C_d is calculated as follow

$$\omega_d = \frac{2(N + 1 - d)}{N(N + 1)}, d \in [1, N] \tag{4}$$

Finally, the reference SPN r is estimated as the weighted average of the N composite SPNs

$$r = \sum_{d}^{N} \omega_d C_d, d \in [1, N] \tag{5}$$

Notice that different size S of image and SPN blocks will affect the estimated reference SPN r. Applying smaller S is more likely to achieve better performance. However, when S is too small (*e.g.*, 16×16 pixels) the computational complexity may become unacceptable. In order to find an proper setting of S, the performance in terms of the Receiver Operating Characteristic curve (ROC) curves of our proposed method measured on three different sizes of S (*i.e.* 16×16, 32×32 and 64×64 pixels) is shown in Figure 2. As illustrated in Figure 2, the results with three sizes are very similar, which implies the impact of SPN block size S is relatively low for reference SPN estimation. In this work, we use $S = 64 \times 64$ as it is 2 and 4 times faster than $S = 32 \times 32$ and $S = 16 \times 16$ pixels.

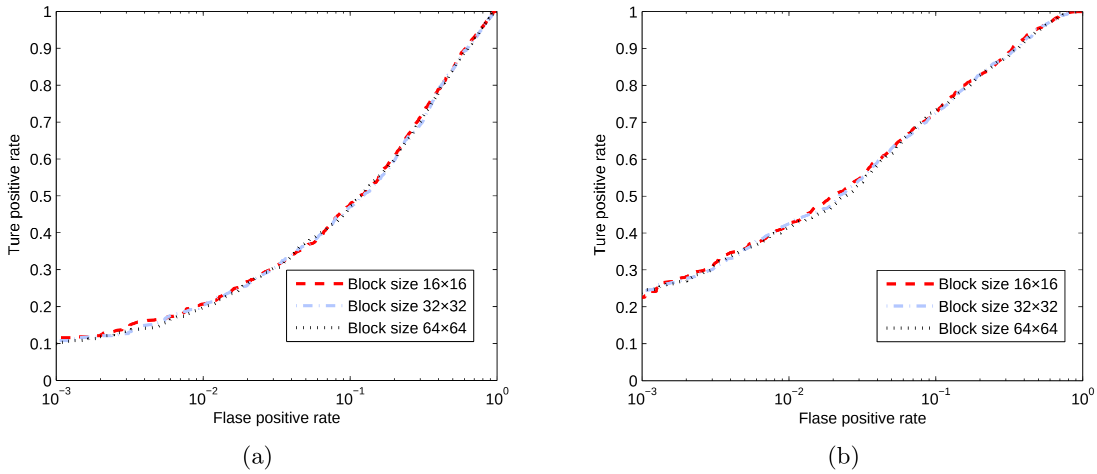


Figure 2. The ROC performance with respect to three different SPN block sizes. (query/reference image size: (a) 128×128 pixels. (b) 256×256 pixels. reference image number: 15)

2.3 Detection Statistics

Normalized Cross Correlation is applied in this work to measure the similarity between the query SPN q and the estimated reference SPN r. The model is formulated as

$$corr\,(q, r) = \frac{(q - \overline{q})\,(r - \overline{r})}{\|q - \overline{q}\|\,\|r - \overline{r}\|} \tag{6}$$

where $\overline{q}$ and $\overline{r}$ are the means of q and r, respectively.

3. EXPERIMENTS

In order to validate the capability of our method, we test our method on the Dresden Image dataset [8] and compare it with the state-of-the-art schemes [1,5], of which the reference SPN is estimated by averaging. A total of 2400 images from 16 camera devices are involved in this experiment, each device responsible for 150. These 16 camera devices belong to four camera models, so each camera model has 3∼5 different devices. Table 1 lists the cameras used in the dataset.

Table 1. Cameras used in our experiment

Cameras	Alias	Resolution	Cameras	Alias	Resolution
Canon_Ixus70_A	C11	3072×2304	Samsung_L74wide_A	C31	3072×2304
Canon_Ixus70_B	C12	3072×2304	Samsung_L74wide_B	C32	3072×2304
Canon_Ixus70_C	C13	3072×2304	Samsung_L74wide_C	C33	3072×2304
Nikon_CoolPixS710_A	C21	4352×3264	Olympus_mju_1050SW_A	C41	4352×3264
Nikon_CoolPixS710_B	C22	4352×3264	Olympus_mju_1050SW_B	C42	4352×3264
Nikon_CoolPixS710_C	C23	4352×3264	Olympus_mju_1050SW_C	C43	4352×3264
Nikon_CoolPixS710_D	C24	4352×3264	Olympus_mju_1050SW_D	C44	4352×3264
Nikon_CoolPixS710_E	C25	4352×3264	Olympus_mju_1050SW_E	C45	4352×3264

All images are the natural scene pictures in daily life which were taken under a wide variety of natural indoor and outdoor sight. It means none exactly smooth and bright images are available for estimating reference SPN. For each camera device, we randomly separate 150 images into two sub-image datasets, the reference image

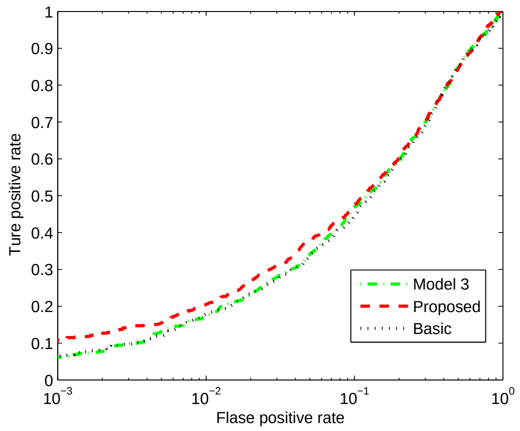


Figure 3. The ROC curves on images with size of 128×128 pixels and 15 reference images.

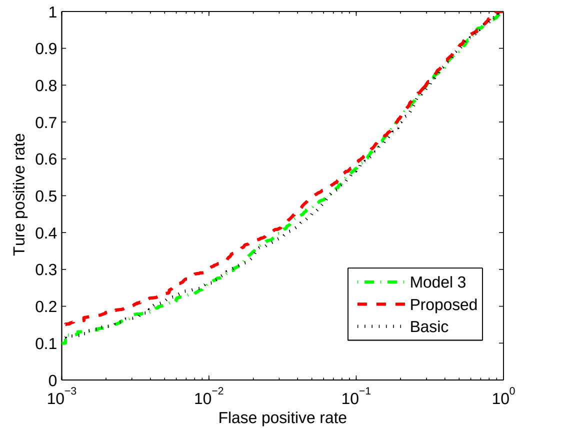


Figure 4. The ROC curves on images with size of 128×128 pixels and 30 reference images.

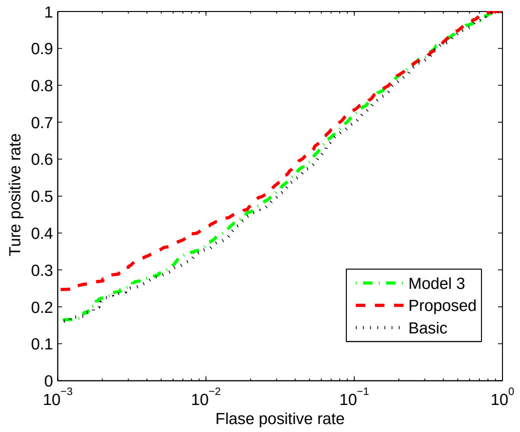


Figure 5. The ROC curves on images with size of 256×256 pixels and 15 reference images.

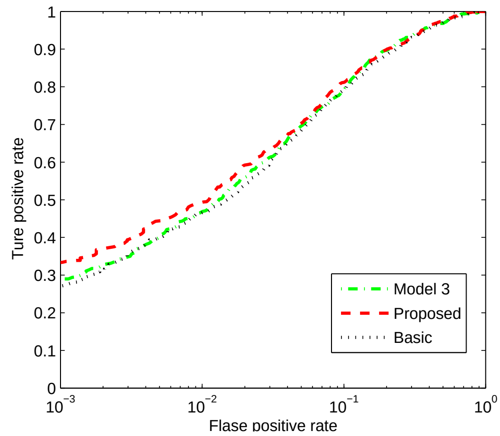


Figure 6. The ROC curves on images with size of 256×256 pixels and 30 reference images.

dataset and test image dataset, each with 30 and 120 images respectively. Images in reference image dataset are used for reference SPN estimation and images in test image dataset are for identification as test samples. Instead of using the full size photo, both reference and test images are of three different sizes (*i.e.* 128×128, 256×256 and 512×512 pixels) cropped from the center of entire image.

The corresponding experimental results in terms of the overall Receiver Operating Characteristic curve are used for presenting the performance. For each chosen camera, we first estimate reference SPN using $L = 15$ and $L = 30$ images from the reference image dataset for different methods respectively. Then, 120 test images of this camera are selected as the positive samples and 1800 test images of the other 15 cameras (each responsible for 120) are selected as the negative samples. Totally we get 120×15 positive and 1800×15 negative samples of correlation values for the overall 16 cameras. For each given detection threshold, the number of true positive decisions and the number of false positive decisions for each camera are recorded respectively, and then they were summed up to obtain the overall True Positive Rate (TPR) and overall False Positive Rate (FPR). Finally, the overall ROC curve can be created by plotting the FPR versus overall TPR in conjunction with different threshold setting.

The overall ROC performance of the proposed method compared with Basic method from [1] and Li's method [5] (*i.e.* Model 3) on different image sizes and different reference numbers are shown in Figure 3 - 8. The reference SPN of Basic method and Li's Model 3 method are estimated by averaging. In order to show the detail of the ROC curves with low FPR, the horizontal axis of the ROC curve is in logarithmic scale. The experimental results

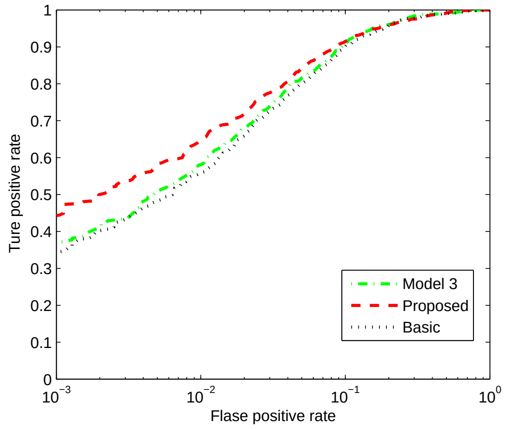


Figure 7. The ROC curves on images with size of 512×512 pixels and 15 reference images.

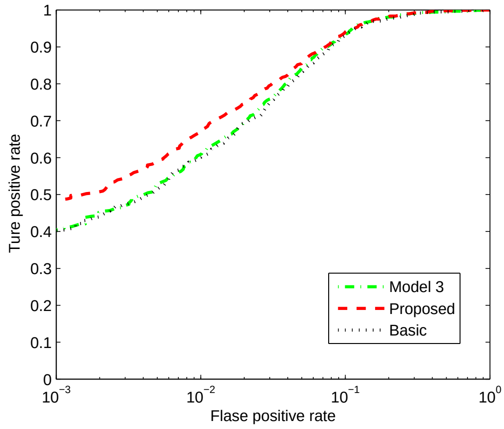


Figure 8. The ROC curves on images with size of 512×512 pixels and 30 reference images.

Table 2. The TPR of three methods at a low FPR of 10^{-3}

Method	128×128		256×256		512×512	
	15	30	15	30	15	30
Proposed	0.104	0.154	0.240	0.334	0.443	0.483
Model 3	0.063	0.109	0.173	0.290	0.369	0.405
Basic	0.065	0.113	0.168	0.275	0.342	0.401

show that, on the challenging Dresden Image Database, the proposed method outperforms the other methods [1,5], especially for a few number of reference images.

Table 2 shows the TPR of different methods at a low FPR of 10^{-3}. It shows that the TPR of the proposed method is always the largest at a low FPR. In addition, we can see that the difference between TPR of proposed method on 15 and 30 reference images is smaller than other methods. For example, on images with size of 128×128, the TPR of proposed method drops 32.5%, from 0.154 at 30 reference images to 0.104 at 15 reference images. The TPR of the Model 3 and Basic method decrease 47.8% and 42.5% respectively. It implies that the impact of losing reference images on the proposed method is not as serious as other methods.

4. CONCLUSION

In this paper, we introduced a measurement based on the smoothness and brightness to evaluate the quality of SPN blocks for the reference SPN estimation. Based on this measurement, a novel reference SPN estimator is then proposed to improve the performance of source device identification. By weighting the SPN blocks according to their quality, the proposed estimator can achieve reliable reference SPN from the limited number of natural images with scene details. Experimental results show that our method achieves better results than the state-of-the-art schemes, especially when the number of available reference images is few. These results demonstrate that our proposed technique is more practical for solving the problem of source device identification in the absence of the imaging camera than the methods based on the averaged reference SPN.

Based on the concept of this work, a further enhancement on the query SPN may be achieved by adaptively weighting the SPN blocks according to their quality, which would be the main focus of our future work.

REFERENCES

[1] J. Lukas, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," *IEEE Transactions on Information Forensics and Security,* **vol. 1**(2), pp. 205–214, 2006.

[2] M. Mihcaak, I. Kozintsev, and K. Ramchandran, "Spatially adaptive statistical modeling of wavelet image coefficients and its application to denoising," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, **vol. 6**, pp. 3253–3256, 1999.

[3] M. Chen, J. Fridrich, M. Goljan, and J. Lukas, "Determining image origin and integrity using sensor noise," *IEEE Transactions on Information Forensics and Security,* **vol. 3**(1), pp. 74–90, 2008.

[4] M. Goljan and J. Fridrich, "Digital camera identification from images - estimating false acceptance probability," in *in Proc. of the 8th Int. Workshop on Digital-forensics and Watermarking*, November 10-12, 2008.

[5] C.-T. Li, "Source camera identification using enhanced sensor pattern noise," *IEEE Transactions on Information Forensics and Security,* **vol. 5**(2), pp. 280–287, 2010.

[6] C.-T. Li and R. Satta, "Empirical investigation into the correlation between vignetting effect and the quality of sensor pattern noise," *Computer Vision, IET* **vol. 6**(6), pp. 560–566, 2012.

[7] G. Wu, X. Kang, and K. Liu, "A context adaptive predictor of sensor pattern noise for camera source identification," in *Proc. of the 19th IEEE Int. Conf. on Image Processing*, pp. 237–240, 2012.

[8] T. Gloe and R. Böhme, "The dresden image database for benchmarking digital image forensics," *Journal of Digital Forensic Practice,* **vol. 3**(2-4), pp. 150–159, 2010.