Identifing source camera using guided image estimation and block weighted average

Le-Bing Zhang, Fei Peng, Min Long

PII: S1047-3203(16)30268-1
DOI: http://dx.doi.org/10.1016/j.jvcir.2016.12.013
Reference: YJVCI 1917

To appear in: *J. Vis. Commun. Image R.*

Received Date: 9 September 2016
Revised Date: 16 December 2016
Accepted Date: 21 December 2016

Please cite this article as: L-B. Zhang, F. Peng, M. Long, Identifing source camera using guided image estimation and block weighted average, *J. Vis. Commun. Image R.* (2016), doi: http://dx.doi.org/10.1016/j.jvcir.2016.12.013

Identifing source camera using guided image estimation and block weighted average

Le-Bing Zhang[a], Fei Peng[a,*], Min Long[b]

[a]School of Computer Science and Electronic Engineering, Hunan University, 410082 Changsha, China
[b]College of Computer and Communication Engineering, Changsha University of Science and Technology, 410112 Changsha, China

Abstract

Sensor pattern noise (SPN) has been widely used in source camera identification. However, the SPN extracted from natural image may be contaminated by its content and eventually introduce side effect to the identification accuracy. In this paper, an effective source camera identification scheme based on guided image estimation and block weighted average is proposed. Before the SPN extraction, an adaptive SPN estimator based on image content is implemented to reduce the influence of image scene and improve the quality of the SPN. Furthermore, a novel camera reference SPN construction method is put forward by using some ordinary images, instead of the blue sky images in the previous schemes, and a block weighted average approach is used to suppress the influence of the image scenes in the reference SPN. Experimental results and analysis indicate that the proposed method can effectively identify the source of the natural image, especially in actual forensics environment with a small number of images.

Keywords: Source camera identification, Guided image filtering, Block weighted average, Sensor pattern noise

1. Introduction

With the rapid development of digital imaging devices, portable video camera, digital camera and mobile phone have been widely used in daily life. Digital image acquisition, publishing and sharing are becoming popular information transmission and exchange means in modern social network. Meanwhile, some powerful and easy-to-use digital image processing software provide simple tools to retouch digital images. As it brings convenience to us, the security concerns of the digital image have been attracted wide attention in the past decade, especially in the field of judicial and criminal investigation. Therefore, as a kind of multimedia security technology, digital image forensics can be used to verify the originality, authenticity and reliability of digital images.

Image source identification is an important branch of digital image forensics. It can accurately and reliably match the specific digital image with its source. Typical image source identification contains two categories. One is image acquisition pipeline identification. The target is to distinguish images taken from digital camera, images created by computer graphics or images produced by scanners; the other is the identification of camera brands or types, and it aims to identify which brand or individual camera is used for a given image. In this paper, the work will be focused on the second type, which is also called as source camera identification.

Natural image is the reflection of the light from the natural scene obtained by the camera. The acquisition of an image with a digital camera is illustrated in Fig. 1. The light reflected from the natural scene first passes through a camera lens. The function of the lens is to filter out other colors of light, and only permit red, green and blue to pass through it. CCD sensor will transform light intensity into voltage/current. For cost consideration, only one color is recorded at each position of the CCD with a color filter array (CFA) pattern. After A/D transform, the rest two colors are obtained by a CFA interpolation algorithm. After that, a sequence of image post-processing operations, such as white balancing, Gamma correction, contrast enhancement and JPEG compression, are performed to generate a digital

*Corresponding author
Email addresses: zhanglebing@hnu.edu.cn (Le-Bing Zhang), eepengf@gmail.com (Fei Peng), caslongm@gmail.com (Min Long)

image. Different brands or types of digital cameras use different means to process image information, so different digital camera has its own unique properties generated in the imaging process, which is the key to the identification of image devices. These features include lens aberration [1, 2], CFA interpolation artifacts [3], sensor pattern noise (SPN) [4–13], statistical characteristics [14], JPEG compression [15, 16] etc. In addition, image quality evaluation index and high order statistics in wavelet domain [17, 18] have also been exploited to identify source camera.

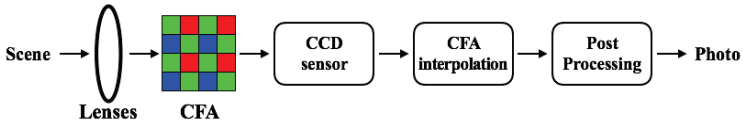


Fig. 1**.** The imaging process of a digital camera.

Among them, SPN generated by digital camera in the imaging process has become a research hotspot in digital image forensics, and it is widely used in image source camera identification and image tamper forensics [19, 20]. Generally, SPN is composed of two main components: fixed pattern noise (FPN) and photo-response non-uniformity noise (PRNU). PRNU is resulted from the hardware defects and the inhomogeneity of silicon wafers during the CMOS/CCD sensor manufacturing process. Therefore, different camera has different PRNU, and it is not affected by temperature or humidity, just like the unique fingerprint of an individual imaging sensor. Therefore, SPN can be considered to be a unique and stable trace of each camera, and it is well suited for camera source identification. In this paper, extraction of stable and accurate SPN is investigated, and a novel source camera identification scheme based on the guided image estimator and block weighted average is proposed to improve the identification performance.

The main contributions of this paper include:

(1) An effective guided image estimator is used for the extraction of SPN, and it can reduce the influence of image scene and improve the quality of the extracted SPN.

(2) A block weighted average approach is proposed to construct camera reference SPN from arbitrary natural images.

(3) Based on the above SPN extraction and camera reference SPN construction method, an effective source camera identification scheme is proposed, and the experimental results verify the effectiveness of it.

The rest of the paper is organized as follows: the related works are introduced in the Section 2. The proposed source camera identification scheme is described in Section 3. Experimental results and analysis are provided in Section 4. Finally, some conclusions are drawn in Section 5.

2. Related works

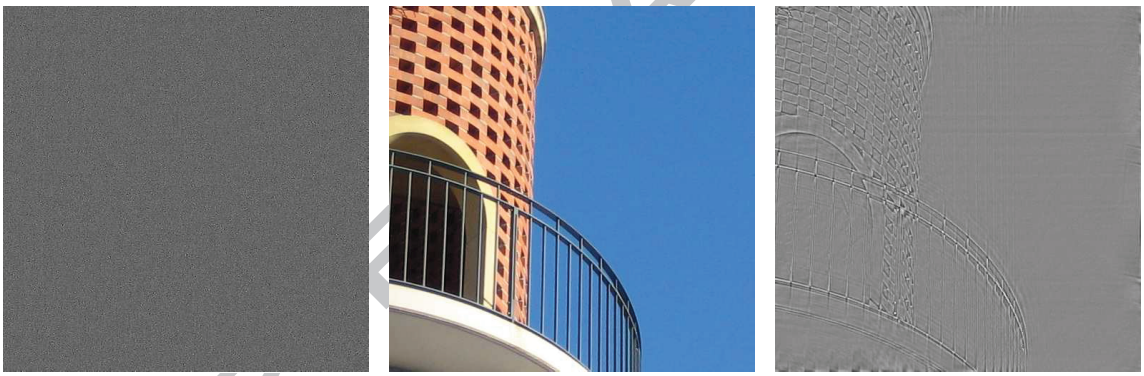
Currently, the general idea of source camera identification based on SPN is: a camera reference SPN is first constructed, and then the SPN of a testing image is extracted. After that, compare two SPNs and the final decision will be made for whether the testing image is obtained from this camera or not.

Lukas et al. [4] first proposed a method to utilize SPN for source camera identification. Since SPN is a kind of high frequency signal, a wavelet based denoising filter is used to obtain the SPN of the testing image. Meanwhile, camera reference SPN is constructed by averaging multiple SPNs extracted from some images whose camera model and brand are known. Based on the correlation between the testing image's SPN and the camera reference SPN, the image source camera can be determined. It provides an effective method for the source camera identification. After that, the researches of camera source forensics based on SPN are mainly focused on two aspects. The first attempts to enhance the quality of the extracted SPN, and the second is to improve the correlation methods. Chen et al. [6] proposed to calculate camera reference SPN by using maximum likelihood estimation, and two preprocessing operations including zero-mean and Wiener filter are implemented to further remove the artifacts of the imaging processing. However, the influence of the image scene is still strong. Li [7] assumed that strong signal in SPN is more likely affected by the content of the image, and 6 enhancement models are proposed to suppress the high frequency signal of SPN for reducing the influence of image scene. Recently, Li [9] proposed a spectrum equalization method to reduce the influence of the non-unique artifacts in the reference SPN. It detects and suppresses peaks to equalize the amplitude spectrum of the reference SPN based on the local features, and periodic interference is removed. It

is similar to the previous SPN enhancement method in [7]. Although it can suppress the image content with high frequency in SPN, it may lose useful high frequency information of SPN. Chierchia et al. [10] proposed a block matched 3D filter (BM3D) [21] to replace the traditional wavelet-based denosing filtering method. Kang et al. [11] proposed a context-adaptive interpolation operator, and Zeng et al. [12] proposed a content adaptive guided image filter for suppressing edge effect of image scene content. Satta [13] proposed a SPN reliability map and then used to weight SPN pixels during matching. All these methods can improve the quality of SPN to a certain degree. However, it is still very difficult to totally remove the image scene in SPN.

For the correlation calculation between the testing image's SPN and the camera reference SPN, Lukas et al. [4] proposed normalized cross-correlation (NCC) to measure the similarity. Goljan et al. [22] proposed the peak to correlation energy (PCE) measure to reduce the influence of periodic noise. Kang et al. [23] proposed to use the correlation over circular cross-correlation norm (CCN) to reduce the false positive rate of detection. Different correlation calculation methods can be combined with different SPN extraction methods to improve the overall detection accuracy. In addition, fast, accurate and effective camera source forensics method based on SPN [24, 25] for large scale dataset is becoming a recent development trend.

Since the SPN is only related to the camera imaging devices, it should be irrelevant to the content of the scene. So the SPN contaminated by the scene cannot accurately reflect the essential characteristics of the camera, which will eventually influence the accuracy of the camera source identification. However, the SPN extracted in the existed camera source identification methods is inevitably affected by the content of the image scene. An example of SPN and reference SPN are shown in Fig. 2.



(a) (b) (c)

Fig. 2. An example of SPN and reference SPN. (a) Reference SPN taken from flat-field images. (b) Image of a natural scene. (c) SPN extracted from (b).

As seen from Fig. 2 (c), it can be found that the extracted SPN contains the scene of the natural image, while the ideal SPN should look like Fig. 2 (a). Therefore, how to effectively eliminate the interference of the image scene on the SPN is important for the source camera identification.

In the previous literatures [6, 7, 11], in order to accurately obtain camera reference SPN, some blue sky images with no scene are generally used for the estimation of the reference SPN. However, in most cases, camera reference SPN can be only extracted from a few ordinary images, which contain different kinds of scene details. Therefore, how to effectively construct the camera reference SPN to identify the source of digital images in actual forensics environment is becoming an urgent problem to be resolved.

To countermeasure the above problems, an effective source camera identification scheme based on guided image estimation and block weighted average is proposed.

3. Description of the proposed scheme

Considering the strong dependence between the extracted SPN and the content of the image scene in the traditional methods, a content based adaptive estimator is investigated. By using the difference between the estimated image and

the original image, the influence of image scene on the SPN is reduced and the SPN information left by the camera during the imaging process is effectively preserved. Here, guided image filter is implemented to extract the SPN information of the image.

3.1. Guided image filter

Guided image filter [26] is a general linear translation invariant filter and its output is a local weighted linear combination of the input image according to a guidance image. Assuming that I is an input image, G is a guidance image and q is an output image. The filtering can be described as:

$$q_i = \sum_j W_{ij}(G) I_j \,, \tag{1}$$

$$W_{ij}(G) = \frac{1}{|\omega|^2} \sum_{k:(i,j)\in\omega_k} (1 + \frac{(G_i - \mu_k)(G_j - \mu_k)}{\sigma_k^2 + \varepsilon}) \,, \tag{2}$$

where $W_{ij}(\cdot)$ is the kernel function, μ_k and σ_k are the mean and variance of G in ω_k, and $|\omega|$ is the number of pixels in ω_k .

The core idea of the guided image filter is that q is obtained from G with a local linear transformation. Equation (1) can be rewritten as:

$$q_i = a_k G_i + b_k \,, \qquad \forall i \in \omega_k \,, \tag{3}$$

where ω_k represents a square window with a radius r, and its center is located at k, a_k and b_k are linear coefficients.

To determine the linear coefficients a_k and b_k, the input image I and the output image q should fulfill

$$q_i = I_i - n_i \,, \tag{4}$$

and minimize the error of the linear transformation. Then a_k, b_k can be written as

$$(a_k,\ b_k) = argmin \sum_{i\in\omega_k} ((a_k G_i + b_k - I_i)^2 + \varepsilon a_k^2) \,, \tag{5}$$

where ε is a regularization parameter penalizing large a_k.

Linear ridge regression model is implemented for Equation (5) to resolve the optimal solution of a_k and b_k. Then

$$a_k = \frac{1}{|\omega|} \sum_{i\in\omega_k} (\frac{G_i I_i - \mu_k \overline{I_k})}{\sigma_k^2 + \varepsilon}) \,, \tag{6}$$

$$b_k = \overline{I_k} - a_k \mu_k \,, \tag{7}$$

$$\overline{I_k} = \frac{1}{|\omega|} \sum_{\mathrm{i}\in\omega_k} I_i \,. \tag{8}$$

Substitute a_k and b_k into the Equation (3), and it can be rewritten as:

$$q_i = \frac{1}{|\omega|} \sum_{k|\mathrm{i}\in\omega_k} (a_k G_i + b_k) = \overline{a_i} G_i + \overline{b_i} \,, \tag{9}$$

where $G = I$, and $\overline{b_i} = \frac{1}{|\omega|} \sum_{k\in\omega_i} b_k$.

When $I = G$, it means that the input image is the same as the guidance image. In this case, Equation (6),(7) can be represented as $a_k = \sigma_k^2/(\sigma_k^2 + \varepsilon)$ and $b_k = (1 - a_k)\mu_k$. Obviously, if $\varepsilon = 0$, then $a_k = 1$, $b_k = 0$. If $\varepsilon > 0$, there exist σ_k, ε so that $a_k \approx 1$ and $b_k \approx 0$ hold in the image regions with complex texture; While in the smooth region of the image, there exist σ_k, ε so that $a_k \approx 0$ and $b_k \approx {}_k$ hold. It reflects that the guided image filter exhibits the edge-preserving properties, which can be used as an image content adaptive estimator to estimate the content of the image scene.

3.2. SPN extraction based on guided image estimation

By employing the property of edge-preserving properties, the SPN can be extracted according to the following steps:

***Step* 1.** To suppress the contents of the image scene, the image I is used as input image and guidance image of the guided image filter. Subsequently, the estimated image $G(I, I)$ is obtained, and the difference R is:

$$R = I - G(I, I) \ , \tag{10}$$

where $G(\cdot, \cdot)$ represents the guided image filtering function, and R represents the residual image.

***Step* 2.** For the residual image R, a pixel-wise adaptive Wiener filtering with a window size of 3×3 is performed to it, the SPN of the residual image can be calculated from

$$W(i, j) = R(i, j)\frac{\sigma_0^2}{\hat{\sigma}^2(i, j) + \sigma_0^2} \ , \tag{11}$$

where $\hat{\sigma}^2$ represents the local variance of the original noise-free image, and σ_0^2 represents the overall variance of the SPN. Here, σ_0^2 is set as 9, and $\hat{\sigma}^2$ is obtained by using the maximum a posteriori probability estimation, according to [11]. It is described as:

$$\hat{\sigma}^2(i, j) = \max(0, \frac{1}{m^2} \sum_{(p,q)\in N_m} R^2(p, q) - \sigma_0^2) \ , \tag{12}$$

where m is the size of a neighbourhood N_m for each pixel, here $m = 3$.

3.3. Construction of the reference SPN based on block weighted average

Currently, there are two conventional methods for constructing a camera reference SPN. One is Lukas's method [4], which constructs the camera reference SPN through averaging the SPNs of multiple images of the camera. The camera reference SPN is calculated from

$$S = \frac{1}{N} \sum_{k=1}^{N} R_k \ , \tag{13}$$

where S is the camera reference SPN, and R_k is a SPN extracted from the k^{th} image.

Another is Chen's method [6], where Maximum Likelihood Estimator is implemented for the construction of SPN. It can be defined as:

$$S = \frac{\sum_{k=1}^{N} R_k I_k}{\sum_{k=1}^{N} (I_k)^2} \ . \tag{14}$$

For the above methods, the reference SPN is usually recommended to be extracted from more than 50 blue sky images. Therefore, the extracted SPN is similar to the ideal SPN. However, considering the actual forensics situations, camera reference SPN can only be extracted from a few ordinary images, which contain different kinds of scene details. Thus, the SPN is inevitable disturbed by the content of the scene. Due to the difference of R_k, the conventional methods cannot achieve the minimum mean squared error. For a natural image, it usually contains smooth regions and textural regions. Generally, smooth region is good for the extraction of SPN, while the textural region interfere the calculation of SPN. Meanwhile, the variance of the textural region is usually larger than that of the smooth region. Therefore, by assigning large weight to the smooth region and assigning small weight to the textural region, a novel estimation approach for camera reference SPN is proposed based on block weighted average. It is illustrated in Fig. 3.

Assuming N ordinary images are provided for the construction of camera reference SPN, the details are described as follows:

***Step* 1.** For each image I_k ($k = 1, 2, ..., N$), obtain its SPN W_k according to Section 3.2. After that, the SPN W_k is divided into $P \times Q$ blocks, and then calculate the variance $\sigma_k^2(p, q)$ of each block $BW_k(p, q)$ ($k = 1, 2, ..., N$,

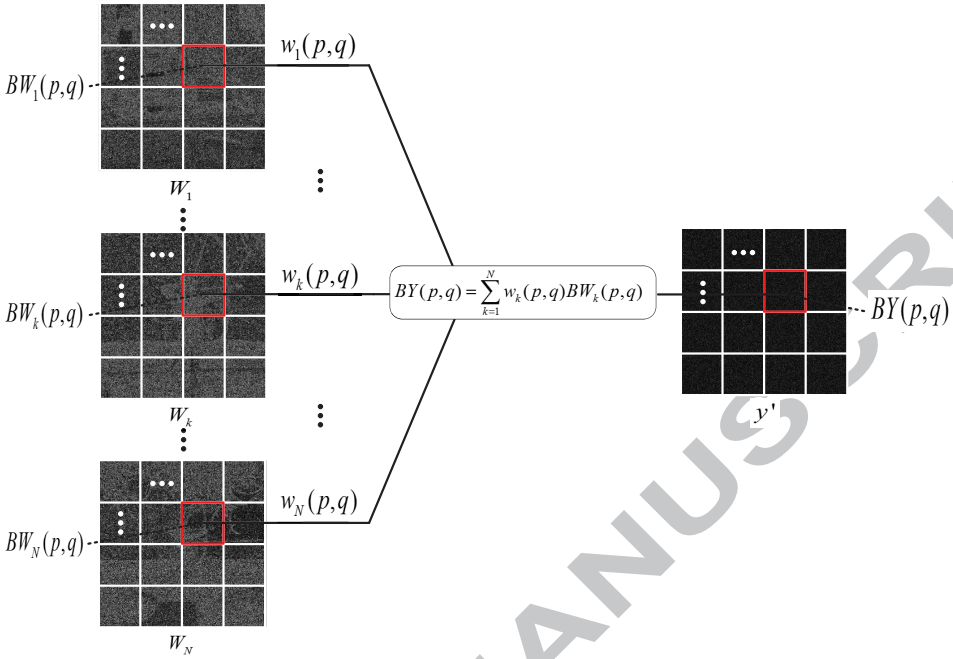


Fig. 3. The diagram for constructing camera reference SPN.

$p = 1, 2, ..., P, q = 1, 2, ..., Q$), where $BW_k(p,q)$ represents the k^{th} SPN's block located at the p^{th} row and the q^{th} column. Here, $P = Q = 2$ is used in the following experiments, which will be analyzed in Section 4.2.1.

***Step* 2.** The local reference SPN $BY(p,q)$ is constructed by multiplying $BW_k(p,q)$ with a corresponding weight, which is depended on $\sigma_k^2(p,q)$. It is represented as:

$$BY(p,q) = \sum_{k=1}^{N} w_k(p,q)BW_k(p,q) \ . \tag{15}$$

The optimal weight w_k is

$$w_k(p,q) = \frac{\frac{1}{\sigma_k^2(p,q)}}{\sum_{i=1}^{N} \frac{1}{\sigma_i^2(p,q)}}, k = 1, 2, ..., N. \tag{16}$$

***Step* 3.** The whole reference SPN y' is acquired by combining these local reference SPN $BY(p,q)$ together.

***Step* 4.** To further suppress the unwanted artifacts caused by camera processing operations such as CFA interpolation, a pre-procession similar to [6] is adopted. Therefore, the final camera reference SPN y is

$$y = WF(ZM(y')) \ , \tag{17}$$

where $ZM(\cdot)$ represents zero-mean operation in every row and column, and $WF(\cdot)$ represents Wiener filtering in the Fourier domain.

4. Experimental results and analysis

In the following experiments, excluding some particular cameras, which contain unexpected artifacts [27], 1350 natural images and 150 flat-field images from 6 cameras are chosen from Dresden Image Database [28] to form the image dataset. The camera model, image format and original resolution are listed in Table 1.

Table 1. Detail of the dataset used in experiment

Brand	Resolution	Number of images
Canon Ixus55	$3,264 \times 2,448$	250
Nikon D200	$3,872 \times 2,592$	250
Olympus Mju 1050SW	$3,648 \times 2,736$	250
Panasonic DMC FZ50	$3,648 \times 2,736$	250
Samsung L74 wide	$3,072 \times 2,304$	250
Sony DSC H50	$3,456 \times 2,592$	250

For simplicity, only the center region of the image is considered. Thus, the testing center sizes include 128×128, 256×256, 512×512 and 1024×1024. Based on this dataset, the performance of the proposed method is compared with the Basic method [4], MLE method [6], Li-Model5 method [7], PCAI8 method [11], and CAGIF method [12].

PCE correlation has proved to be more stable than NCC when images are undergone geometrical transformations such as rotation or resizing. However, the purpose of this work is to verify the effectiveness of eliminating the interference of image scene content on SPN, so the geometric transformation of testing image is not considered. Therefore, NCC is used to measure the similarity between the camera reference SPN and noise residue for each method. It is defined as:

$$\rho(x, y) = corr(x, y) = \frac{(x - \overline{x}) \cdot (y - \overline{y})}{\|x - \overline{x}\| \cdot \|y - \overline{y}\|} , \tag{18}$$

where $\|\cdot\|$ represents L_2 norm, and $\bar{x}$, $\bar{y}$ represent the means of the SPN extracted from the testing image x and the camera reference SPN y with a same size $M \times N$, respectively.

To further evaluate the experimental results, ROC curves is implemented for the performance analysis. For a given detection threshold, true positive (TP) and false positive (FP) of each camera are calculated, and then the total true positive and false positive can be achieved. True positive rate (TPR), false positive rate (FPR) and the accuracy are defined as:

$$TPR = \frac{TP}{TP + FN} , \tag{19}$$

$$FPR = \frac{FP}{FP + TN} , \tag{20}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} . \tag{21}$$

4.1. Experimental results

The experiments are conducted with two situations. In the first situation, the camera reference SPN can be only extracted from a few ordinary images, which contain different kinds of scene details. The second case is that the camera reference SPN is extracted from some flat-field images in the image database.

4.1.1. Experiments with reference SPN constructed from ordinary images

In this part, the camera reference SPN is extracted from 25 ordinary images. The ordinary images contain a wide variety of indoor and outdoor scenes. For each camera, 25 ordinary images are randomly chosen from 225 ordinary images to estimate reference SPN, and the other 200 images are used as testing images. The identification accuracies of different methods with different image sizes are listed in Table 2.

From table 2, it can be seen that the average accuracy of the proposed method outperforms the other methods in different image sizes, which indicates the good identification performance of the proposed method.

The overall ROC curves of CAGIF [12], PCAI8 [11], Li-Model5 [7], MLE [6], Basic [4] and the proposed method are shown in Fig. 4.

Table 2. Accuracy of different methods

Image size	Methods	Canon Ixus55	Nikon D200	Olympus mju1050W	Panasonic DMC-FZ50	Samsung L74wide	Sony DSC-H50	Average accuracy
128 × 128	Basic[4]	90.25%	87.25%	86.25%	89.92%	86.58%	95.17%	89.24%
	MLE[6]	89.42%	86.17%	86.58%	89.42%	85.92%	96.58%	89.01%
	Li-Model5[7]	**92.75%**	**93.42%**	**89.17%**	83.33%	**90.42%**	94.08%	90.53%
	PCAI8[11]	92.17%	89.67%	87.33%	**93.25%**	87.92%	97.17%	91.25%
	CAGIF[12]	91.08%	87.83%	85.50%	89.00%	86.08%	96.00%	89.25%
	Proposed Method	92.17%	90.17%	88.00%	93.17%	88.75%	**97.33%**	**91.60%**
256 × 256	Basic[4]	95.50%	91.83%	90.00%	95.75%	91.75%	99.00%	93.97%
	MLE[6]	96.42%	90.75%	90.00%	95.75%	90.17%	99.08%	93.69%
	Li-Model5[7]	**98.25%**	**95.58%**	91.08%	83.33%	**97.00%**	96.58%	93.64%
	PCAI8[11]	97.75%	94.67%	92.17%	97.33%	93.67%	98.92%	95.75%
	CAGIF[12]	96.25%	90.50%	90.92%	94.08%	91.33%	98.83%	93.65%
	Proposed Method	98.00%	95.50%	**93.08%**	**97.50%**	93.58%	**99.17%**	**96.14%**
512 × 512	Basic[4]	**99.92%**	96.08%	93.92%	98.00%	98.00%	99.50%	97.57%
	MLE[6]	99.83%	95.42%	95.92%	98.00%	97.58%	**100%**	97.79%
	Li-Model5[7]	99.83%	96.42%	92.08%	83.33%	**99.92%**	97.17%	94.79%
	PCAI8[11]	99.83%	98.00%	96.08%	98.25%	99.17%	**100%**	98.56%
	CAGIF[12]	99.58%	96.67%	95.92%	97.67%	97.42%	99.92%	97.86%
	Proposed Method	99.83%	**98.67%**	**96.33%**	**98.33%**	98.83%	**100%**	**98.67%**
1024 × 1024	Basic[4]	99.92%	97.92%	94.67%	**98.75%**	99.83%	99.83%	98.49%
	MLE[6]	**100%**	98.17%	97.08%	98.25%	99.92%	**100%**	98.90%
	Li-Model5[7]	**100%**	97.00%	94.17%	85.67%	**100%**	97.83%	95.78%
	PCAI8[11]	**100%**	99.50%	**97.67%**	98.50%	99.92%	**100%**	99.26%
	CAGIF[12]	99.92%	99.42%	97.33%	**98.75%**	99.92%	**100%**	99.22%
	Proposed Method	**100%**	**99.67%**	97.50%	98.58%	**100%**	**100%**	**99.29%**

As seen from Fig. 4, the overall ROC curves indicate that the proposed method can achieve good performance when different sizes of the center images are used. Furthermore, in camera source forensics application, it is important to minimize the probability of false acceptance rate. Thus, the ROC with small FPR (10^{-3}) is more desirable.

Table 3 shows the TPRs of different methods when FPR = 10^{-3}. It indicates that the TPR of the proposed method outperform those of the methods in [4, 6, 7, 11, 12].

Table 3. The TPR of the different methods when FPR = 10^{-3} (%)

Method	Image size			
	128 × 128	256 × 256	512 × 512	1024 × 1024
Basic[4]	36.42	75.00	92.75	97.58
MLE[6]	40.33	73.25	95.08	98.50
Li-Model5[7]	25.75	55.25	71.92	88.08
PCAI8[11]	48.58	83.25	**98.00**	98.75
CAGIF[12]	40.33	75.58	95.08	98.75
Proposed Method	**50.50**	**85.25**	97.42	**98.83**

4.1.2. Experiments with reference SPN constructed from flat-field images

To further verify capability of the proposed method in suppressing the scene details, the camera reference SPN is extracted from 25 flat-field images similar to previous literatures [6, 7, 11]. Table 4 shows the TPR of different methods when FPR = 10^{-3}. The overall ROC curves when the size of center image is 256 × 256 are shown in Fig. 5.

As seen from Fig. 5, the AUC (Area Under Curve) of the proposed method is the largest among all the methods, which indicates the good performance of it.

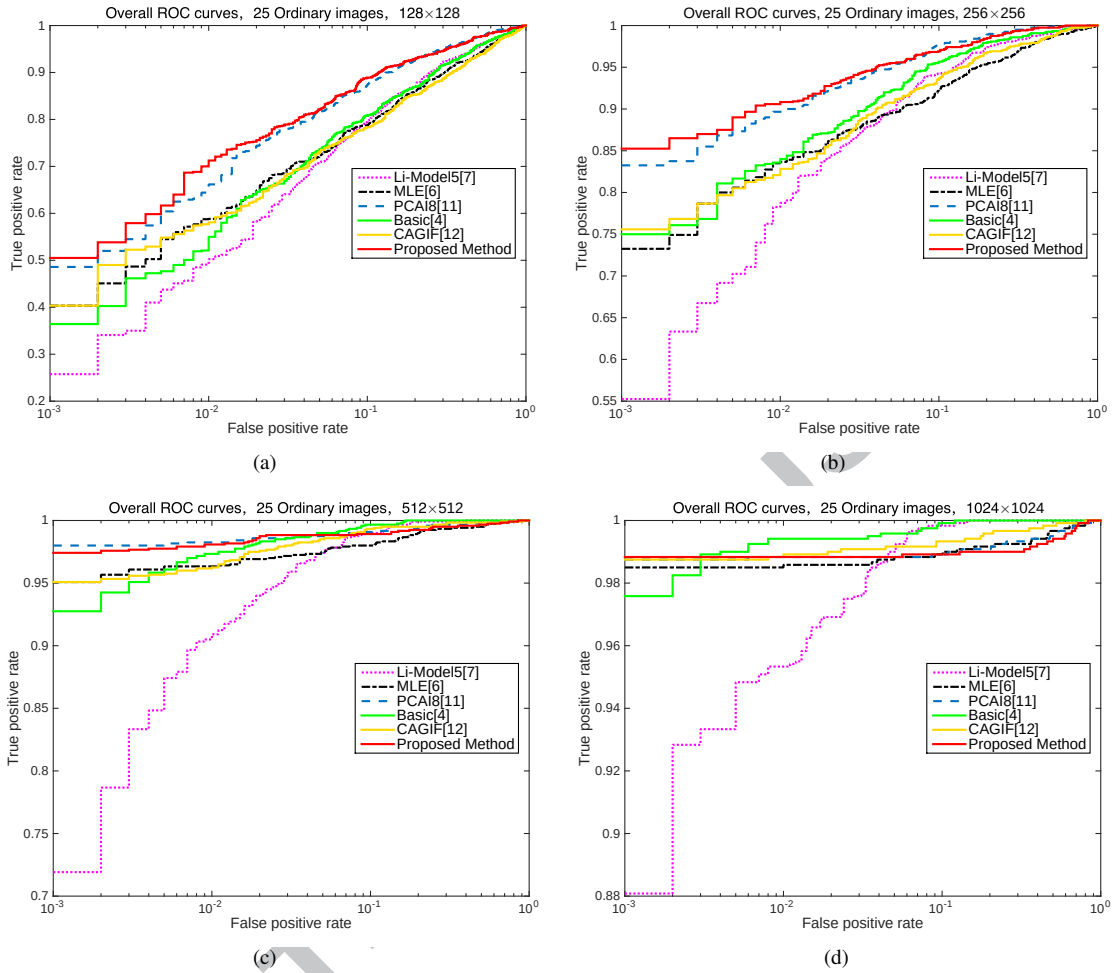


Fig. 4. The overall ROC curves of different methods. (a) Image size: 128×128. (b) Image size: 256×256. (c) Image size: 512×512. (d) Image size: 1024×1024.

Table 4. The TPR of the different methods when FPR = 10^{-3} (%)

Method	**Image size**			
	128×128	256×256	512×512	1024×1024
Basic[4]	48.33	72.58	80.42	83.50
MLE[6]	47.17	80.58	95.67	98.58
Li-Model5[7]	20.67	59.67	73.50	77.08
PCAI8[11]	54.17	81.50	94.50	98.58
CAGIF[12]	53.00	84.17	96.08	**98.75**
Proposed Method	**55.25**	**86.50**	**96.33**	98.67

The above experimental results both indicate that the proposed method has better performance in camera source identification than those of the existed methods in [4, 6, 7, 11, 12], whether the camera reference SPN is extracted from ordinary images or flat-field images.

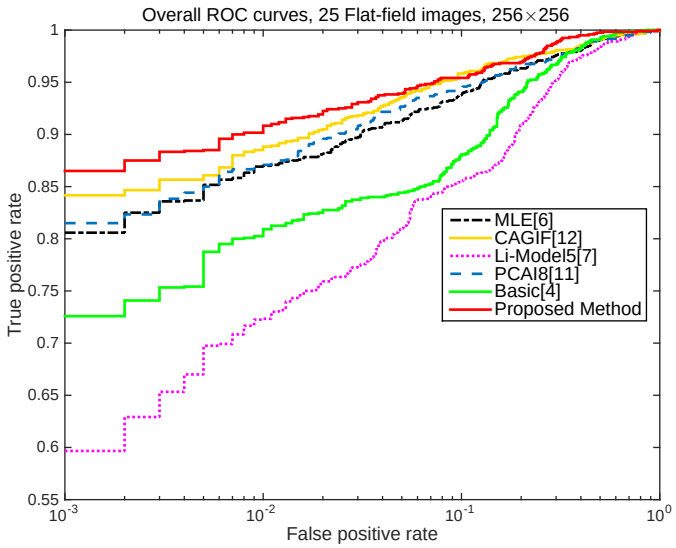


Fig. 5. The overall ROC curves of image size 256×256.

4.2. Performance analysis

4.2.1. Selection of the block size of the camera reference SPN

From the experimental results, it can be found that the identification results are affected by the image size. Furthemore, the block size of the camera reference SPN can also affect the identification results. Generally, the detection accuracy increases as the block size increases. However, when the block size exceeds a certain threshold, the accuracy will be decreased or tended to be stable. In order to demonstrate this rule, different block sizes for images with different sizes are investigated, and the results of image size 512×512 are shown in Fig. 6.

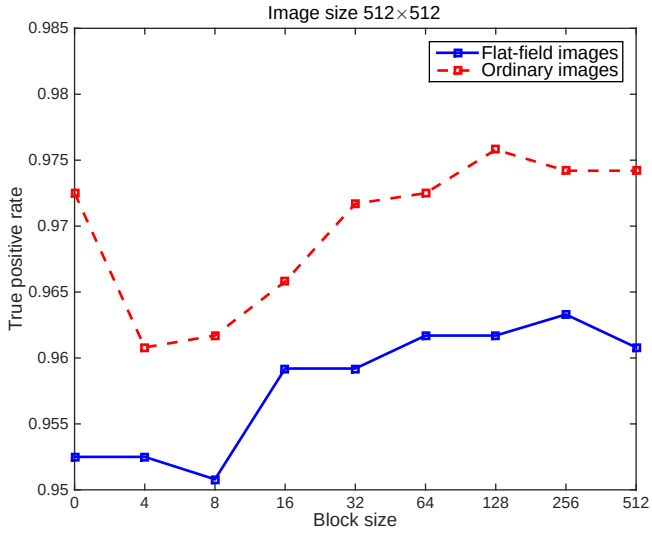


Fig. 6. The TPR of different block sizes.

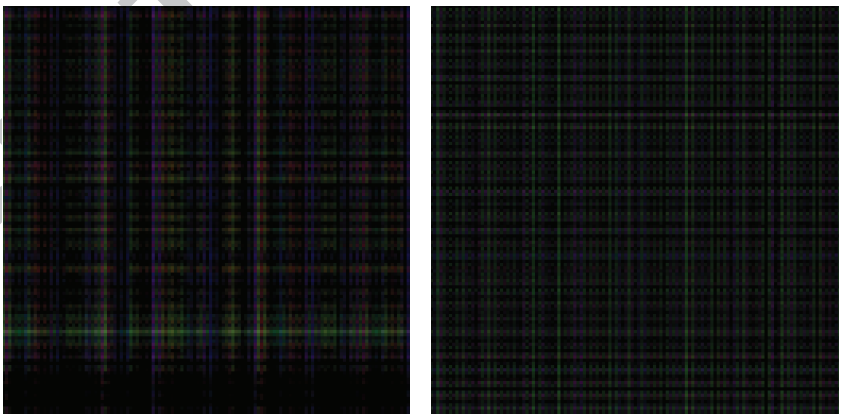
As seen from Fig. 6, the blue curve represents the camera reference SPN extracted from 25 flat-field images, while the red curve is the camera reference SPN extracted from 25 ordinary images, which contain the content of image scenes. Y axis represents TPR when FPR is 10^{-3}. X axis represents different block sizes. When the block size is 0, it

means that block weighted average is not used. For flat-field images, the largest TPR 96.33% can be achieved when the block size is 256×256. While for ordinary images, the largest TPR 97.58% can be achieved when the block size is 128×128. Both outperform the performance without using block weighted average. The results illustrate the rule mentioned above. Meanwhile, considering the balance of two situations, the best average accuracy can be achieved when the block size is 256×256, which is just the half of the image size. Therefore, considering the comprehensive performance of the camera reference SPN constructed from different image sizes and the time consumption, the image is recommended to be segmented into four parts. That is, $P = Q = 2$.

4.2.2. Analysis of the performance with different kinds of camera reference SPN

Comparing the experimental results in Section 4.1 (Table 3 and Table 4), it can be found that the performance of most methods, when the reference SPN is extracted from flat-field images, are better than the situation that it is extracted from ordinary images. It is mainly because the camera reference SPN extracted from ordinary images contains more residual information of the image scene than that extracted from the flat-field images, which will impose impacts on the correlation between the reference SPN and the SPN of the testing images, especially for the images with small size. This is the main reason why some blue sky images with no image scene content were generally used for the estimation of reference SPN in the previous literatures [6, 7, 11]. However, the proposed method can achieve relatively stable performance in the same image size when the reference SPN is extracted from ordinary images or flat-field images, which indicated that the proposed method is more suitable for the real forensics environment.

Meanwhile, comparing with the situation that the reference SPN is extracted from flat-field images, it is found that the performance of the Basic [4] and Li-Model5 [7], when the reference SPN is extracted from ordinary images, are significantly improved when the image size is large. For example, with an image size 512×512, the TPR of the Basic [4] is 92.75% when the reference SPN is extracted from ordinary images. Compared with the situation that the reference SPN is extracted from flat-field images, the improvement is 12.33%. The main reason is that the extracted reference SPN in Basic [4] and Li-Model5 [7] contain other imaging processing artifacts, which is named as linear pattern (the CFA and color interpolation component) [6], while this kind of artifacts are removed in the other methods ([6, 11, 12] and the proposed method). The linear pattern also make some contribution to the improvement of the camera brand or model identification. As known to us, the estimation of CFA interpolation is more accurate in non-smooth regions. Therefore, the reference SPN extracted from ordinary images provide more useful information of the camera. Here, the linear pattern of the Cannon Ixus55 camera are shown in Fig. 7.



(a) (b)

Fig. 7. Detail of the liner pattern for Cannon Ixus55. (a) Linear pattern extracted from 25 flat-field images. (b) Linear pattern extracted from 25 ordinary images.

As seen from Fig. 7, it can be found that the liner pattern extracted from ordinary images is uniform compared with that extracted from flat-field images. It means that the CFA interpolation extracted from ordinary images is well estimated, which can provide a stable and uniform signal to improve the identification. Therefore, the performance of

the Basic [4] and Li-Model5 [7] are significant improved when the reference SPN is extracted from ordinary images with large size.

4.3. The capability of resisting JPEG compression

As JPEG compression is generally used for digital images. To evaluate the influence of JPEG compression on the identification performance, different center sizes of the testing images in the dataset are compressed with quality factor (QF) 90, 80 and 70. Table 5 shows the TPR of different methods when FPR = 10^{-3} (the reference SPN extracted from 25 ordinary images), and the overall ROC curves of these methods are shown in Fig. 8, where the image size is 512×512, and QF = 90.

Table 5. The TPR of the different methods when FPR = 10^{-3} (%)

Image size	QF	Basic[4]	MLE[6]	Li-Model5[7]	PCAI8[11]	CAGIF[12]	Proposed Method
128×128	90	34.08	25.33	21.25	32.83	35.42	**43.00**
	80	21.58	18.85	11.58	25.75	16.67	**27.50**
	70	17.00	11.83	11.00	17.75	16.08	**20.42**
256×256	90	66.75	58.58	46.92	70.58	68.67	**75.00**
	80	59.67	54.17	19.92	56.08	59.42	**59.92**
	70	46.92	41.17	15.17	44.92	45.00	**47.42**
512×512	90	91.92	88.92	71.42	92.42	91.08	**93.75**
	80	83.92	81.58	44.5	**85.25**	84.25	84.58
	70	**77.67**	71.67	16.17	74.17	72.75	75.08
1024×1024	90	98.33	98.67	88.83	98.5	**98.75**	**98.75**
	80	98.08	97.83	82.5	97.83	**98.42**	98.25
	70	**97.33**	96.42	61.17	94.92	96.75	95.92

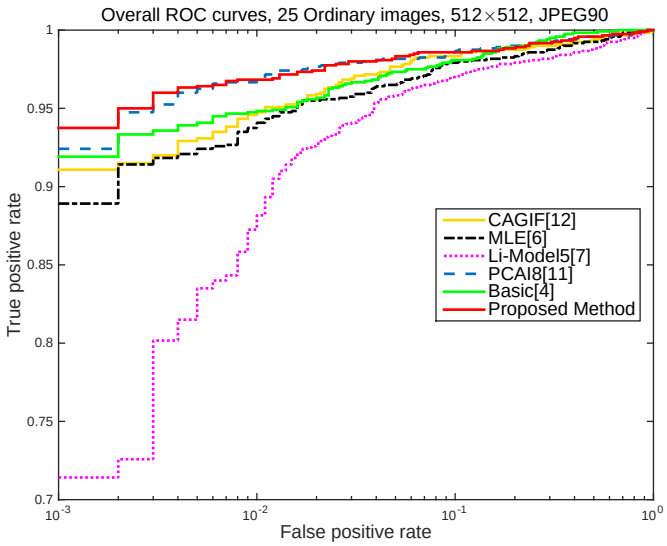


Fig. 8. The ROC curves with JPEG quality factor of 90.

As seen from Table 5, it can be found that the proposed method can still achieve good performance with JPEG compression in most cases, although the performance have declined with the reduction of QF. Compared with the images without compression (seen Table 3), it can be found that the performance of the proposed methods is less affected when QF = 90. The results illustrate that the proposed method is stable and robust in resisting JPEG compression with high quality factor compared with the other methods.

5. Conclusion and future work

In this paper, a source camera identification scheme based on the guided image estimation and block weighted average is proposed. It uses a guidance image to reduce the influence of the image scene and improve the quality of the SPN, and the reference SPN is constructed by block weighted average of the extracted SPNs. Experimental results show that the proposed method achieves fairly good performance compared with some existed schemes. Furthermore, with the situation that camera reference SPN can be only extracted from a few ordinary images, it still can get satisfactory performance.

Our future work opportunities include the extension of more sophisticated camera reference SPN construction strategy and improvement of the correlation calculation between the SPNs, such as clique-graph matching [29, 30], and more sophisticated machine learning method, such as multi-task structural learning [31, 32], to improve the identification performance. In addition, image search [33] and multimedia answer generation [34] methods for source camera identification will be utilized in network big data application.

Acknowledgements

This work was supported in part by project supported by National Natural Science Foundation of China (Grant No. 61572182, 61370225), project supported by Hunan Provincial Natural Science Foundation of China (Grant No. 15JJ2007), and supported by the Scientific Research Plan of Hunan Provincial Science and Technology Department of China (2014FJ4161).

The authors thank Prof. Xiangui Kang from Sun Yat-sen University for providing the source code of CAGIF method.

References

[1] K. San Choi, E. Y. Lam, K. K. Wong, Source camera identification using footprints from lens aberration, in: Electronic Imaging 2006, International Society for Optics and Photonics, 2006, pp. 60690J–60690J.

[2] M. K. Johnson, H. Farid, Exposing digital forgeries through chromatic aberration, in: Proceedings of the 8th workshop on Multimedia and security, ACM, 2006, pp. 48–55.

[3] A. C. Popescu, H. Farid, Exposing digital forgeries in color filter array interpolated images, IEEE Transactions on Signal Processing 53 (10) (2005) 3948–3959.

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

[5] Y. Sutcu, S. Bayram, H. T. Sencar, N. Memon, Improvements on sensor noise based source camera identification, in: 2007 IEEE International Conference on Multimedia and Expo, IEEE, 2007, pp. 24–27.

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

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

[8] J. Fridrich, Sensor defects in digital image forensic, in: Digital Image Forensics, Springer, 2013, pp. 179–218.

[9] X. Lin, C.-T. Li, Preprocessing reference sensor pattern noise via spectrum equalization, IEEE Transactions on Information Forensics and Security 11 (1) (2016) 126–140.

[10] G. Chierchia, S. Parrilli, G. Poggi, C. Sansone, L. Verdoliva, On the influence of denoising in PRNU based forgery detection, in: Proceedings of the 2nd ACM workshop on Multimedia in Forensics, Security and Intelligence, ACM, 2010, pp. 117–122.

[11] X. Kang, J. Chen, K. Lin, P. Anjie, A context-adaptive SPN predictor for trustworthy source camera identification, EURASIP Journal on Image and Video Processing 2014 (1) (2014) 1–11.

[12] H. Zeng, X. Kang, Fast source camera identification using content adaptive guided image filter, Journal of Forensic Sciences 61 (2) (2016) 520–526.

[13] R. Satta, Sensor pattern noise matching based on reliability map for source camera identification, in: International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP 2015), Vol. 1, 2015, pp. 222–226.

[14] F. Peng, D.-l. Zhou, M. Long, X.-m. Sun, Discrimination of natural images and computer generated graphics based on multi-fractal and regression analysis, AEU - International Journal of Electronics and Communications 71 (1) (2017) 72–81.

[15] M. J. Sorrell, Digital cannera source identification through JPEG quantisation, Multimedia forensics and security (2008) 291–335.

[16] E. J. Alles, Z. J. Geradts, C. J. Veenman, Source camera identification for heavily JPEG compressed low resolution still images, Journal of forensic sciences 54 (3) (2009) 628–638.

[17] B. Sankur, O. Celiktutan, I. Avcibas, Blind identification of cell phone cameras, Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX 6505 (2007) 1H–1I.

[18] P. Sutthiwan, J. Ye, Y. Q. Shi, An enhanced statistical approach to identifying photorealistic images, in: International Workshop on Digital Watermarking, Springer, 2009, pp. 323–335.

[19] G. Chierchia, S. Parrilli, G. Poggi, L. Verdoliva, C. Sansone, PRNU-based detection of small-size image forgeries, in: International Conference on Digital Signal Processing (DSP), 2011, pp. 1–6.
[20] G. Chierchia, G. Poggi, C. Sansone, L. Verdoliva, A Bayesian-MRF approach for PRNU-based image forgery detection, IEEE Transactions on Information Forensics and Security 9 (4) (2014) 554–567.
[21] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, Image denoising by sparse 3-D transform-domain collaborative filtering, IEEE Transactions on image processing 16 (8) (2007) 2080–2095.
[22] M. Goljan, Digital camera identification from images–estimating false acceptance probability, in: International Workshop on Digital Watermarking, Springer, 2008, pp. 454–468.
[23] X. Kang, Y. Li, Z. Qu, J. Huang, Enhancing source camera identification performance with a camera reference phase sensor pattern noise, IEEE Transactions on Information Forensics and Security 7 (2) (2012) 393–402.
[24] M. Goljan, M. Chen, P. Comesaña, J. Fridrich, Effect of compression on sensor-fingerprint based camera identification, Electronic Imaging 2016 (8) (2016) 1–10.
[25] D. Valsesia, G. Coluccia, T. Bianchi, E. Magli, Compressed fingerprint matching and camera identification via random projections, IEEE Transactions on Information Forensics and Security 10 (7) (2015) 1472–1485.
[26] K. He, J. Sun, X. Tang, Guided image filtering, in: European conference on computer vision, Springer, 2010, pp. 1–14.
[27] T. Gloe, S. Pfennig, M. Kirchner, Unexpected artefacts in PRNU-based camera identification: A'dresden image database'case-study, in: Proceedings of the on Multimedia and security, ACM, 2012, pp. 109–114.
[28] T. Gloe, R. Böhme, The dresden image database for benchmarking digital image forensics, Journal of Digital Forensic Practice 3 (2-4) (2010) 150–159.
[29] W.-Z. Nie, A.-A. Liu, Z. Gao, Y.-T. Su, Clique-graph matching by preserving global and local structure, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 07-12-June, 2015, pp. 4503–4510.
[30] A.-A. Liu, W.-Z. Nie, Y. Gao, Y.-T. Su, Multi-modal clique-graph matching for view-based 3D model retrieval, IEEE Transactions on Image Processing 25 (5) (2016) 2103–2116.
[31] A.-A. Liu, Y.-T. Su, P.-P. Jia, Z. Gao, T. Hao, Z.-X. Yang, Multipe/single-view human action recognition via part-induced multitask structural learning, IEEE Transactions on Cybernetics 45 (6) (2015) 1194–1208.
[32] A.-A. Liu, Y.-T. Su, W.-Z. Nie, M. Kankanhalli, Hierarchical clustering multi-task learning for joint human action grouping and recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence (2016). doi:10.1109/TPAMI.2016.2537337.
[33] L. Nie, M. Wang, Z.-J. Zha, T.-S. Chua, Oracle in Image Search: A content-based approach to performance prediction, ACM Transactions on Information Systems 30 (2) (2012) 1–23.
[34] L. Nie, M. Wang, Y. Gao, Z.-J. Zha, T.-S. Chua, Beyond text QA: Multimedia answer generation by harvesting web information, IEEE Transactions on Multimedia 15 (2) (2013) 426–441.

Highlights

- Guided image estimator can reduce the influence of image scene and improve the SPN quality
- Reference SPN is constructed from any natural images using block weighted average
- Experimental results illustrate the effectiveness of the proposed identification scheme