Anti-Forensics of Camera Identification and the Triangle Test by Improved Fingerprint-Copy Attack
Haodong Li, Weiqi Luo, Quanquan Rao, Jiwuu Huang

Abstract—The fingerprint-copy attack aims to confuse camera identification based on sensor pattern noise. However, the triangle test shows that the forged images undergone fingerprint-copy attack would share a non-PRNU (Photo-response nonuniformity) component with every stolen image, and thus can detect fingerprint-copy attack. In this paper, we propose an improved fingerprint-copy attack scheme. Our main idea is to superimpose the estimated fingerprint into the target image dispersely, via employing a block-wise method and using the stolen images randomly and partly. We also develop a practical method to determine the strength of the superimposed fingerprint based on objective image quality. In such a way, the impact of non-PRNU component with every stolen image, and thus can detect fingerprint-copy attack. The experiments evaluated on 2,900 images from 4 cameras show that our fingerprint-copy attack is difficultly detected. The experiments show that the forged images undergone fingerprint-copy attack scheme can effectively fool camera identification, and significantly degrade the performance of the triangle test simultaneously.

Index Terms—Anti-forensics, camera identification, photoresponse nonuniformity, sensor fingerprint, triangle test

I. INTRODUCTION

With the various powerful image editing software, such as GIMP and Adobe Photoshop, modifying digital images becomes increasingly easy, which would inevitably lead to some potential moral and/or legal consequences. Nowadays, digital image forensics become an important issue. Image forensics involve many technical methods for different applications such as detecting image compression, exposing image processing history, revealing tampered images, differentiating between photorealistic and photographic images, identifying the acquisition component or device of an image, and so on. As one of the important issues in image forensics, source camera identification aims to identify which camera was used to capture the given image. Kurosawa et al. [17] carried out the identification utilizing fixed pattern noise (FPN) appearing on dark frames; Kharrazi et al. [18] proposed a feature set based on the color of pixel and image quality metrics. In [13], Lukáš et al. proposed a novel source camera identification method based on PRNU (Photo-response nonuniformity). By averaging the noise components of sufficient images from a given camera, the obtained result can be regarded as an estimation of the PRNU component, which can serve as a reliable and unique fingerprint for the camera. In [19], Chen et al. proposed a method to estimate the PRNU factor based on maximum-likelihood approach, and shown that the method outperformed that proposed in [13]. Besides, several improved PRNU-based methods have been proposed for source camera identification. For example, Li [20] proposed an enhanced method to suppress some unwanted components within the noise components. Hu et al. [21] proposed a way to construct camera fingerprints from all of the three color channels via using the characteristics of the color filter array (CFA) structure. Kang et al. [22] introduced a new detection method to lower the false positive rate. Bayram et al. [23] tried to reduce the computational expensive of conventional approaches through binarizing the sensor fingerprints. In [24] and [25], two methods are proposed to estimate the sensor fingerprints based on local information of image via block-wise strategies.

On the opposite side of forensics, however, a wise attacker can perform some anti-forensic operations to confuse the corresponding forensic methods. Studying the anti-forensic methods is very important for researchers, since it can help to discover the limitations of current forensic methods and further improve their performance. Up to now, several typical anti-forensic works have been reported. For instance, Stamm and Liu [26] proposed a scheme to conceal the traces left by digital image compression. In [27] and [28], some methods are proposed to trick the resampling detector and the sensor pattern noise based image forgeries detector, respectively. In [29], the authors proposed a method to remove the PRNU noise within an image so that it would not be matched with its source camera. Some methods such as [30] and [31] try to confuse the source camera identification as described above via fingerprint-copy attack. These methods firstly estimate the fingerprint of a camera $C_A$ from some stolen images, and then superimpose it into a target image taken by a different camera $C_E$ to disguise the resulting image as one taken by camera $C_A$. Most existing PRNU based camera identification algorithms would incorrectly identify that the resulting forged image is taken by camera $C_A$ rather than camera $C_E$.

Forensics and anti-forensics is like the cat-and-mouse game. Recently, Goljan et al. [30] proposed the triangle test method to expose the fingerprint-copy attack. They pointed out that there would be a common non-PRNU component shared by those stolen images and the resulting forged images after performing the fingerprint-copy attack. Based on such a property, it can identify the stolen images from some candidate images.
using the individual test, determine whether a set of candidate images contain some stolen images using the pooled test, and directly detect the forged images using the multiple forgeries test.

In this paper, we propose an anti-forensic method to deceive both camera identification and the triangle test. The main idea is to randomly select just a small part of stolen images for fingerprint estimation, and then superimpose the estimated fingerprint into the target image block by block. By doing so, we are able to reduce impact of the non-PRNU components from the stolen images, and thus it would deceive the triangle test. Moreover, we propose a practical way to determine the strength of the superimposed fingerprint based on adjusting PSNR, so we can overcome the drawback of conventional fingerprint-copy attack that the true fingerprint needs to be available. In our experiments, 2,900 images from three different camera brands (four camera individuals) and three different kinds of the triangle tests have been included. The experimental results have shown that the proposed method outperforms the conventional fingerprint-copy attack method [30] and degrades the performance of triangle test significantly.

The rest of this paper is organized as follows. Section II gives a brief overview of some related techniques on camera identification based on PRNU. Section III presents the proposed anti-forensic method, including superimpose the estimated fingerprint dispersedly and determine the fingerprint strength according to PSNR. Section IV shows the experimental results and discussions. Finally, the concluding remarks and future works are given in Section V.

II. RELATED WORKS

In this section, we will introduce some related works including the camera identification method based on PRNU, the conventional fingerprint-copy attack method and the triangle test method. Beforehand, some notations used in this paper are given: everywhere a boldface symbol represents a matrix, e.g. \( \mathbf{X} \); for two matrices \( \mathbf{X} \) and \( \mathbf{Y} \) with the same dimensions, their element-wise product and element-wise division are denoted as \( \mathbf{Z} = \mathbf{X} \mathbf{Y} \) and \( \mathbf{Z} = \mathbf{X} / \mathbf{Y} \) respectively.

A. Camera Identification Based on PRNU

Due to the manufacturing imperfection and the nonuniformity of silicon wafers, the recorded pixels on camera sensors vary from each other even if the sensors are exposed to the same illumination. Therefore, a kind of pattern noise called PRNU exists in all sensor-based cameras, and it is introduced to every image taken by camera. Such noise is independent of the environment. More importantly, the noises of all images taken by a same camera are highly correlated, while the correlations among the noises of those images taken by different cameras would be much weaker. Based on this property, therefore, the pattern noise can be regarded as a unique fingerprint of a digital camera.

In [19], Chen et al. proposed a maximum likelihood method to estimate the PRNU multiplicative factor \( \hat{\mathbf{K}} \) of a camera \( \mathcal{C}_A \)

\[
\hat{\mathbf{K}} = \frac{\sum_{i=1}^{N} \mathbf{W}_i \mathbf{I}_i}{\sum_{i=1}^{N} \mathbf{I}_i^2} \tag{1}
\]

where \( \mathbf{I}_1, \ldots, \mathbf{I}_N \) are the \( N \) images taken by camera \( \mathcal{C}_A \) that are used to estimate the fingerprint, and \( \mathbf{W}_i = \mathbf{I}_i - F(\mathbf{I}_i) \) is the noise residual of \( \mathbf{I}_i \), where \( F \) is the wavelet denoising filter [32]. For a given image \( \mathbf{J} \), the presence of the camera fingerprint can be evaluated by the correlation detector:

\[
\rho = \text{corr}(\mathbf{W}_j, \hat{\mathbf{K}}) \tag{2}
\]

where \( \hat{\mathbf{K}} \) is the estimated fingerprint of camera \( \mathcal{C}_A \) computed by Eq. (1). Thus, camera identification can be achieved with a detector based on the statistic \( \rho \). By setting a false alarm rate \( P_{fa} \) (i.e. the probability of that those images taken by other cameras are identified as being taken by camera \( \mathcal{C}_A \)), some training images taken by other cameras are used to determine the detection threshold \( t_1 \). Then for a testing image \( \mathbf{J} \), its corresponding \( \rho \) is obtained based on Eq. (2). If \( \rho > t_1 \), \( \mathbf{J} \) is decided to be taken by camera \( \mathcal{C}_A \); otherwise, \( \mathbf{J} \) is not taken by camera \( \mathcal{C}_A \). It is noted that the PRNU based correlation detection is an important method in camera identification, its reliability and effectiveness have been verified in [13] and [19].

B. Fingerprint-copy Attack

The purpose of fingerprint-copy attack [30] is to confuse those threshold-based camera identification algorithms. Assume that Alice is the victim and Eve is the attacker. Alice has posted some images taken by her camera \( \mathcal{C}_A \) on some websites, and Eve steals \( N \) of these images. In order to make a forgery, Eve can firstly suppress the fingerprint within a target image \( \mathbf{J} \) taken by a different camera \( \mathcal{C}_E \). Next, she estimates a fake fingerprint \( \hat{\mathbf{K}}_E \) of camera \( \mathcal{C}_E \) with Eq. (1) using the stolen images. Finally, she superimposes the fake fingerprint \( \hat{\mathbf{K}}_E \) into \( \mathbf{J} \) and makes a forgery \( \mathbf{J}' \)

\[
\mathbf{J}' = [\mathbf{J}(1 + \alpha \hat{\mathbf{K}}_E)] \tag{3}
\]

where the symbol \( [\cdot] \) is the rounding and truncating operation (a real number is rounded to its nearest integer, and it will be truncated to 0 or 255 if it is outside the range of [0,255]), and the parameter \( \alpha > 0 \) denotes the fingerprint strength. In order to create a good forgery, Eve must set a proper parameter \( \alpha \) to make the response of the camera identification detector (e.g. the correlation detector using Eq. (2)) on \( \mathbf{J}' \) “natural” as if \( \mathbf{J}' \) was indeed taken by camera \( \mathcal{C}_E \). In this case, the fingerprint strength \( \alpha \) should be adjusted according to the true fingerprint \( \mathbf{K} \) of camera \( \mathcal{C}_A \).

Please note that two issues should be further considered in this fingerprint-copy attack strategy. Firstly, is it a good strategy to use all the \( N \) stolen images for estimating the whole fake fingerprint \( \hat{\mathbf{K}}_E \)? Secondly, how to adjust the fingerprint strength \( \alpha \) to make the correlation computed by Eq. (2) on \( \mathbf{J}' \) “natural”? Since Eve cannot access the true fingerprint \( \mathbf{K} \) of camera \( \mathcal{C}_A \), it seems unreasonable for her to obtain a natural \( \alpha \).
in real applications. To avoid this assumption, we will develop a method in Section III via adjusting PSNR to determine the proper fingerprint strength.

C. Triangle Test

In order to prove her innocence, Alice can perform the triangle test [30] to expose the existence of fingerprint-copy attack.

One possible scenario is that Alice only gets a forged image $J'$, and all the $N$ images stolen by Eve belong to a large database, containing $N_c$ ($N_c > N$) images, which is available to Alice. In such a case, Alice tries to identify the stolen images from the image database so as to prove $J'$ is a forgery. Since Alice has the camera $C_A$, she can obtain an estimated fingerprint $\hat{K}_A$ of camera $C_A$ via Eq. (1) using some well designed images (e.g. blue sky images) which are not in the image database $N_c$. Then for each candidate image $I$ in the database, she computes the three correlations between the noise residual of $I$, the noise residual of $J'$, and the estimated fingerprint $\hat{K}_A$, denoted as $c_{IJ'}, c_{I,\hat{K}_A}$, and $c_{J',\hat{K}_A}$ respectively, as shown in Fig. 1.

In [30], the authors obtained an estimation of $c_{IJ'}$ (denoted as $\hat{c}_{IJ'}$) from the two other correlations $c_{I,\hat{K}_A}$ and $c_{J',\hat{K}_A}$. They further showed that for a forged image $J'$, the $c_{IJ'}$ and $\hat{c}_{IJ'}$ were highly linear dependent if $I$ was not used by Eve. If the image $I$ was previously used by Eve, however, such a linear property would be destroyed and $c_{IJ'}$ would present a higher value (Please refer to Fig. 2(a) for example), since $I$ and $J'$ share a non-PRNU component between each stolen image $I$ and $J'$.

In order to prove her innocence, Alice can perform the triangle test on $W_{IJ'}$, $W_{\hat{K}_A}$, and $\hat{K}_A$ without accessing the $N$ stolen images used by Eve.

III. THE PROPOSED FINGERPRINT-COPY ATTACK SCHEME

To design a successful fingerprint-copy attack, the two following requirements should be carefully considered.

(a) Confuse the camera identification. Since Eve's original intention is to frame Alice as the owner of the forged image $J'$. Therefore, after superimposing a fake fingerprint into a target image, it should guarantee that the camera identification detectors cannot distinguish the forged image $J'$ from the original ones.

(b) Fool the triangle test. To avoid being detected, the performance of the triangle test evaluated on the resulting forgeries should degrade significantly after using the fingerprint-copy attack.

There is a tradeoff between the two requirements. Usually, the more strength of a fake fingerprint is superimposed, the easier to confuse the camera identification (i.e. requirement #a), however, the better detection performance of the triangle test (i.e. requirement #b) would be obtained. Based on the analysis in Section II.C we find that the key idea of the triangle test is to detect the non-PRNU component from the stolen images. If we can reduce the impact of the non-PRNU component on the triangle test, it is expected that the performance of the triangle test will drop significantly.

A. A Block-wise and Randomized Fingerprint-copy Scheme

Unlike the conventional fingerprint-copy attack described Section II.B in this subsection we introduce a block-wise and randomized fingerprint-copy scheme. Assuming that we have stolen $N$ images taken by Alice's camera, we produce the fake fingerprint for different blocks of the target image by just using a randomly selected subset of the $N$ stolen images. The details are as follows.

Firstly, we divide all stolen images $I_1, \ldots, I_N$ (assuming that their sizes are the same as the target one) and the target image $J$ into $n$ non-overlapping blocks with the size of $l \times l$. For each block $J_b$ of image $J$, where $b = 1, 2, \ldots, n$, we randomly select $r (r < N)$ stolen images $I_1', \ldots, I_r'$, and calculate $\hat{K}_{E,b}$ according to Eq. (1) using the corresponding blocks $I_1', \ldots, I_r'$ of the selected stolen images, then we superimpose this fake fingerprint block $\hat{K}_{E,b}$ into the corresponding block $J_b$ within the target image,

$$J_b' = [J_b(1 + \alpha_b \hat{K}_{E,b})]$$

where $\alpha_b$ is the fingerprint strength for the $b$th block. We will discuss how to select a proper $\alpha_b$ in Section III-B.

In such a way, the fake fingerprint superimposed into the target image $J$ is dispersed, so the common non-PRNU component between each stolen image $I$ and the forged image $J'$ is also dispersed. As a result, it is expected that the impact of the non-PRNU component on the triangle test will be reduced. Based on our experiments (please see Section IV-C to Section...
On the other hand, it is noted that the dispersion of fingerprint IV -E), the performance of the triangle test is significantly de-

Section IV -B), at the same level of PSNR, the proposed

identification. Based on our experimental results (please s ee

has slight effect on the performance for confusing the camer a

compared to the conventional method [30].

B. Determining the Fingerprint Strength

An important factor in Eq. [3] is the fingerprint strength \( \alpha_0 \). Since a forger does not know the true fingerprint \( \mathbf{K} \) in practice, we try to determine \( \alpha_0 \) based on the objective image quality measured by PSNR as the following equation,

\[
\alpha_0^* \in \{ \alpha_0 | \text{PSNR}(\mathbf{J}_b^*, \mathbf{J}_b) = A \} \tag{5}
\]

where PSNR\((\mathbf{J}_b^*, \mathbf{J}_b)\) stands for the PSNR between an image block after and before superimposing the fake fingerprint, and \( A \) is a target PSNR value. In other words, we select \( \alpha_0^* \) so that the PSNR between the forged image block \( \mathbf{J}_b^* \) and the source image block \( \mathbf{J}_b \) is equal to the given PSNR value \( A \). In practice, it is not easy to obtain the solution \( \alpha_0^* \) directly. Besides, there would be several \( \alpha_0 \) that satisfy PSNR\((\mathbf{J}_b^*, \mathbf{J}_b) = A \) due to the rounding/truncation errors incurred in generating \( \mathbf{J}_b^* \). Therefore, we use binary search to approximately find out \( \alpha_0^* \), which can ensure PSNR\((\mathbf{J}_b^*, \mathbf{J}_b) \) is as close to \( A \) as possible.

Based on the work in [30], we note that the PSNR between the original image and the forged image after the fingerprint-
copy attack with the “natural” parameter \( \alpha \) would approximately fall in the range of [47.6 dB, 58.7 dB]. Therefore, we can select the value of \( A \) within this range. Usually, the larger

the \( A \) to be selected, the smaller the \( \alpha_0^* \) is obtained, and vice versa.

C. Implementation of the Proposed Method

Based on the experiments and discussions in the Section III-A and III-B we summarize the procedures for creating a forged image \( \mathbf{J}' \) as follows.

(a) Divide the target image \( \mathbf{J} \) and all the stolen images \( \mathbf{I}_1, \ldots, \mathbf{I}_N \) into \( n \) non-overlapping small blocks. For each block \( \mathbf{J}_b \) \((b = 1, 2, \ldots, n)\), randomly select \( r \) \((r \leq N)\) corresponding stolen image blocks \( \mathbf{I}_{r,b} \) to estimate \( \mathbf{K}_{E,b} \) according to Eq. (1).

(b) For each block \( \mathbf{J}_b \), we set \( A \) (the target PSNR) for adjusting fingerprint strength, and then we find the proper fingerprint strength \( \alpha_0^* \) according to Eq. (5), finally superimpose the estimated fingerprint \( \mathbf{K}_{E,b} \) into \( \mathbf{J}_b \) with Eq. (3) (set \( \alpha_0^* \) as \( \alpha_0^* \) to obtain a forged block \( \mathbf{J}_b^* \)).

(c) After processing all the blocks \( \mathbf{J}_b \) \((b = 1, 2, \ldots, n)\) as described in step (b), we combine all \( \mathbf{J}_b^* \) \((b = 1, 2, \ldots, n)\) together to generate a forged image \( \mathbf{J}' \).

By choosing proper parameters \((i.e. r \text{ and } A)\), we can achieve a good tradeoff between the two requirements described previously. Comparing with the conventional fingerprint-copy attack [30], the main difference of the proposed method described above is that we superimpose the estimated fingerprint into a target image block by block rather than the whole image, and we just randomly use some stolen images to estimate the corresponding fingerprint for each forged image block rather than all stolen images. Furthermore, we obtain the proper fingerprint strength via adjusting the PSNR rather than using a natural \( \alpha \) which is not available in practice.

An image forgery made by the proposed method with parameter \( N = 100, A = 50 \) and \( r = 10 \) is illustrated in Fig. 2. Since the PSNR between this forgery and its original counterpart is 50 dB, no perceptible artifacts are introduced in the resulting image. This forgery can successfully confuse the camera identification and fool the triangle test simultaneously. More comparative results will be given in the Section IV.

IV. EXPERIMENTAL RESULTS

All our experiments are conducted on 2,900 images taken by four different digital cameras, including Nikon D7000 #1, Nikon D7000 #2, Canon EOS 400D, and Pentax K20D. The number of images for each camera and their imaginary owners are shown in Table I. Please note that there are two digital cameras with the same model, Nikon D7000 \((i.e. \mathbf{C}_A \text{ and } \mathbf{C}_E)\), since Eve may create the image forgery according to the model of Alice’s camera to avoid introducing other detectable artifacts, such as inconsistent CFA [33] and quantization matrix [34]. Besides, camera \( \mathbf{C}_E \) and \( \mathbf{C}_E3 \) with different brands and models are also included in our experiments.

In order to reduce the computational complexity, all images are center cropped with the size of 1024 \( \times \) 1024, which are then JPEG compressed with a high quality factor 90 just as did in
TABLE I
IMAGE DATABASES USED IN OUR EXPERIMENTS

| Camera name | Camera brand | Number of images | Owner |
|-------------|--------------|------------------|-------|
| CA1         | Nikon D7000 #1 | 2000             | Alice |
| CE1         | Nikon D7000 #2 | 300              | Eve   |
| CE2         | Canon EOS 400D | 300              | Eve   |
| CE3         | Pentax K20D   | 300              | Eve   |

TABLE II

| Block size | Pfa (%) | PD (%) |
|------------|---------|--------|
| 8          | 96.33   | 24.56  |
| 16         | 99.33   | 44.22  |
| 32         | 100     | 54.19  |
| 64         | 100     | 58.29  |
| 128        | 100     | 56.55  |
| 256        | 100     | 49.35  |
| 512        | 100     | 31.21  |

(a) l = 8

(b) l = 32

(c) l = 512

(d) Conventional method (l = 1024)

Fig. 3. True correlation \( c_{IJ} \) vs. estimated correlation \( \hat{c}_{IJ} \) for the forged image #198 in \( CE_1 \). The blue circles, black dots, and red dashed lines denote the images used by Eve, the images not used by Eve, and the detection thresholds, respectively. The dash-dotted lines are fitted lines of the black dots. Please note that the results for other images are similar based on our experiments.

We assume that the attacker Eve steals \( N \) images from \( CA \) and creates some forgeries for each target image from \( CE_1 \), \( CE_2 \) and \( CE_3 \) using the conventional fingerprint-copy attack \([30]\), and the proposed method with different parameters. For a fair comparison, we also adjust the \( \alpha \) in Eq. (3) based on PSNR like Eq. (5) when applying the conventional method, since Eve cannot obtain the true fingerprint \( K \) of camera \( CA \) in real forensic cases. As it did in \([30]\), the target images are slightly denoised by the wavelet filter \([32]\) with parameter \( \sigma = 1 \) to suppress their own fingerprints and other possible artifacts before superimposing the fake estimated fingerprint.

### A. Parameter Selection

1) Selection of Parameter \( l \): Since the proposed method works in a block-wise manner, the parameter \( l \) (i.e. block size) would have impact on the anti-forensic performance. In order to select a proper value of \( l \), we have conducted the following experiments.

By setting \( N=100, A=50, r=10 \), and \( l \in \{8, 16, 32, 64, 128, 256, 512\} \), we create 300×7 (with regard to different \( l \)) forged images from camera \( CE_1 \) using the proposed method. Then we investigate the performance of camera identification and individual triangle test, respectively. For the forged images, we aim to increase the false alarm rate of camera identification (Pfa) and decrease the correct detection rate of triangle test (PD). The experimental results are shown in Table II. From Table II it is observed that we obtain Pfa = 100% when \( l \geq 32 \), meaning that the camera identification is successfully deceived. When \( l \) is very small (e.g. 8) or very large (e.g. 512), the obtained PD of triangle test is relatively smaller than those obtained with a middle range \( l \) (e.g. 1024 or 256). Due to a small \( l \) would lead to poor performance for deceiving the camera identification, it may be better to choose a large \( l \).

However, our further analysis indicates that a too large \( l \) would bring some potential risks. As illustrated in Fig. 8 the true correlation values (blue circles) for relatively small \( l \) (i.e. Fig. 8(a) and Fig. 8(b)) are always close to the detection threshold (red dashed line), and most of them are less than \( 8 \times 10^{-3} \). However, some of the true correlation values for a large \( l \) (i.e. Fig. 8(c)) are significantly larger than the threshold.

[30] to avoid the inconsistence of JPEG quantization matrices.
Comparing Fig. 3(a) and Fig. 3(b), it is observed that many true correlation values (larger than $1 \times 10^{-2}$ in this experiment) obtained with $l = 512$ are even larger than those obtained with the conventional method ($l = 1024$). These values can be regarded as outlier data, which can be used as a strong sign for image forgery. Such a case is not shown if we just focus on the $P_\beta$ of triangle test, but the one who attempts to perform anti-forensics must pay attention to it in a real scenario. To avoid such traces, $l$ should not be too large. Thus, we select $l = 32$ in the following experiments.

2) Selection of Parameter $A$ and $r$: There are two important parameters in the proposed method, i.e., the target PSNR $A$ for setting the parameter $\alpha_b$ in Eq. (5) and the number of randomly selected images $r$.

It is noted that the PSNR between the images before and after the fingerprint-copy attack would approximately fall in the range of $[47.6 \, \text{dB}, \, 58.7 \, \text{dB}]$ (please refer to Table I in [30]). Therefore, we set $A$ as 50 dB and 55 dB to evaluate the performance of the proposed method. Since the proposed method ensures the PSNR for each image block is as close to $A$ as possible, the PSNR for the whole image is also very close to the $A$.

Another parameter need to be determined is $r$. To confuse the threshold-based correlation detector, we should guarantee that the correlation between the noise residual $\mathbf{w}_J$ of the forged image and the true fingerprint $\mathbf{K}$ of camera $\mathcal{C}_A$ is large enough. Usually, for a given $A$ (i.e. the strength of the estimated fingerprint superimposed into $\mathbf{J}$ is given), the more stolen images we used for estimating the fingerprint, the larger the correlation we can obtain, while the poorer performance for attacking the triangle test (refer to Table IV [17] for more details). In order to select a proper $r$, the following experiments are conducted.

Firstly, we determine the threshold $t_1$ for the threshold-based correlation detector. We collect 300 blue sky images taken by camera $\mathcal{C}_A$ and obtain a good estimation $\hat{\mathbf{K}}$ of the true fingerprint of camera $\mathcal{C}_A$. Then we use the 2,000 images from camera $\mathcal{C}_A$ as positive samples and collect 2,000 images from the Dresden Image Database (available at [http://forensics.inf.tu-dresden.de/didingdb]) taken by 10 different cameras (including 5 different camera models, denote as $\mathcal{C}_{\text{other}}$) as negative samples. We calculate the correlations of these images with $\hat{\mathbf{K}}$ according to Eq. (2). Finally, we set the false alarm rate $P_{fa}$ as $10^{-3}$ to obtain the threshold $t_1$ ($t_1 = 0.0046$ in this experiment). If the correlation of an image is larger than $t_1$, the detector would identify the image as being taken by camera $\mathcal{C}_A$. In Fig. 4, the boxplots of the correlations for all the images from different cameras are illustrated. It is observed that almost all images from camera $\mathcal{C}_A$ (see the leftmost boxplot) and from other cameras (i.e. $\mathcal{C}_{\text{other}}$, $\mathcal{C}_{E1}$, $\mathcal{C}_{E2}$ and $\mathcal{C}_{E3}$, see the two boxplots in the middle) are clearly separated by the threshold $t_1$.

By setting $N = 100$, we try to analyze the effect of the parameter $r$ on the correlation detection. Please note that similar results can be obtained for other values, such as $N = 20, 50, 200$ and 300 based on our experiments. Fig. 5 shows the $P_{fa}$ of the 900 forged images (i.e. the percentage of those forged images whose correlations are larger than the threshold $t_1$) with increasing the parameter $r$ from 1 to 20. It can be observed that the $P_{fa}$ would increase with increasing the parameter $r$. At the beginning, e.g., $r$ ranging from 1 to 9, the $P_{fa}$ increases sharply. When $r$ is larger than 15, it increases slightly, which means that there is no need to use all the stolen images for fingerprint estimation, since the more stolen images to be used for estimating the fake fingerprint, the more easily the triangle test detects the resulting forgeries. Based on the above analysis, we evaluate the proposed method with the parameter $r \in \{10, 15\}$ in the following experiments.

B. Experiment #1: Confusing Camera Identification

In this experiment, we will evaluate the proposed method for confusing the camera identification. In this case, each of the image forgeries is tested by the correlation detector, and a higher value of $P_{fa}$ means the better anti-forensic performance.

The experimental results evaluated on those forgeries from the three different cameras are shown in Table III. It can be observed that the parameter $A$ is the most influential factor for the detection performance, since it determines the fingerprint strength. Usually, the smaller $A$ we use, the larger the fingerprint strength is, and thus the better performance for confusing the camera identification will be achieved. When $A = 50$ dB, both the conventional method and the proposed method can achieve very high $P_{fa}$ (all larger than 97%). As an example, we show the correlations for the forged images from $\mathcal{C}_{E1}$, $\mathcal{C}_{E2}$ and $\mathcal{C}_{E3}$ made by the proposed method with
TABLE III

$P_{fa}$ (%) of Camera identification for the forgeries from three cameras.

| $N$ | 20 | 50 | 100 | 200 | 300 |
|-----|----|----|-----|-----|-----|
| A = 50 | Method in [30] | 100 | 100 | 100 | 100 |
| | Proposed $(r = 15)$ | 100 | 100 | 100 | 100 |
| | Proposed $(r = 10)$ | 100 | 99.67 | 100 | 99.67 |
| A = 55 | Method in [30] | 97.67 | 99.67 | 100 | 100 |
| | Proposed $(r = 15)$ | 96.00 | 96.00 | 96.00 | 95.67 |
| | Proposed $(r = 10)$ | 94.00 | 93.67 | 94.33 | 93.00 |
| A = 50 | Method in [30] | 96.67 | 99.53 | 100 | 100 |
| | Proposed $(r = 15)$ | 99.33 | 99.33 | 99.33 | 99.00 |
| | Proposed $(r = 10)$ | 99.00 | 99.00 | 99.00 | 98.67 |
| A = 55 | Method in [30] | 92.67 | 97.67 | 98.67 | 99.00 |
| | Proposed $(r = 15)$ | 83.00 | 84.00 | 85.33 | 82.67 |
| | Proposed $(r = 10)$ | 79.33 | 79.33 | 80.33 | 76.67 |
| A = 50 | Method in [30] | 99.67 | 99.67 | 99.67 | 99.67 |
| | Proposed $(r = 15)$ | 99.00 | 98.67 | 99.00 | 98.67 |
| | Proposed $(r = 10)$ | 98.33 | 98.00 | 98.00 | 97.67 |
| A = 55 | Method in [30] | 89.67 | 96.00 | 99.00 | 99.67 |
| | Proposed $(r = 15)$ | 86.67 | 86.33 | 87.33 | 84.67 |
| | Proposed $(r = 10)$ | 80.67 | 82.00 | 82.33 | 81.33 |

**Fig. 6.** Correlations $c_{fj}$ vs. $c_{ej}$ for the forged image shown in Fig. 2 using the conventional method [30] and the proposed method. The parameters are $N = 100, r = 10, A = 50$ dB. The dash-dotted line is the fitted straight line of the black dots.

The decrements for the proposed method are severer comparing with the conventional method. Besides, we find that the changes slightly for the parameter $N$. Based on Table III, giving A and r, the changes of $P_{fa}$ for the proposed method are less than 4% when N ranges from 20 to 300. Overall, the experimental results show that we can set A as 50 dB to make almost all forged images successfully confuse the camera identification.

**C. Experiment #2: Attacking Individual Triangle Test**

In this experiment, we will evaluate the proposed method for attacking individual triangle test. In this scenario, Alice try to detect the forgeries via identifying those images stolen by Eve, so a good anti-forensic method should reduce the detection rate $P_D$ of individual test as far as possible. In our experiments, we firstly estimate an approximate fingerprint $K_j$ using $N_j = 70$ blue sky images from camera $C_j$. [30] has pointed out that it is not necessary for Alice to work with a better quality fingerprint than Eve, and $N_j$ is recommended from 15 to 70. For each forged image $J_f$, the individual test is performed on 600 images taken by camera $C_i$. Among the 600 images, $N$ ($N \in \{20, 50, 100, 200, 300\}$ in our experiments) stolen images used by Eve are treated as candidate images, and the remaining $600-N$ images are used to estimate the pdf $f_{fj}(x)$ for $J_f$ (refer to Section III-C for the details of $f_{fj}(x)$).

In Fig. 6 we plot the $c_{fj}$ versus $c_{ej}$ for two forgeries created from the image example illustrated in Fig. 2 with the conventional method [30] and the proposed method, respectively. Here, the parameters are set as $N = 100, A = 50$ dB, and $r = 10$. Compared with the Fig. 6(a), it is observed that the separations between the images used and those not used by Eve significantly deteriorate after applying the proposed method. When $P_{fa}$ (i.e., the false alarm rate of triangle test) is set as $10^{-3}$, the $P_D$ for the conventional method is as high as 96%, while the $P_D$ for the proposed method is only 39%.

To further show the effectiveness of the proposed method, we conduct the individual test on all the forgeries made with different methods and parameters. The average $P_D$ evaluated on 300 forged images from each of the three cameras are shown in Table IV. Overall, the proposed method outperforms the conventional method significantly, especially when $N$ is large. For instance, when $N > 50$ and $A = 50$ dB, the average decrements of $P_D$ for the proposed method are over 48% comparing with the conventional method. Besides, the larger
the number of stolen images \( N \) is, the poorer detection performance for the individual triangle test we achieve. When \( N > 100 \), \( P_D \) for the proposed method with \( A = 55 \text{ dB} \) are less than 4\%, meaning that the individual triangle test would almost become useless in these cases.

D. Experiment #3: Attacking Pooled Triangle Test

In this experiment, we will evaluate the proposed method for attacking the pooled triangle test. Firstly, we randomly select 100 target images from each of the camera \( C_{E1}, C_{E2} \) and \( C_{E3} \) respectively, and create totally 1,800 forgeries using the selected images with the proposed method and the conventional method and the parameters \( N \in \{100, 200, 300\} \), \( A = 50 \text{ dB} \), and \( r = 10 \). Please note that the two methods with these parameters can obtain similar good results (all \( P_{fa} > 96\% \)) on confusing camera identification according to the experiment \#1. For each forged image \( J' \), the pdf \( f_{J'}(x) \) is estimated from 500 images taken by camera \( C_{E} \) which are not used by Eve. Assume that the \( N \) stolen images are a subset of the \( N_c \) candidate images \( I \) from camera \( C_A \) (\( N_c \geq N \)). For a given ratio \( N/N_c, k = 60 \) images are randomly selected out of the \( N_c \) candidate images and their statistics \( c_{I,J'} - \lambda c_{I,J'} - \eta \) are computed. If the \( p \)-value of these statistics is less than \( P_{fa} = 10^{-3} \), we determine \( J' \) is a forgery. Like what was done in [30], we repeat the process of random selecting \( k \) images and making a decision on \( J' \) 10,000 times, and finally obtain the \( P_D \) for \( J' \), i.e., the probability of that \( J' \) is correctly detected as a forgery over the 10,000 times.

Fig. 7 shows the ratio \( N/N_c \) (please note that \( N \) is set as 100, 200 or 300, \( N_c \) would change with different ratio of \( N/N_c \)) versus the average \( P_D \) on 100 images from the three cameras. It is observed that the conventional method [30] can be easily detected by the pooled test, especially when \( N/N_c > 0.5 \) (\( P_D > 88\% \) in such cases). However, the proposed method can decrease the \( P_D \) significantly. For a given \( N \), here \( N = 100, 200, 300 \), the decrements of \( P_D \) averaging over the three cameras with all ratio \( N/N_c \) are 10.0\%, 22.1\%, and 29.9\%, respectively. From Fig. 7 it is also observed that the \( P_D \) will increase when the ratio \( N/N_c \) ranging from 0 to 1. When \( N/N_c < 0.4 \) and \( N \geq 200 \), the average \( P_D \) is smaller than 53.4\% for the proposed method. Even when \( N/N_c = 1 \) (i.e., the set of the stolen images for Eve is exactly the same as the set of candidate images used in the pool test for Alice, and this case seems unlikely to happen in practice for detectors), the average \( P_D \) for the proposed method is still less than 88.3\% when \( N \geq 200 \), while the average \( P_D \) for the conventional method [30] is larger than 95.9\%.

### Table IV

| \( N \)  | 20   | 50   | 100  | 200  | 300  |
|---------|------|------|------|------|------|
| \( A = 50 \) | Method in [30] | 99.30 | 97.74 | 92.10 | 80.12 | 67.48 |
|         | Proposed \((r = 15)\) | 98.07 | 89.55 | 62.97 | 26.30 | 11.65 |
|         | Proposed \((r = 10)\) | 97.02 | 85.22 | 54.19 | 18.06 | 7.25 |
| \( A = 55 \) | Method in [30] | 97.20 | 98.72 | 63.80 | 47.89 | 27.58 |
|         | Proposed \((r = 15)\) | 92.67 | 50.47 | 18.95 | 3.17 | 1.20 |
|         | Proposed \((r = 10)\) | 77.33 | 39.65 | 12.03 | 1.98 | 0.88 |

Fig. 7. The average \( P_D \) of the pooled test for the conventional method [30] and the proposed method as a function of the ratio \( N/N_c \) on 100 images at \( P_{fa} = 10^{-3} \), where \( N \) is the number of images used by Eve to estimate the fake fingerprint, and \( N_c \) is the number of candidate images. The parameters are \( A = 50 \text{ dB} \) and \( r = 10 \).
In this paper, we first analyze the limitations of the conventional fingerprint-copy attack, and then propose an improved fingerprint-copy attack scheme via estimating the fake fingerprint from a randomly selected subset of the stolen images and superimposing it into the target image in a dispersive manner. The proposed scheme achieves a good tradeoff between the two requirements as described in Section III. The extensive experimental results evaluated on 2,900 images from 4 different cameras have shown that the proposed method can successfully confuse the camera identification (refer to Experiment #1) and attack the triangle test in three different forensic scenarios (refer to Experiments #2, #3 and #4 respectively). It significantly outperforms the conventional fingerprint-copy attack [30], especially when the number of stolen images is large enough, e.g. $N > 100$.

Though the proposed method is effective, it still has some limitations. Firstly, like most current anti-forensics works aiming at some certain forensic methods (such as [26], [27], and [31]), the proposed method is designed against one of the most popular camera identification methods [19] and the triangle test [30]. Thus it is difficult to attack other camera identification methods that are not based on PRNU. Secondly, the local textural information within the image has not been fully considered in the proposed scheme. It is promising that one can obtain better anti-forensic results by setting different fingerprint strengths for different image blocks according to the textural complexity. In the future, we will try to develop an adaptive way for adjusting the fingerprint strength to further improve the proposed scheme.

### V. Concluding Remarks

In this section, we will evaluate the proposed method for attacking the multiple forgeries triangle test. The number of the images stolen by Eve is set as $N \in \{20, 50, 100, 200, 300\}$, where the $N$ stolen images are belong to 600 images from camera $\mathcal{C}_A$. For each forgery $J$, the pdf $f_J(x)$ is estimated by $600-N$ images not used by Eve just as we did in experiment #2, and the candidate images in this case are the other 299 ($=300-1$) forgeries from the same camera generated with the same method and parameters. Thus the $P_D$ of $J$ means the probability of that the 299 forgeries are correctly detected by the multiple forgeries test. For each method with some given parameters, we perform multiple forgeries test on 100 forgeries from camera $\mathcal{C}_{E1}$, $\mathcal{C}_{E2}$ and $\mathcal{C}_{E3}$ respectively. The average $P_D$ are listed in the Table V.

| $N$ | 20 | 50 | 100 | 200 | 300 |
|-----|----|----|-----|-----|-----|
| $\mathcal{C}_{E1}$ | Method in [30] | 99.97 | 99.97 | 100 | 99.90 | 99.97 |
| $A = 50$ | Proposed ($r = 15$) | 99.73 | 93.61 | 76.19 | 54.98 | 42.11 |
| | Proposed ($r = 10$) | 98.36 | 86.15 | 63.11 | 39.13 | 24.82 |
| $A = 55$ | Method in [30] | 73.78 | 71.87 | 74.32 | 69.97 | 72.54 |
| | Proposed ($r = 15$) | 66.92 | 35.15 | 16.02 | 4.38 | 2.37 |
| | Proposed ($r = 10$) | 52.81 | 21.91 | 6.19 | 2.64 | 1.64 |
| $\mathcal{C}_{E2}$ | Method in [30] | 72.74 | 72.41 | 72.51 | 71.47 | 71.74 |
| $A = 50$ | Proposed ($r = 15$) | 68.43 | 51.40 | 45.52 | 44.45 | 44.21 |
| | Proposed ($r = 10$) | 60.00 | 46.92 | 44.35 | 42.51 | 44.05 |
| $A = 55$ | Method in [30] | 89.36 | 89.60 | 89.97 | 89.00 | 87.32 |
| | Proposed ($r = 15$) | 88.33 | 74.88 | 65.72 | 60.77 | 59.73 |
| | Proposed ($r = 10$) | 83.68 | 70.30 | 62.58 | 58.70 | 60.03 |

### E. Experiment #4: Attacking Multiple Forgeries Triangle Test

From Table V it is observed that the $P_D$ for the conventional method with $A = 50$ dB are larger than 98% even when $N = 300$. Overall, the proposed method can effectively degrade the $P_D$, especially when $N$ is large. Taking $A = 50$, $N = 300$ and $r = 10$ as an example, the average decrement for the proposed method is 52.3% comparing to the conventional method. It is also observed that the proposed method performs better on camera $\mathcal{C}_{E1}$ than camera $\mathcal{C}_{E2}$ and $\mathcal{C}_{E3}$. For example, when $\mathcal{C}_{E1}$ and $\mathcal{C}_{E3}$, the proposed method of camera $\mathcal{C}_{E1}$ is 6.19%, while the $P_D$ for camera $\mathcal{C}_{E2}$ and $\mathcal{C}_{E3}$ are 44.35% and 62.58%, respectively. The possible reason may be the effects of sensor noises and/or compression artifacts introduced by the digital cameras. Since camera $\mathcal{C}_{E1}$ and $\mathcal{C}_{E3}$ are with the same model, most hardware and software employed within the two cameras may be very similar or exactly the same. Therefore, it is expected that the non-PRNU components of their resulting images are relatively similar compared with those images from different brands and models of camera $\mathcal{C}_{E2}$ and $\mathcal{C}_{E3}$, thus the forgeries from camera $\mathcal{C}_{E2}$ and $\mathcal{C}_{E3}$ are easier to be detected. In our future research, we will further analyze the key factors inside different cameras that influence the detection performance of the multiple forgeries triangle test.

### REFERENCES

[1] M. Stamm, M. Wu, and K. Liu, “Information forensics: An overview of the first decade,” IEEE Access, vol. 1, pp. 167–200, 2013.

[2] H. Farid, “Exposing digital forgeries from JPEG ghosts,” IEEE Trans. Information Forensics and Security, vol. 4, no. 1, pp. 154–160, Mar. 2009.

[3] Z. Liu, J. He, X. Tang, and C.-K. Tang, “Fast, automatic and fine-grained tampered JPEG image detection via DCT coefficient analysis,” Pattern Recognition, vol. 42, no. 11, pp. 2492–2501, 2009.
[4] W. Luo, J. Huang, and G. Qiu, “JPEG error analysis and its applications to digital image forensics,” IEEE Trans. Information Forensics and Security, vol. 5, no. 3, pp. 480–491, Sep 2010.

[5] A. Popescu and H. Farid, “Exposing digital forgeries by detecting traces of resampling,” IEEE Trans. Signal Processing, vol. 53, no. 2, pp. 758–767, 2005.

[6] H.-D. Yuan, “Blind forensics of median filtering in digital images,” IEEE Trans. Information Forensics Security, vol. 6, no. 4, pp. 1335–1345, 2011.

[7] L. Li, J. Xue, Z. Tian, and N. Zheng, “Moment feature based forensic detection of resampled digital images,” in Proc. 21st ACM Int. Conf. on Multimedia, New York, NY, USA: ACM, 2013, pp. 569–572.

[8] Z. He, W. Lu, W. Sun, and J. Huang, “Digital image splicing detection based on markov features in DCT and DWT domain,” Pattern Recognition, vol. 45, no. 12, pp. 4292–4299, 2012.

[9] X. Qiu, H. Li, W. Luo, and J. Huang, “A universal image forensic strategy based on steganalytic model,” in Proc. 2nd ACM Workshop on Information Hiding and Multimedia Security, New York, NY, USA: ACM, 2014, pp. 165–170.

[10] Z. Liang, G. Yang, X. Ding, and L. Li, “An efficient forgery detection algorithm for object removal by exemplar-based image inpainting,” Journal of Visual Communication and Image Representation, vol. 30, pp. 75–85, 2015.

[11] S. Lyu and H. Farid, “How realistic is photorealistic?” IEEE Trans. Signal Processing, vol. 53, no. 2, pp. 845–850, 2005.

[12] A. Dirik, S. Bayram, H. Sencar, and N. Memon, “New features to identify computer generated images,” in Proc. IEEE Int. Conf. Image Processing, vol. 4, 2007, pp. IV – 433 – IV – 436.

[13] J. Lukáš, 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.

[14] N. Khamna, A. Mikkilineni, G. Chiu, J. Allebach, and E. Delp, “Scanner identification using sensor pattern noise,” in Proc. SPIE, Security, Steganography, and Watermarking of Multimedia Contents IX, vol. 6505, 2007, p. 65051K.

[15] A. Swaminathan, M. Wu, and K. Liu, “Nonintrusive component forensics of visual sensors using output images,” IEEE Trans. Information Forensics and Security, vol. 2, no. 1, pp. 91–106, Mar. 2007.

[16] O. Celiktutan, B. Sankur, and I. Avciibas, “Blind identification of source cell-phone model,” IEEE Trans. Information Forensics and Security, vol. 3, no. 3, pp. 553–566, 2008.

[17] K. Kurosawa, K. Kuroki, and N. Saitoh, “CCD fingerprint method - identification of a video camera from videotaped images,” in Proc. IEEE Int. Conf. Image Processing, vol. 3, 1999, pp. 537–540.

[18] M. Kharrazi, H. Sencar, and N. Memon, “Blind source camera identification,” in Proc. IEEE Int. Conf. Image Processing, vol. 1, Oct. 2004, pp. 709–712.

[19] M. Chen, J. Fridrich, M. Goljan, and J. Lukáš, “Determining image origin and integrity using sensor noise,” IEEE Trans. Information Forensics and Security, vol. 3, no. 1, pp. 74–90, 2008.

[20] C.-T. Li, “Source camera identification using enhanced sensor pattern noise,” IEEE Trans. Information Forensics and Security, vol. 5, no. 2, pp. 280–287, Jun. 2010.

[21] Y. Hu, C. Jian, and C.-T. Li, “Using improved imaging sensor pattern noise for source camera identification,” in Proc. IEEE Int. Conf. Multimedia and Expo, Jul. 2010, pp. 1481–1486.

[22] X. Kang, Y. Li, Z. Qu, and J. Huang, “Enhancing source camera identification performance with a camera reference phase sensor pattern noise,” IEEE Trans. Information Forensics Security, vol. 7, no. 2, pp. 393–402, Apr. 2012.

[23] S. Bayram, H. Sencar, and N. Memon, “Efficient sensor fingerprint matching through fingerprint binarization,” IEEE Trans. Information Forensics and Security, vol. 7, no. 4, pp. 1404–1413, 2012.

[24] F. Gharibi, F. Akhlaghian, J. RavanJamjah, and B. ZahirAzami, “Using the local information of image to identify the source camera,” in 2010 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 2010, pp. 515–519.

[25] R. Li, C.-T. Li, and Y. Guan, “A reference estimator based on composite sensor pattern noise for source device identification,” in Proc. IS&T/SPIE Conference on Media Watermarking, Security, and Forensics, vol. 9028, 2014, p. 90280O.

[26] M. Stamm and K. Liu, “Anti-forensics of digital image compression,” IEEE Trans. Information Forensics and Security, vol. 6, no. 3, pp. 1050–1065, Sep. 2011.

[27] T. Gloe, M. Kirchner, A. Winkler, and R. Böhme, “Can we trust digital image forensics?” in Proc. 15th ACM Int. Conf. on Multimedia, New York, NY, USA: ACM, 2007, pp. 78–86.

[28] R. Böhme and M. Kirchner, “Counter-forensics: Attacking image forensics,” in Digital Image Forensics, H. T. Sencar and N. Memon, Eds. Springer New York, 2013, pp. 327–366.

[29] A. Dirik and A. Karakciuk, “Forensic use of photo response non-uniformity of imaging sensors and a counter method,” Optics Express, vol. 22, no. 1, pp. 470–482, Jan. 2014.

[30] M. Goljan, J. Fridrich, and M. Chen, “Defending against fingerprint-copy attack in sensor-based camera identification,” IEEE Trans. Information Forensics and Security, vol. 6, no. 1, pp. 227–236, 2011.

[31] R. Caldelli, I. Amerini, and A. Novi, “An analysis on attacker actions in fingerprint-copy attack in source camera identification,” in Proc. IEEE Int.Workshop on Information Forensics and Security, Dec. 2011, pp. 1–6.

[32] M. Mihcak, 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, Mar. 1999, pp. 3253–3256.

[33] A. Popescu and H. Farid, “Exposing digital forgeries in color filter array interpolated images,” IEEE Trans. Signal Processing, vol. 53, no. 10, pp. 3948–3959, Oct. 2005.

[34] H. Farid, “Digital image ballistics from JPEG quantization,” Department of Computer Science, Dartmouth College, Tech. Rep., 2006.