LETTER

Sensor-Pattern-Noise Map Reconstruction in Source Camera Identification for Size-Reduced Images

Joji WATANABE[†a)], *Nonmember*, **Tadaaki HOSAKA**[††], *and* **Takayuki HAMAMOTO**[†], *Members*

SUMMARY   For source camera identification, we propose a method to reconstruct the sensor pattern noise map from a size-reduced query image by minimizing an objective function derived from the observation model. Our method can be applied to multiple queries, and can thus be further improved. Experiments demonstrate the superiority of the proposed method over conventional interpolation-based magnification algorithms.
key words: *digital forensics, source camera identification, sensor pattern noise, size-reduced image, reconstruction*

1. Introduction

Interest in digital image forensics has grown sharply since the emergence of digital photography and internet photo distribution. One of the hot topics in this area is source camera identification (SCI), in which the unique device used to capture a given query image is determined [1]–[5]. Goljan et al. [1], [2] proposed a powerful method for SCI based on sensor pattern noise (SPN) which is inherent and time-invariant in each pixel. Due to SPNs, some pixels always indicate brighter/darker intensities for the constant amount of incident light. The main component of SPNs is, in many cases, photo-response non-uniformity, which is the per-pixel discrepancies in light sensitivity that result from imperfections and inhomogeneity in fabricated wafers. Although this method offers a high degree of accuracy (even when identifying one among several same-model cameras), it requires that size-reduced query images be enlarged to their original size. This is a task for which traditional interpolation-based magnification algorithms, such as the nearest-neighbor and bilinear methods, are ill-suited, as they tend to degrade identification precision. This tendency becomes more pronounced as the degree of magnification increases.

In this letter, we attempt to overcome this difficulty. Based on an image observation model, we derive the cost function that the desirable SPN map should minimize. Then, the SPN map is reconstructed from size-reduced query images by solving this optimization problem. A particular advantage of our method is that it can be naturally extended for use on a set of query images, and can thus be made more accurate. Experiments demonstrate the effectiveness of the

proposed method, even when query images are greatly reduced in size.

2. SPN-Based Source Camera Identification

In this section, we describe the method of SCI proposed in [1]. This approach also forms the fundamental framework for our method.

Goljan et al. extracted an approximate SPN component by simply subtracting the denoised image from the raw image. Thus, the SPN component $n \in \mathbb{R}^N$ ($\mathbb{R}^N$ represents the N-dimensional Euclidean space) referred to as the SPN map in this letter, can be expressed as

$$n(u) = u - g(u), \tag{1}$$

where the vector $u \in \mathbb{R}^N$ is the array for pixel values of a raw image (grayscale, in this study), and the vector $g(u) \in \mathbb{R}^N$ represents the denoised version of u that can be generated by using a simple Gaussian filter, the wavelet-based filter [6] as used in [1], and so on.

The aim of SCI is to decide whether the query image $q \in \mathbb{R}^N$ has been captured by a specific camera c. To solve this problem, the SPN map of the target camera c needs to be learnt in advance for reference. Given multiple images captured with camera c, Goljan et al. defined the reference pattern $r_c \in \mathbb{R}^N$ as the average of SPN maps derived by Eq. (1) for all learning images. The averaging operation suppresses the influence of specific scene-contents and random noises. Discrimination of a query image q is performed based on the normalized cross-correlation (NCC) between the reference pattern r_c and the SPN map $n(q)$, defined as

$$\rho(q, r_c) = \frac{n(q) \cdot r_c}{\|n(q)\| \, \|r_c\|}, \tag{2}$$

where $\| \cdot \|$ represents the l_2-norm. If the value ρ is higher than a predefined threshold, the query image q is determined to have been captured by the target camera c.

This method assumes that the reference pattern r_c and a query image q have identical dimensions. For an artificially size-reduced query image (a common situation in actual SCI usage), Goljan and Fridrich insisted [2] that the query image needs to be interpolated to match the dimensions of the reference data. However, well-known interpolation methods tend to degrade identification accuracy, particularly for greatly size-reduced query images, as will be shown later in our experiments.

3. Proposed Method

In our method, we reconstruct the SPN map from size-reduced images by solving an inverse problem derived from the observation model. Our observation model is similar to that of reconstruction-based super-resolution [7]. An artificially size-reduced (low-resolution) query image $l \in \mathbb{R}^M$ ($M < N$) is assumed to be a down-sampled version of the original (high-resolution) image $h \in \mathbb{R}^N$. This model is expressed as

$$l = Dh, \tag{3}$$

where the matrix $D \in \mathbb{R}^{M \times N}$ is a linear representation of the algorithm used to down-sample the image.

Substituting Eq. (1) into l and h in Eq. (3) yields

$$n_l + g_l(l) = D\{n_h + g_h(h)\}, \tag{4}$$

where n_h is the raw SPN map to be estimated and n_l is regarded as its low-dimensional version. Because desirable denoising filters generally depend on the resolution, the two filters are expressed by the separate indices. Here, we assume that $g_l(l)(= g_l(Dh)$, i.e., a denoised version of the down-sampled image) and $Dg_h(h)$ (i.e., a down-sampled version of the denoised image) are approximately equal. This assumption is validated for most real images by tuning the kernel settings of the denoising filters, as illustrated by Fig. 1, in which the distribution of differences (Fig. 1 (b)) between $g_l(l)$ and $Dg_h(h)$ for a given image (Fig. 1 (a)) is shown, and the difference is less than 0.2 for 94.5% of pixels. This approximation leads to

$$n_l = Dn_h + \varepsilon, \tag{5}$$

where the vector $\varepsilon \in \mathbb{R}^M$ represents the approximation error.

Our goal is to find the vector n_h that satisfies Eq. (5) as precisely as possible among natural SPN maps. To achieve this task, we define the following objective function to minimize with respect to n_h:

$$f(n_h) = \{w \cdot (n_l - Dn_h)\}^2 + \lambda \|n_h\|^2, \tag{6}$$

where the vector $w \in \mathbb{R}^M$ represents the weight for each pixel, the second term of Eq. (6) regularizes the solution, and the parameter λ balances both the terms. Because the SPN map obtained using Eq. (1) extracts not only true SPN components but also scene-dependent textures, we try to suppress the influence of textured regions by defining each element of w as

$$w_i = \exp\left(-\frac{v_i^2}{2\beta^2}\right), \quad (i = 1, 2, \ldots, M) \tag{7}$$

where v_i expresses the variance of intensities within the small region centered at pixel i in the query image l, and parameter β controls the sensitivity. From the viewpoint of Bayesian inference, the regularization term of Eq. (6) corresponds to the prior distribution defined as
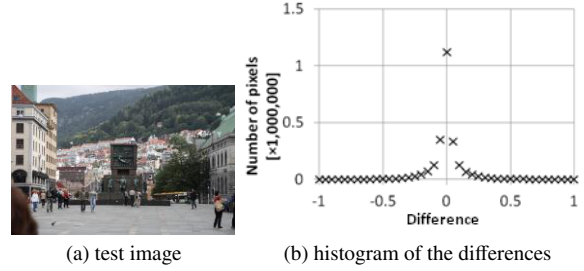

(a) test image   (b) histogram of the differences

Fig. 1 Distribution of differences between $g_l(l)$ and $Dg_h(h)$.

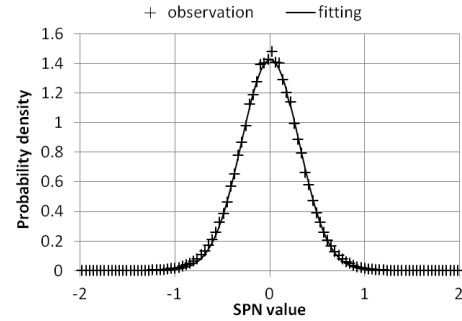**Fig. 2** Distribution of each element of a reference pattern and its Gaussian approximation.

$$p(n_h) \propto \exp\left(-\frac{\|n_h\|^2}{2\sigma^2}\right), \tag{8}$$

where the parameter σ is the variance. The validity of this regularization can be empirically confirmed for most real images, as illustrated in Fig. 2, in which the distribution of each element of a reference pattern r_c (SPN map averaged over 50 images) for a certain camera is shown, and it is well approximated by a Gaussian distribution.

We can apply this method to cases in which multiple query images are simultaneously identified. In this study, such multiple query images are assumed to be captured by the same camera, and we can simply rewrite the objective function as

$$f'(n_h) = \sum_{k=1}^{K} \{w_k \cdot (n_{lk} - D_k n_h)\}^2 + \lambda \|n_h\|^2, \tag{9}$$

where K represents the number of query images. We can easily minimize the above function with respect to n_h by using the conjugate gradient method. Once the SPN map n_h is obtained, we identify whether the query images have been captured by the target camera according to the NCC criterion given in Eq. (2).

4. Experiments

Evaluation experiments were conducted using images captured by five cameras of varying maker and model (Nikon D70, Canon EOS 400D DIGITAL, Canon EOS DIGITAL REBEL XSi, Pentax K20D, and Leica M9 Digital Camera) which are published in the BOSS Web page [8]. This database includes images for a variety of subjects such as

Fig. 3 Examples of images used in the experiments.

Fig. 4 Average EER with respect to the reduction ratio when using a single query image.

Fig. 5 Average EER with respect to the number of query images.

scenery, animals, and buildings, as shown in Fig. 3. To suppress the influence of the number of pixels, all images were cropped to a size of 3000×2000 pixels.

The reference SPN maps were computed using 50 images for each camera. In this process, we utilized the usual Gaussian filter as a denoising filter $g(\cdot)$.

The down-sampling procedure to generate size-reduced query images was as follows:

(i) We consider a 2-dimensional space on which pixels of the original (high-resolution) image correspond to grid points.

(ii) We set the sampling points on this space in accordance with the reduction ratios. Note that these sampling points are not necessarily on the grid points.

(iii) The intensity at each sampling point is interpolated by the weighted average over the surrounding pixels. The weights are calculated by using the Gaussian kernel whose center is located at each sampling point.

(iv) Extracting these interpolated intensities at the sampling points gives us the size-reduced image.

Assuming that the matrix D characterizing the above reduction process is known, the reconstruction was performed with parameters $\lambda = 0.01$, $\beta^2 = 3$. We evaluated the performance of our proposed method using equal error rate (EER), which is the value at which the false acceptance rate and the false rejection rate are made equal by varying the identification threshold.

4.1 Experiments with a Single Query Image

For each camera, 100 query images were prepared. The average EER with respect to reduction ratios $\sqrt{M/N}$ (={0.25, 0.29, 0.33, 0.4, 0.5}) is shown in Fig. 4 along with the results when nearest-neighbor and bilinear algorithms were used for magnifying a query image. Due to properties suitable for reconstructing the SPN map (i.e., suppression of scene-dependent textures by the weight w and of over-fitting by the regularization), the identification accuracy of the proposed method represents an improvement over conventional methods at every reduction ratio. However, when the reduction ratio falls below 0.3, the accuracy deteriorates remark-
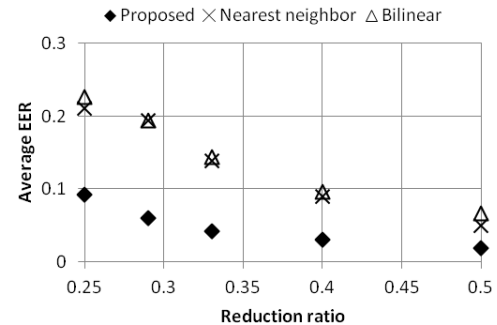
ably because the number of sampling points is insufficient to reconstruct the desirable SPN map. We hope that the identification accuracy at these lower reduction ratios can be improved by using multiple query images.
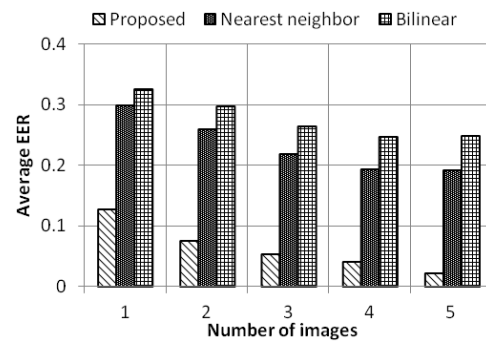
4.2 Experiments with Multiple Query Images

For each camera, 100 greatly size-reduced images, whose reduction ratios are in the range of [0.16, 0.25], were collected. These images were divided into 20 sets of five images. For a given $K(\leq 5)$, we used every combination of images within each set as queries, yielding a number of trials totaling 20_5C_K.

Figure 5 shows the average EER with respect to the number of query images identified at one time. For the two kinds of conventional methods, SPN maps were separately derived for each query image and averaged. We see that by increasing the number of query images, we can achieve more drastic improvements over conventional methods. This implies that an increase in sampling points helps to reconstruct a better SPN map. Examples of the SPN maps reconstructed with one, three, and five query images, are shown in Fig. 6 where SPN values are shifted and amplified so that they are expressed in the range of [0, 255]. It can be confirmed that the resolutions of SPN maps are enhanced and more pixels of high/low light-sensitivity are reconstructed as the number of query images increases.
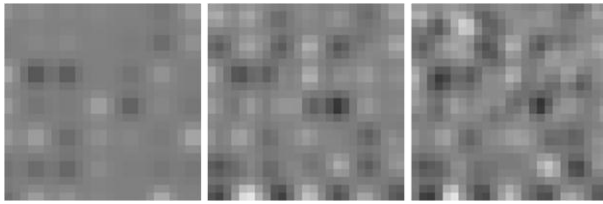
(a) one query image   (b) three query images   (c) five query images

Fig. 6 Reconstructed SPN maps with respect to the number of query images: A part of the SPN map is clipped in size of 25×25.

4.3 Discussion on Estimation of the Matrix D

In real situations, it is necessary to estimate the matrix D from the given query images. When the matrix D is not precise, the quality of reconstruction degrades as is the case with the super-resolution, and consequently, the value of NCC decreases. Given that the down-sampling process is modeled with some parameters (in our case, the variance of the Gaussian kernel), it is possible to estimate the matrix D by exploring the model parameters that maximize the value of NCC. This task can be performed by using a variety of optimization algorithms.

5. Conclusions and Future Works

We proposed a method to reconstruct the SPN map from size-reduced images for SCI. Our reconstruction method is based on the image observation model and can be applied to the case of multiple query images. In the evaluation experiments, our proposed method showed higher identification accuracy than conventional interpolation-based magnification algorithms, and was further improved by using multiple query images.

In future work, we will address a variety of common image-processing transformations, including trimming, rotation, and paint insertion. Another direction is to investigate the robustness of the proposed method to nonlinear processing such as JPEG compression.

References

[1] J. Lukas, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," IEEE Trans. Information Forensics and Security, vol.1, no.2, pp.205–214, 2006.

[2] M. Goljan and J. Fridrich, "Camera identification from cropped and scaled image," Proc. SPIE Electronic Imaging, Forensics, Security, Steganography, and Watermarking of Multimedia Contents X, vol.6819, pp.0E-1–0E-13, 2008.

[3] E.J. Alles, Z.J. Geradts, and C.J. Veenman, "Source camera identification for low resolution heavily compressed images," Proc. Int. Conf. on Computational Sciences and Its Applications, pp.557–567, 2008.

[4] Y. Sutcu, S. Bayram, H.T. Sencar, and N. Memon, "Improvements on sensor noise based source camera identification," Proc. IEEE Int. Conf. Multimedia and Expo, pp.24–27, 2007.

[5] Y. Long and Y. Huang, "Image based source camera identification using demosaicking," Proc. IEEE 8th Workshop Multimedia Signal Processing, pp.419–424, 2006.

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

[7] S.C. Park, M.K. Park, and M.G. Kang, "Super-resolution image reconstruction: A technical overview," IEEE Signal Process. Mag., vol.20, no.3, pp.21–36, 2003.

[8] http://exile.felk.cvut.cz/boss/BOSSFinal/