Colour-Decoupled Photo Response Non-Uniformity for Digital Image Forensics

Chang-Tsun Li[1] and *Yue Li*[2]

[1]Department of Computer Science, University of Warwick, Coventry CV4 7AL, UK
c-t.li@warwick.ac.uk

[2]College of Software, Nankai University, Tianjin, China
liyue80@nankai.edu.cn

Abstract

The last few years have seen the use of photo response non-uniformity noise (PRNU), a unique fingerprint of imaging sensors, in various digital forensic applications such as source device identification, content integrity verification and authentication. However, the use of a colour filter array for capturing only one of the three colour components per pixel introduces colour interpolation noise, while the existing methods for extracting PRNU provide no effective means for addressing this issue. Because the *artificial* colours obtained through the colour interpolation process is not directly acquired from the scene by *physical* hardware, we expect that the PRNU extracted from the physical components, which are free from interpolation noise, should be more reliable than that from the artificial channels, which carry interpolation noise. Based on this assumption we propose a Couple-Decoupled PRNU (CD-PRNU) extraction method, which first decomposes each colour channel into 4 sub-images and then extracts the PRNU noise from each sub-image. The PRNU noise patterns of the sub-images are then assembled to get the CD-PRNU. This new method can prevent the interpolation noise from propagating into the physical components, thus improving the accuracy of device identification and image content integrity verification.

Keywords: Image forensics, colour filter array, photo response non-uniformity noise, demosaicing, image authentication, source device identification

I. Introduction

As digital multimedia processing hardware and software become more affordable and their functionalities become more versatile, their use in our everyday life becomes ubiquitous. However, while most of us enjoy the benefits these technologies have to offer, the very same set of technologies can also be exploited to manipulate contents for malicious purposes. Consequently, the credibility of digital multimedia when used as evidence in legal and security domains will be constantly challenged and has to be scientifically proved. After over 15 years of intensive research, digital watermarking [1, 2, 3, 4, 5, 6, 7] has been accepted as an effective way of verifying content integrity in a wide variety of applications and will continue to play an important role in multimedia protection and security. However, because of the need of proactive embedding of extra information in the host media, digital watermarking is only applicable when such a data embedding mechanism is available and the application standards/protocols are followed. Given this limitation it is of no doubt that unwatermarked multimedia will keep on being produced. Moreover, there has been an enormous number of existing unwatermarked media in circulation and it is in no way that digital watermarking can be of help in carrying out digital investigation when these pieces of unwatermarked multimedia are the objects in question. In the light of these issues, the recent years have seen an increasing number of publications on digital forensics [8, 9, 10, 11, 12, 13], which rely on extracting device signature left in the images/videos by acquisition devices [14, 15, 16, 17] to verify the credibility and identify the source of digital images/videos. A device signature may take the form of sensor pattern noise (SPN) [11, 14, 18, 19, 20], camera response function [21], re-sampling artefacts [22], colour filter array (CFA) interpolation artefacts [12, 16], JPEG compression [13, 23], lens aberration [24, 25], sensor dust [8], etc. Other device and image attributes such as binary similarity measures, image quality measures and higher order wavelet statistics have also been exploited to identify and classify source

devices [17, 26, 27].

A. Photo Response Non-Uniformity (PRNU)

Among so many types of intrinsic device signatures, sensor pattern noise [11, 14, 19, 20, 28] have drawn much attention due to its feasibility in identifying not only device models of the same make, but also individual devices of the same model [9, 11, 14]. Sensor pattern noise is mainly caused by imperfections during the manufacturing process of semiconductor wafers and slight variations in which individual sensor pixels convert light to electrical signal [29]. It is this uniqueness of manufacturing imperfections and non-uniformity of photo-electronic conversion that makes sensor pattern noise capable of identifying imaging sources to the accuracy of individual devices. The reader is referred to [29] for more details in relation to sensor pattern noise.

The dominating component of sensor pattern noise is photo response non-uniformity (PRNU) [14, 29]. However, the PRNU can be contaminated by various types of noise introduced at different stages of the image acquisition process. Figure 1 demonstrates the image acquisition process [30]. A colour photo is represented in three colour components (i.e., R, G, and B). For most digital cameras, during the image acquisition process, the lenses let through the rays of the three colour components of the scene, but for every pixel only the rays of one colour component is passed through the CFA and subsequently converted into electronic signals by the sensor. This colour filtering is determined by the CFA. After the conversion, a colour interpolation function generates the electronic signals of the other two colour components for every pixel according to the colour intensities of the neighbouring pixels. This colour interpolation process is commonly known as demosaicking [31, 32, 33]. The signals then undergo additional signal processing such as white balance, gamma correction and image enhancement. Finally, these signals are stored

in the camera's memory in a customised format, primarily the JPEG format.

In acquiring an image, the signal will inevitably be distorted when passing through each process and these distortions result in slight differences between the scene and the camera-captured image [14]. As formulated in [11], a camera output model can be expressed as

$$I = g^{\gamma}\left[(1+K)\Psi + \Lambda + \Theta_s + \Theta_r\right]^{\gamma} + \Theta_q \tag{1}$$

where I is the output image, and Ψ is the input signal of the scene, g is the colour channel gain, γ (= 0.455) is the gamma correction factor, K is the zero-mean multiplicative factor responsible for the PRNU, and Λ, Θ_s, Θ_r, Θ_q stand for dark current, shot noise, read-out noise and quantisation (lossy compression) noise, respectively. In Eq. (1), Θ_s and Θ_r are random noise and Λ is the fixed pattern noise (FPN) that is associated with every camera and can be removed by subtracting a dark frame from the image taken by the same camera [34]. Since Ψ is the dominating term in Eq. (1), after applying Taylor expansion to Eq. (1) and keeping the first two terms of the expansion,

$$I = I^{(0)} + \gamma \cdot I^{(0)} \cdot K + \Theta \tag{2}$$

where $I^{(0)}$ is the denoised image and Θ is the ensemble of the noises, including Λ, Θ_s, Θ_r and Θ_q. The PRNU pattern noise K can then be formulated as

$$K = \frac{W - \Theta}{\gamma \cdot I^{(0)}} \tag{3}$$

where

$$W = I - I^{(0)} \tag{4}$$

is the noise residual obtained by applying a denoising filter on image *I*. Although various denoising filters can be used, the wavelet-based denoising process (i.e., the discrete wavelet

transform followed by a Wiener filtering operation), as described in Appendix A of [14], has been reported as effective in producing good results.

B. Use of PRNU in Device Identification

The basic idea of using the PRNU noise pattern in device identification can be described as follows.

1) First, for each imaging device *d*, the noise residual patterns are extracted using Eq. (5) from a number of low-contrast images taken by device *d* and then the PRNU is estimated using the ML estimation procedure adopted by Chen *et. al*. [11], i.e.,

$$K_d = \frac{1}{\gamma} \frac{\sum_{s=1}^{S} W_{d,s} \cdot I_{d,s}}{\sum_{s=1}^{S} (I_{d,s})^2}, \tag{5}$$

where S is the number of images involved in the calculation, γ is the gamma correction factor ($\gamma \approx 0.455$), $I_{d,s}$ is the s-th image taken by device d and $W_{d,s}$ is the noise residual extracted from $I_{d,s}$. Note the multiplication operation in Eq. (5) is element-wise.

2) Secondly, the noise residual W_I of image I under investigation is extracted using Eq. (5) and compared against the reference PRNU K_d of each device d available to the investigator in the hope that it will match one of the reference fingerprints, thus identifying the source device that has taken the image under investigation [14]. The normalised cross-correlation

$$\rho(I \cdot K_d, W_I) = \frac{(I \cdot K_d - \overline{I \cdot K_d}) \cdot (W_I - \overline{W_I})}{\left\| I \cdot K_d - \overline{I \cdot K_d} \right\| \cdot \left\| W_I - \overline{W_I} \right\|} \tag{6}$$

is used to compare the noise W_I against the reference fingerprint K_d, where $\bar{\bullet}$ is the mean function. Note in Eq. (6), instead of using K_d, we used $I \cdot K_d$ as suggested in

[11]. Again the multiplication operation in Eq. (6) is element-wise.

Given the PRNU-based approaches' potential in resolving device identification problem to the accuracy at individual device level, it is important that the PRNU extracted is as close to the genuine pattern noise due to the sensor as possible. Since for most cameras, only one of the three colours of each pixel is physically captured by the sensor while the other two are artificially interpolated by the demosaicking process, this inevitably introduce noise with power stronger than that of the genuine PRNU. We can see from Eq. (2), (3) and (4) that the accuracy of both PRNU K and noise residual W depends on the denoising operation applied to I in obtaining $I^{(0)}$. However, as mentioned earlier that the most common method [11, 14, 15, 18] of obtaining $I^{(0)}$ is to apply the discrete wavelet transform followed by a Wiener filtering operation directly to the entire image I without differentiating physical components from artificial components and, as a result, allowing the interpolation noise in the artificial components to contaminate the real PRNU in the physical components. Addressing this shortcoming is the motivation of this work. In this work, we will look at the impact of demosaicking on PRNU fidelity in Section II and propose an improved formula for extracting PRNU in Section III. In Section IV, we present some experiments on device identification and image content integrity verification to validate the proposed PRNU extraction formula. Section V concludes this work. Because the PRNU is formulated in Eq. (3) and (5) as a function of the noise residual W (i.e., Eq. (4)), in the rest of the work we will use the two terms, PRNU and noise residual, interchangeably whenever there is no need to differentiate them.

II. Demosaicking Impact on PRNU Fidelity

In this work, we call the colour components physically captured by the sensor as *physical colours* and the ones artificially interpolated by the demosaicking function as *artificial*

colours. Due to the fact that demosaicking is a key deterministic process that affects the quality of colour images taken by many digital devices, demosaicking has been rigorously investigated [31, 32, 33, 35, 36]. Most demosaicking approaches group the missing colours before applying an interpolation function. The grouping process is usually content-dependent, e.g., edge-adaptive or non-adaptive, hence the accuracy of colour interpolation result is also content-dependent [37]. For example, in a homogeneous area, because of the low variation of the colour intensities of neighbouring pixels, the interpolation function can more accurately generate artificial components [30]. Conversely, in inhomogeneous areas, the colour variation between neighbouring pixels is greater, thus the interpolation noise is also more significant. This indicates that the PRNU in physical colour components is more reliable than that in the artificial components. However, the existing method for extracting PRNU as formulated in Eq. (4) and (5) based on the definition of the output image model in Eq. (1) does not take this into account [11]. To extract the PRNU using Eq. (4) and (5), the discrete wavelet transform followed by a Wiener filtering operation is applied. The main problem inherent to Eq. (4) is that it involves the whole image plane, which contains both artificial and physical components, in one noise residual extraction process. However, each coefficient of the wavelet transform used in the noise residual extraction process involves multiple pixels and thus both artificial and physical components. As a result the interpolation noise gets diffused from the artificial components into the physical ones. For example, in the red colour component/plane of an image taken by a camera with a Bayer CFA, only one fourth of the pixels' red colour are physical and for each pixel with physical red colour all its 8-neighbours' red colours are artificial. When wavelet transform is applied during the noise residual extraction process the interpolation noise residing in the artificial components propagates into the physical components. Therefore it is desirable to devise a noise residual extraction method that can prevent the artificial components from contaminating the reliable

PRNU residing in the physical components with the interpolation noise.

III. Formulation of Colour Decoupled PRNU (CD-PRNU)

In this section, we will discuss the formulation and extraction of CD-PRNU. First, a mathematical model for the CD-PRNU is derived and then an extraction algorithm is proposed to extract the noise residual that is to be used for estimating the final CD-PRNU, without prior knowledge about the CFA.

A. Mathematical Model of CD-PRNU

A generic demosaicking process is to convolve an interpolation matrix with an image block of the same size centred at the pixel where the artificial colour is to be calculated [10, 16, 31]. Although the 2×2 Bayer CFA is the most common CFA pattern, to make the proposed CD-PRNU versatile and applicable to cameras adopting different CFA patterns, we makes no assumption about the CFA pattern, F, except that it is a 2 × 2 square array. Let Ω be an interpolation matrix with $2N+1 \times 2N+1$ coefficients and $\Psi = \{\Psi_c \mid c \in \{R, G, B\}\}$ be a $X \times Y$-pixel input signal from the scene consisting of three colour components, R (red), G (green) and B (blue) before colour interpolation. That is to say that for each pixel $\Psi(x, y)$, only *one* of the three colour components takes a value *physically* captured by the sensor and this colour is determined by the colour configuration of the CFA pattern F. The other two colour components are to be determined by the demosaicking process. For each colour component of a pixel, $\Psi_c(x, y)$, $c \in \{R, G, B\}$, can be determined according to

$$\Psi_c(x,y) = \begin{cases} \Psi_c(x,y), & \text{if } F(x \bmod 2, y \bmod 2) = c \\ \sum_{u,v=-N}^{N} \Omega(u,v) \cdot \Psi_c(x+u, y+v), & \text{otherwise} \end{cases} . \qquad (7)$$

The first part of Eq. (7) means that if the colour component c is the same as the colour that the CFA pattern F allows to pass, i.e., $F(x \bmod 2, y \bmod 2) = c$, then no demosaicking is needed because c has been *physically* captured by the sensor. Otherwise, the second part of Eq. (7) is *artificially* applied to calculate the colour. According to Eq. (7), the image output model of Eq. (1) proposed in [11] can be re-formulated as

$$I = \begin{cases} g^{\gamma}\left[(1+K)\Psi + \Lambda + \Theta_s + \Theta_r\right]^{\gamma} + \Theta_q, & \text{if the colour is } \textit{physical} \\ g^{\gamma}\left[\Omega \cdot (1+K)\Psi + \Lambda + \Theta_s + \Theta_r\right]^{\gamma} + \Theta_q, & \text{if the colour is } \textit{artificial} \end{cases} \tag{8}$$

According to Eq. (3), we know that $\|K\| << 1$ because $\|W\| << I^{(0)}$. Therefore $(1+K) \approx 1$ and if we define the interpolation noise P as $\Omega = 1 + P$, the second part of Eq. (8) becomes $g^{\gamma}\left[(1+P)(1+K)\Psi + \Lambda + \Theta_s + \Theta_r\right]^{\gamma} + \Theta_q \approx g^{\gamma}\left[(1+P)\Psi + \Lambda + \Theta_s + \Theta_r\right]^{\gamma} + \Theta_q$. This is because, for the artificial components, the interpolation noise P is many orders greater than the PRNU K and $\|K\| << 1$, therefore $(1+P)(1+K)$ is virtually equal to $(1+P)$. As a result, Eq. (8) can be re-formulated as

$$I = \begin{cases} g^{\gamma}\left[(1+K)\Psi + \Lambda + \Theta_s + \Theta_r\right]^{\gamma} + \Theta_q, & \text{if the colour is } \textit{physical} \\ g^{\gamma}\left[(1+P)\Psi + \Lambda + \Theta_s + \Theta_r\right]^{\gamma} + \Theta_q, & \text{if the colour is } \textit{artificial} \end{cases} \tag{9}$$

Eq. (9) suggests that in the artificial components, the PRNU is actually the interpolation noise P while, in the physical components, the PRNU remains unaffected by the interpolation noise.

It can also be seen from Eq. (9) that the physical components and artificial components have similar mathematical expression. Hence if the physical and artificial colour components can be separated / decoupled, P can be extracted in the same way as the sensor pattern noise K is extracted (i.e., Eq. (3)). That is

$$P = \frac{W^a - \Theta}{\gamma \cdot I^{(0),a}} \tag{10}$$

where $I^{(0),a}$ is a low-passed filtered version of the artificial components I^a and W^a is the corresponding "sensor pattern noise", which is actually the interpolation noise. We can also use the same ML estimate as in Eq. (5) to extract the reference interpolation noise P_d for a particular device d from *S* low-variation images taken by *d* such that

$$P_d = \frac{1}{\gamma} \frac{\sum_{s=1}^{S} W_{d,s}^a \cdot I_{d,s}^a}{\sum_{s=1}^{S} (I_{d,s}^a)^2}, \tag{11}$$

where $I_{d,s}^a$ is the artificial colour components of the *s*-th low-contrast image taken by device d and $W_{d,s}^a$ is the interpolation noise extracted from $I_{d,s}^a$. We will discuss how the physical and artificial colour components can be decoupled in simple manner without *a priori* knowledge about the CFA pattern in Section III.B.

B. CD-PRNU Extraction Algorithm

According to Eq. (10) and (11), we can extract the sensor pattern noise and interpolation noise, respectively, from the physical and artificial components if the CFA is known. However, manufacturers usually do not provide information about the CFA used by their cameras [30]. Therefore, several methods have been proposed to estimate the CFA [10, 12, 16, 38]. Unfortunately, these methods have to exhaust all of the possible CFA patterns in order to infer/estimate the 'real'/optimal CFA. However, exhaustive search is by no means acceptable. In this work, to extract the CD-PRNU, we first separate the three colour channels I_c, $c \in \{R, G, B\}$ of a colour image I of $X \times Y$ pixels. Most CFA patterns are of 2 × 2 elements and are periodically mapped to the sensors. We know that, for each pixel of I, only one of the three colour components is physical and the other two are artificial, so the second step is, for each channel I_c, we perform a 2:1 down-sampling across both horizontal

and vertical dimensions to get four sub-images, $I_{c,i,j}$, $i,j \in \{0,1\}$, such that

$$I_{c,i,j}(x,y) = I_c(2x+i, 2y+j) \tag{12}$$

where $x \in [0, \lfloor X/2 \rfloor - 1]$ and $y \in [0, \lfloor Y/2 \rfloor - 1]$.

For each colour channel, I_c, without knowing the CFA pattern used by the manufacturer, we do not know (actually we do not have to know) which pixels carry the colour captured physically by the hardware and which are not. But by decomposing I_c into four sub-images, $I_{c,i,j}$, we know that each of the four sub-images either contains only the physical colour or only the artificial colours. By de-coupling the physical and artificial colour components in this fashion *before* extracting the noise residual, we can prevent the artificial components from contaminating the physical components during the DWT process. Eq. (4) is then used to obtain noise residual $W_{c,i,j}$ from each sub-images $I_{c,i,j}$, $\forall i,j \in \{0,1\}$. Finally the CD-PRNU W_c of each colour channel c is formed by combining the four sub-noise residuals $W_{c,i,j}$, $\forall i,j \in \{0,1\}$ such that

$$W_c(x,y) = W_{c,i,j}(\lfloor x/2 \rfloor, \lfloor y/2 \rfloor), \qquad i = x \bmod 2,\ j = y \bmod 2 \tag{13}$$

where $x \in [0, X-1]$, $y \in [0, Y-1]$ and mod is the modulo operation. The framework of the colour decoupled noise residual extraction process is shown in Figure 2 and the procedures are listed in Algorithm 1. Note that Algorithm 1 is for extracting the noise residual pattern *W* from an image *I*. To estimate the CD-PRNU P_d of a particular device *d* and use it as the reference signature of *d*, Eq. (11) is applied.

Algorithm 1. Noise residual extraction algorithm.

==========

Input: original image I

Output: colour decoupled noise residual W

Noise residual extraction algorithm

1) Decompose image I into R, G, and B components, I_R, I_G, and I_B.

2) $\forall c \in \{R,G,B\}$, decompose I_c into four sub-images, $I_{c,0,0}$, $I_{c,0,1}$, $I_{c,1,0}$ and $I_{c,1,1}$ by using Eq. (12).
3) $\forall c \in \{R,G,B\}$, extract $W_{c,0,0}$, $W_{c,0,1}$, $W_{c,1,0}$ and $W_{c,1,1}$ from $I_{c,0,0}$, $I_{c,0,1}$, $I_{c,1,0}$ and $I_{c,1,1}$ by using Eq. (4).
4) $\forall c \in \{R,G,B\}$, generate the colour decoupled noise residual W_c by combining $W_{c,0,0}$, $W_{c,0,1}$, $W_{c,1,0}$ and $W_{c,1,1}$ according to Eq. (13)
5) Combine the colour decoupled noise residual W_R, W_G, W_B to form the final noise residual W.

==========

IV. Experimental Results

In this section, we carry out experiments on source camera identification and image content integrity verification to validate the feasibility of the proposed CD-PRNU in a comparative manner.

A. Source Camera Identification

We have carried out source camera identification tests on 300 2048×1536-pixel photos of natural scenes taken by six cameras (C_1 to C_6), each responsible for 50. The six cameras are listed in Table 1. The reference PRNU (i.e., P_{C_i}, $i \in [1,6]$) of each camera C_i is generated by taking the weighted average of the PRNUs extracted from 30 photos of blue sky according to Eq. (11). For device identification purpose, we need clean PRNUs (which appear as high frequency bands of images) as device fingerprints for comparison against the PRNU extracted from individual images under investigation. The reason blue-sky images are chosen in this work is because blue sky contains less scene details (high frequency signal), thus giving better chance of extracting clean PRNU. Actually, other images with low-variation scenes (i.e., scenes without significant details) can be used instead. Taking the average of the PRNUs from 30 blue sky images is to further reduce variation. Our empirical experience suggests that an average of 20 blue sky images is accurate enough.

Source camera identification requires similarity comparisons among PRNUs (CD-PRNUs) and therefore the feasibility of the chosen similarity metrics is important. Fridrich suggested the use of the Peak to Correlation Energy (PCE) measure in [15], which has been proved to be a more stable detection statistics than normalised cross-correlation when applied to the scenarios in which the images of interest may have undergone geometrical manipulations, such as rotation or scaling. The purpose of this experiment is to demonstrate the capability of the proposed CD-PRNU in dealing with the colour interpolation noise, so geometrical transformations will not be applied in order to prevent biased evaluation from happening. Therefore, in the following experiments, normalised cross-correlation formulated as in Eq. (6) will be used to measure the similarity between PRNUs (CD-PRNUs).

In practice, the normalised cross-correlation has to be greater than a specified threshold for a camera to be identified as the source camera. However, in this experiment, the key point is about demonstrating the different performance of the traditional PRNU and the proposed CD-PRNU. Therefore, a camera is identified as the source camera, if out of the six reference PRNUs (or CD-PRNUs), its reference PRNU (or CD-PRNU) is most similar to the PRNU (or CD-PRNU), W_I, of the image I under investigation.

Because PRNU is often used in content integrity verification, where smaller image blocks have to be analysed, we also compare the performance of the proposed CD-PRNU against that of the traditional PRNU [11] when they are applied to blocks of 5 different sizes cropped from the centre of the full-sized PRNU (CD-PRNU). Table 2 lists the identification rates. Individually speaking, C_1, C_3, C_4, C_5 and C_6 perform significantly better when CD-PRNU is used in all cases, except for a few cases when images are of full size (1536 × 2048 pixels) and the identification rates are close or equal to 100% (1.0000). For C_2, PRNU performs equally well as CD-PRNU when the image size is 192 × 256 pixels

and slightly outperforms CD-PRNU when the block size is 48 × 64 pixels. We suspect that the reason C_2 does not perform as expected is because the CFA pattern is not a 2 × 2 square array as we have assumed. Another reason is that, because the smaller the images, the less data is available, therefore identification results become less reliable. Generally speaking, when the statistics of the six cameras are pooled together, as listed in the *Total* column of Table 2, we can see that CD-PRNU still outperforms PRNU significantly. This has been graphically presented in Figure 3(a). In Figure 3(b), a ROC curve of the performance of PRNU and CD-PRNU are demonstrated. We can see that the CD-PRNU outperforms the PRNU because at all fixed False Positive rate the CD-PRNU's True Positive rate are always higher than that of the PRNU.

For a system with a Pentium Core II 1.3G CPU and 3 GB RAM, it takes 0.526 seconds to compute the similarity between the PRNUs of two images of 2048 × 1536 pixels and 0.567 seconds to calculate the similarity between a pair of CD-PRNUs of the same size. The amount of data processed during the extraction of PRNU and CD-PRNU is the same. Although extracting CD-PRNU requires down-sampling and up-sampling, these two operations are trivial and only incur negligible increase of time complexity.

B. Content Integrity Verification

We also carried out the following three content integrity verification experiments on 640×480-pixel images.

- In the first experiment, we copied an 160×390-pixel area from Image $I.1$ in Figure 4(a), and pasted it at approximately the same location in Image $I.2$ in Figure 4(b) to create the forged Image $I.3$ as shown in Figure 4(c). The images in Figure 4(a) and (b) are taken by Olympus C730.
- In the second experiment, we cropped an 80×100-pixel area from Image $II.1$ in

Figure 5(a), which covers the face of the person, pasted it at the area where the face of another person is in Image *II*.2 in Figure 5(b) to create the forged Image *II*.3 in Figure 5(c). The images in Figure 5(a) and (b) are also taken by the same camera.

- In the third experiment, we cropped a 60×80-pixel area from Image *III*.1 in Figure 6(a) taken by Canon PowerShot A400, which covers the face of the person, pasted it at the area where the face of another person is in Image *III*.2 in Figure 6(b), which is taken by Olympus C730, to create the forged Image *III*.3 in Figure 6(c).

To detect the manipulated areas, we slid a 128×128-pixel window across the PRNU extracted from the image under investigation and another window of the same size across the reference PRNU of the cameras that have taken images *I*.2, *II*.2 and *III*.2. In Chen's method [11], the windows are moved a pixel at a time, which incurs a high computational load. Moreover, this method is not accurate at the pixel level [11]. Therefore, in our experiment, the sliding step/displacement is set to 5 pixels in order to reduce the computational load without sacrificing the accuracy of the integrity verification. Table 3 lists the number of manipulated and non-manipulated blocks of 5×5 pixels in the forged images.

To decide whether a block centred at the window superposed on the image has been manipulated or not, the cross-correlation of the PRNU patterns inside the two windows at the same location was calculated according to Eq. (6). If the cross-correlation is lower than a predetermined threshold *t*, the block in the centre of the window is deemed as manipulated. As discussed in [11], the cross-correlation follows the Generalised Gaussian (GG) distribution, therefore, we use various thresholds defined as $T(t)=\mu-t\cdot\sigma$ to analyse the performance of PRNU and CD-PRNU, where μ and σ are the mean and standard deviation of the correlations distribution, respectively, and $T(t)$ is the threshold. By varying the value of *t*, we can evaluate the integrity verification performance across a wide

range of correlation thresholds $T(t)$. In the following experiments we will allow t to vary independently in the range from 0.0 to 3.0 and use the four metrics, *true positive* (*TP*), *false positive* (*FP*), *true negative* (*TN*) and *false negative* (*FN*) to measure the performance of integrity verifications based on PRNU and CD-PRNU. As t grows, we will obtain lower *TP* and *FP,* while higher *TN* and FN. Let B be an arbitrary block and $M(B)$ and $M_d(B)$ be defined as

$$M(B) = \begin{cases} 1, \textit{if } B \textit{ is manipulated} \\ 0, \textit{otherwise} \end{cases} \tag{14}$$

$$M_d(B) = \begin{cases} 1, \textit{ if } B \textit{ is } \det\textit{ected as manipulated} \\ 0, \textit{otherwise} \end{cases} \tag{15}$$

TP, *FP*, *TN* and *FN* are defined as $TP = |\{B \mid M(B) = 1 \text{ and } M_d(B) = 1\}|$, $TN = |\{B \mid M(B) = 0 \text{ and } M_d(B) = 0\}|$, $FP = |\{B \mid M(B) = 0 \text{ and } M_d(B) = 1\}|$ and $FN = |\{B \mid M(B) = 1 \text{ and } M_d(B) = 0\}|$. Higher *TP* and *TN*, and lower *FP* and *FN* indicate better performance.

According to Chen's predication [11], "the block dimensions impose a lower bound on the size of tampered regions that our algorithm can identify. Thus, we remove all simply connected tampered regions from Z that contain less than 64×64 pixels (one quarter of the number of pixels in the block)". Chen applies erosion and dilation operations with a square 20×20 kernel in order to filter small areas identified as tampered with. The final authentication result is a image with the dilated areas highlighted as the tampered areas. However, the performance of the filtering / dilation operation strongly depends on parameter setting and hence many experiments must be run to obtain the best parameters for filtering. In order to simplify the comparison and to obtain a fair result, we use the raw data without any filtering to calculate the TP, TN, FP and FN. As a result, the experiments on *III*.3 demonstrate that CD-PRNU-based method significantly outperforms the PRNU-based method when the tampered area is about one quarter of the sliding window.

Experiment on Image I.3

Figure 7 shows the performance of the PRNU and CD-PRNU in terms of *TP*, *TN*, *FP* and *FN* when authentication is carried out on image *I*.3 across a range of correlation threshold *T*(*t*). We can see from Figure 7(a) and 7(b) that CD-PRNU generally achieves higher *TP* and *TN* while maintaining lower *FP* and *FN*. A lower correlation (similarity) allows the algorithm to detect more manipulated blocks, leading to higher *TP*. However, a low threshold also results in the situation where more authentic blocks are mistakenly detected as manipulated, giving rise to a higher *FP*. Therefore a ROC curve of *TP rate* with respect to *FP rate* can be used to evaluate the overall performance of the PRNU and CD-PRNU. Let α be the number of manipulated blocks and β be the number of authentic blocks, the ROC is formulated as

$$ROC = \frac{TP/\alpha}{FP/\beta} \tag{16}$$

At the same *false positive rate* (FP/β), which is marked along the horizontal axis of the ROC curve, an algorithm with better performance will have a higher *true positive rate* (TP/α), which is marked vertically. The ROC curves for the integrity verification experiments on image *I*.3 is illustrated as Figure 8. It is clear that the ROC curve of the PRNU-based scheme mostly overlaps with that of Random Guess, which means the authentication result is generally as unreliable as that of a random guess. This is because the area we copied from the source image *I*.1 is at approximately the same location as the original area in image *I*.2; therefore the PRNU pattern noises in the two areas are almost the same. As a result, the scheme cannot detect the manipulated area based on PRNU. By contrast, the CD-PRNU-based scheme results in a curve much higher than the PRNU-based method, which means that by using CD-PRNU manipulated blocks can be detected more reliably.

Experiment on Image II.3

When verifying the integrity of image *II*.3, CD-PRNU's consistently higher *TP* and lower *FN*, as shown in Figure 9(a) and 9(d), again indicate its superiority to PRNU. However, mixed performance in terms of *TN* and *FP* can be seen in Figure 9(b) and 9(c). Albeit their mixed performance in terms of *TN* and *FP*, both PRNU and CD-PRNU can effectively detect the manipulated blocks as their ROC curves have suggested in Figure 10. Figure 10 also shows that the ROC curve of CD-PRNU is still slightly higher than that of PRNU, indicating a slightly better performance of CD-PRNU.

Experiment on Image III.3

When authenticating *III*.3, although the performance of PRNU and CD-PRNU in terms of *TN* and *FP* are mixed, as can be seen in Figure 11(b) and 11(c), CD-PRNU's significantly better performance in terms of *TP* and lower *FN* can still be seen again in Figure 11(a) and 11(d), respectively. When the threshold t is higher than 1.1, the PRNU cannot correctly detect any manipulated blocks (i.e., $TP = 0$ as demonstrated in Figure 11(a)). This poor performance is also reflected in the PRNU's ROC curve in Figure 12 and is due to the fact that he manipulated area is too small (60×80 pixels), which is only about one quarter of the sliding window (128×128 pixels). Chen predicated in [11] that one quarter of the sliding window is the lower bound on the size of tampered regions that our algorithm can identify, and therefore areas smaller than this should be filtered in order to remove the falsely identified noise. The experiment result on III.3 conforms to Chen's observation. Since the tampered area is 60×80 pixels, approximately one quarter of the window, the method based on PRNU can perform no better than a random guess. By contrast, the manipulated blocks can be effectively detected by the CD-PRNU-based

scheme because the areas in question are from two images taken by different cameras and thus contain different interpolation noise. As a result, the CD-PRNU-based method can identify smaller areas.

V. Conclusions

In this work we have pointed out that the use of a colour filter array (CFA) in the image acquisition process can lead to inaccurate extraction of the PRNU, a commonly used *fingerprint* for identifying source imaging devices and image authentication. We have also proposed a simple, yet effective, colour-decoupled PRNU (CD-PRNU) extraction method, which can prevent the CFA interpolation error from diffusing from the *artificial* colour channels into the *physical channels*, thus improving the accuracy of the fingerprint. Moreover, the proposed method requires no *a priori* knowledge about the CFA colour configuration. Experiments on source camera identification and content integrity verification have been carried out to test our proposed CD-PRNU extraction method and significant improvement has been confirmed.

Acknowledgment: This work is partly supported by the EU FP7 Digital Image Video Forensics project (Grant Agreement No. 251677, Acronym: DIVeFor)

References

[1] Y. Yang, X.Sun, H. Yang, C.-T. Li and R. Xiao, "A Contrast-Sensitive Reversible Visible Image Watermarking Technique," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 19, no. 5, pp. 656 - 667, May 2009.

[2] *X. Zhao, A. T.S. Ho, Y. Q. Shi, "*Image Forensics Using Generalised Benford's Law For Improving Image Authentication Detection Rates in Semi-fragile Watermarking," *International Journal of Digital Crime and Forensics*, vol. 2, no. 2, Apr – Jun 2010.

[3] S. Chen and H. Leung, "Chaotic Watermarking for Video Authentication in Surveillance Applications," IEEE Transactions on Circuits and Systems for Video Technology, vol. 18, no. 5, pp. 704 – 709, 2008.

[4] Y. Zhao, P. Campisi and D. Kundur, “Dual Domain Watermarking for Authentication and Compression of Cultural Heritage Images,” *IEEE Transactions on Image Processing*, vol. 13, no. 3, pp. 430 – 448, 2004.

[5] D.P. Mukherjee, S. Maitra and S. T. Acton, “Spatial Domain Digital Watermarking of Multimedia Objects for Buyer Authentication,” *IEEE Transactions on Multimedia*, vol. 6, no. 1, pp. 1 – 15, 2004.

[6] F. Bao, R. H. Deng, B. C. Ooi and Y. Yang, “Tailored Reversible Watermarking Schemes for Authentication of Electronic Clinical Atlas,” IEEE Transactions on Information Technology in Biomedicine, vol. 9, no. 4, pp. 554 – 563, 2005.

[7] M. U. Celik, G. Sharma and A. M. Tekalp, “Lossless Watermarking for Image Authentication: A New Framework and an Implementation,” *IEEE Transactions on Image Processing*, vol. 15, no. 4, pp. 1042 – 1049, 2006.

[8] A. E. Dirik, H. T. Sencar, and N. Memon, “Digital Single Lens Reflex Camera Identification from Traces of Sensor Dust,” *IEEE Trans. Information Forensics Security*, vol. 3, no. 3, pp. 539–552, Sep. 2008.

[9] C.-T Li, "Source Camera Identification Using Enhanced Sensor Pattern Noise," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 2, June 2010.

[10] H. Cao and A. C.Kot, "Accurate Detection of Demosaicing Regularity for Digital Image Forensics," IEEE Transactions on Information Forensics and Security, vol. 4, no. 4 , Part: 2, pp. 899 - 910, 2009.

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

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

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

[14] J. Lukas, J. Fridrich, and M. Goljan, "Digital Camera Identification from Sensor Noise," *IEEE Transactions on Information Security and Forensics*, vol. 1, no. 2, pp. 205-214, 2006.

[15] J. Fridrich, "Digital image forensics," IEEE Signal Processing Magzine, vol. 26, no. 2, pp. 26-37, 2009.

[16] A. Swaminathan, M. Wu, and K. J. R. Liu, "Non-Intrusive Component Forensics of Visual Sensors Using Output Images," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 1, pp. 91-106, 2007.

[17] O. Celiktutan, B. Sankur, and I. Avcibas, “Blind Identification of Cell-Phone Model,” *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 3, pp. 553 - 566, September 2008.

[18] R. Caldelli, I. Amerini, F. Picchioni and A. De Rosa and F. Uccheddu, "Multimedia Forensic Techniques for Acquisition Device Identification and Digital Image Authentication," in *Handbook of Research on Computational Forensics, Digital Crime*

and Investigation: Methods and Solutions, C.-T. Li (Ed.), Hershey, PA: Information Science Reference (IGI Global), Nov. 2009.

[19] N. Khanna, A. K. Mikkilineni and Edward J. Delp, "Scanner Identification Using Feature-Based Processing and Analysis," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 1, pp. 123 – 139, March 2009

[20]H. R. Chennamma and L. Rangarajan "Source Camera Identification Based on Sensor Readout Noise," *International Journal of Digital Crime and Forensics*, vol. 2, no. 3, Jul – Sep 2010.

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

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

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

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

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

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

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

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

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

[30] K. *Nakamura, Image Sensors and Signal Processing for Digital Still Cameras*. Boca Raton, FL: Taylor & Francis Group, 2006.

[31] X. Li, B. Gunturk, and L. Zhang, "Image Demosaicing: A Systematic Survey," *Proc. SPIE*, vol. 6822, pp. 68221J–68221J-15, 2008.

[32] B. K. Gunturk, J. Glotzbach, Y. Altunbasak, R. W. Schafer, and R. M. Mersereau, "Demosaicking: Color Filter Array Interpolation," *IEEE Signal Processing Magazine*, vol. 22, no. 1, pp. 44 - 54, January 2005.

[33] R. Ramanath, W. E. Snyder, G. L. Bilbro, and W. A. Sander III, "Demosaicking Methods for Bayer Color Arrays," *Journal of Electronic Imaging*, vol. 11, no. 3, pp. 306-315, July 2002.

[34] M. Goesele, W. Heidrich, and H.-P. Seidel, "Entropy Based Dark Frame Subtraction," in

Proceedings of Image Processing, Image Quality, Image Capture Systems Conference, 2001, pp. 293-298.

[35] J. Adams, “Interaction between Color Plane Interpolation and Other Image Processing Functions in Electronic Photography,” *Proc. SPIE*, vol. 2416, pp. 144–151, 1995.

[36] J. F. Hamilton Jr. and J. E. Adams, “Adaptive Color Plane Interpolation in Single-Sensor Color Electronic Camera,” U.S. Patent 5 629 734, 1997.

[37] M. Sorell, "Unexpected Artifacts in a Digital Photograph," *International Journal of Digital Crime and Forensics*, vol. 1, no. 1, pp. 45-48, 2009.

[38] S. Bayram, H. T. Sencar, N. Memon, and I. Avcibas, "Source Camera Identification Based on CFA Interpolation," in *Proceedings of the IEEE International Conference on Image Processing*, 2005, pp. III-69 – III-72.

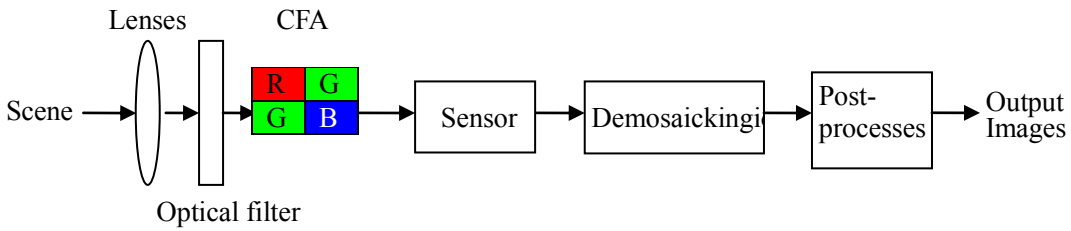


Figure 1. The image acquisition process of a digital camera.

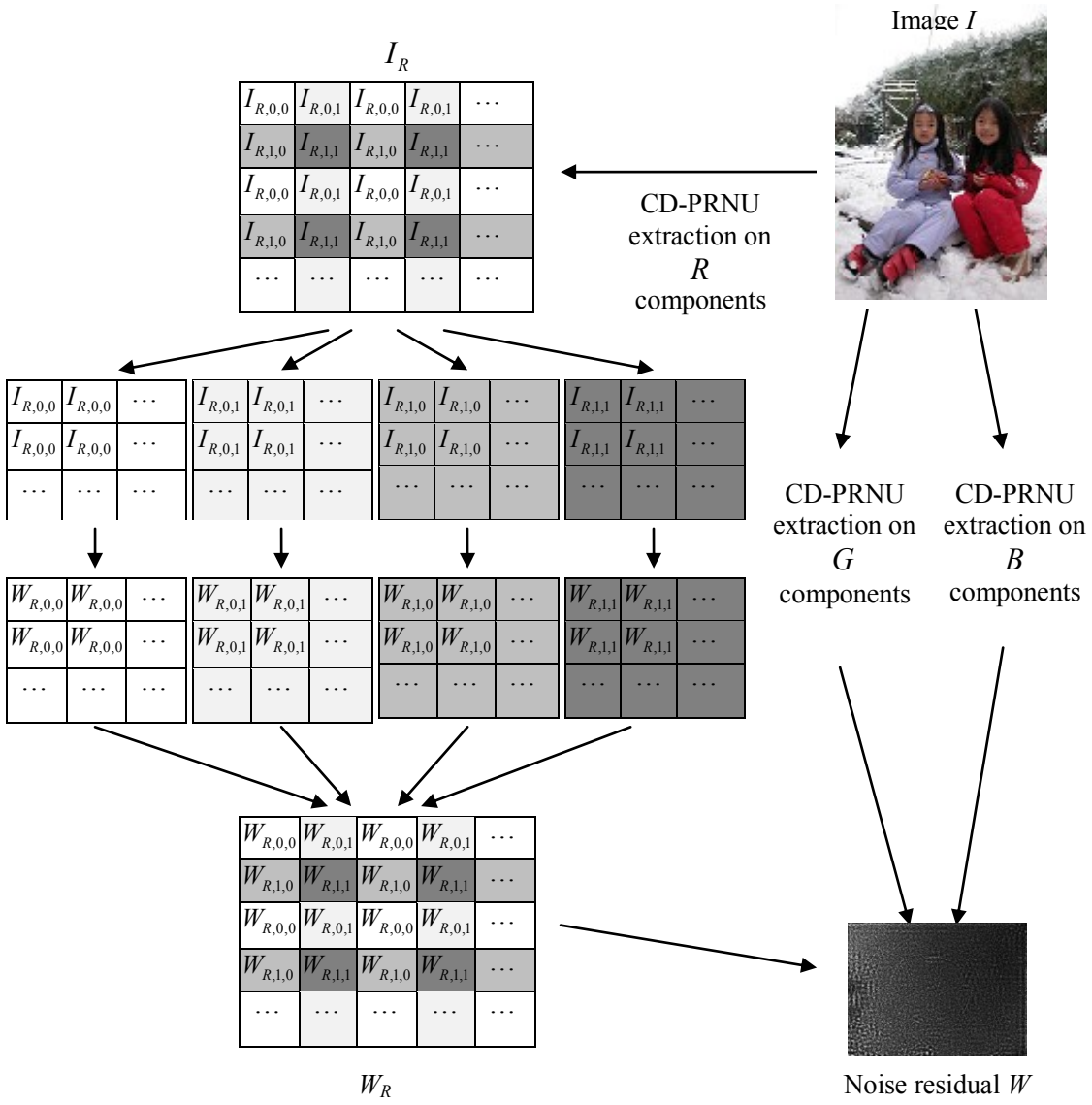
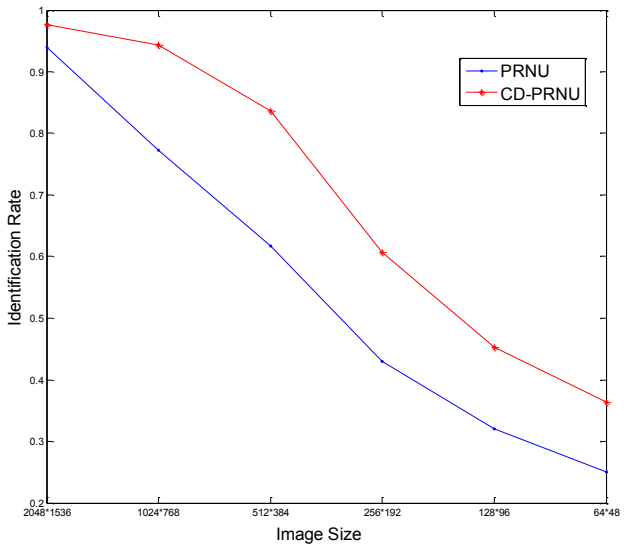
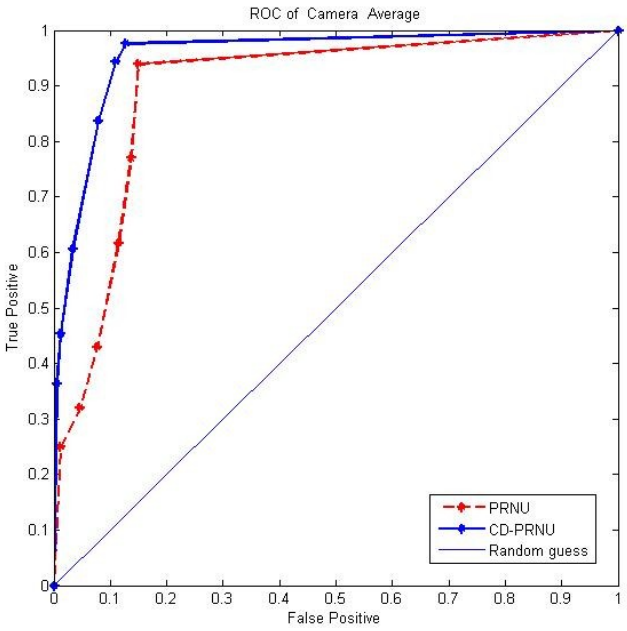


Figure 2. The noise residual extraction process.



(a)



(b)

Figure 3. Performance comparison of source camera identification a) Overall identification rates when CD-PRNU and PRNU are used as fingerprint; b) Overall ROC curve when CD-PRNU and PRNU are used as fingerprint.

(a) Original Image I.1 (b) Original Image I.2 (c) Forged Image I.3

Figure 4. The original image, source image and forged images for the content verification experiments.

(a) Original Image II.1 (b) Original Image II.2 (c) Forged Image II.3

Figure 5. The original image, source image and forged images for the content verification experiments.

(a) Original Image III.1 (b) Original Image III.2 (c) Forged Image III.3

Figure 6. The original image, source image and forged images for the content verification experiments.

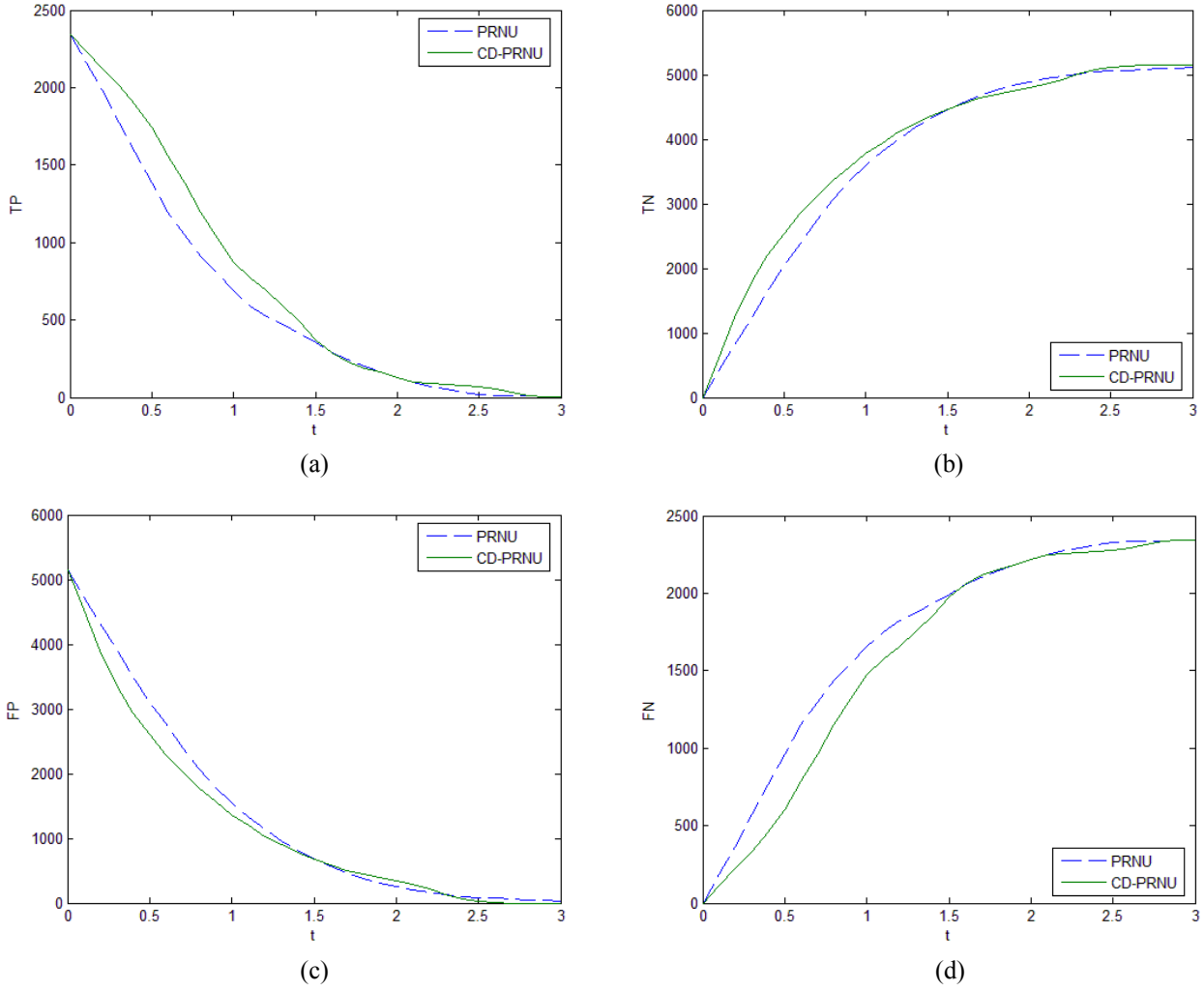


Figure 7. Authentication results on image *I*.3: Integrity verification performance of the PRNU and CD-PRNU in terms of a) TP, b) TN, c) FP and d) FN across a range of correlation threshold $T(t)$, with t varying from 0.0 to 3.0.

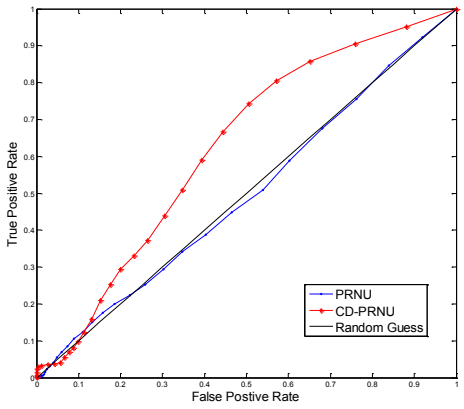


Figure 8. The ROC curve of Truth Positive Rate with respect to *False Positive Rate* of PRNU and CD-PRNU when authentication is performed on image *I*.3

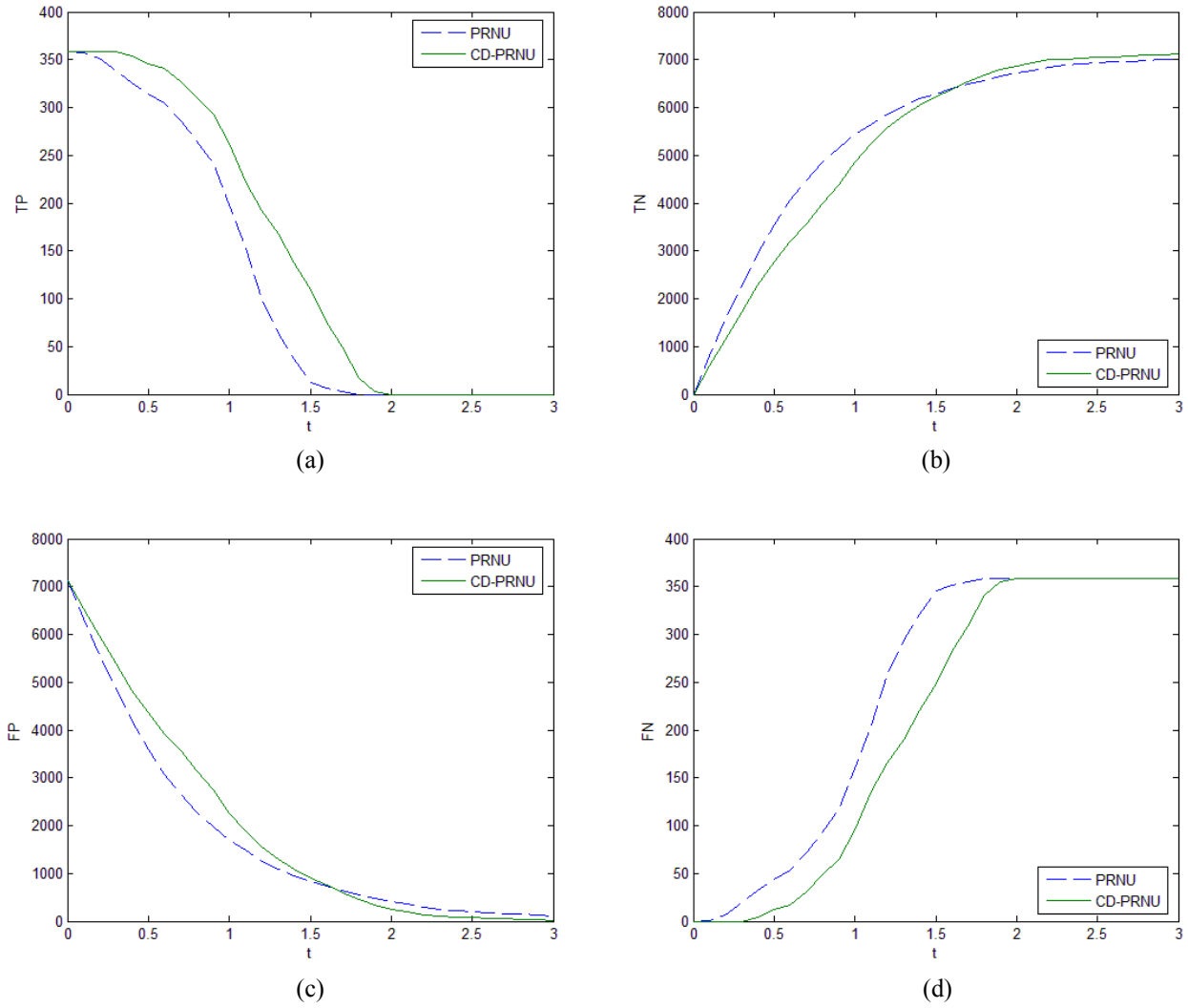


Figure 9. Authentication results on image *II*.3: Integrity verification performance of the PRNU and CD-PRNU in terms of a) TP, b) TN, c) FP and d) FN across a range of correlation threshold $T(t)$, with t varying from 0.0 to 3.0.

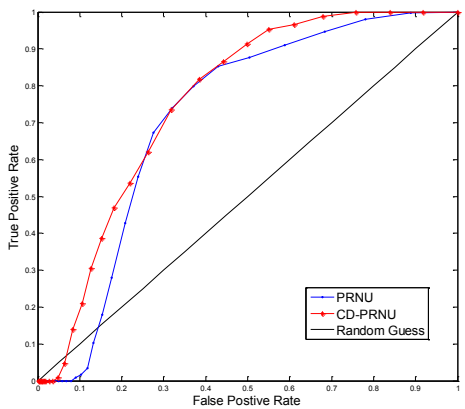


Figure 10. The ROC curve of Truth Positive Rate with respect to *False Positive Rate* of PRNU and CD-PRNU when authentication is performed on image *II*.3

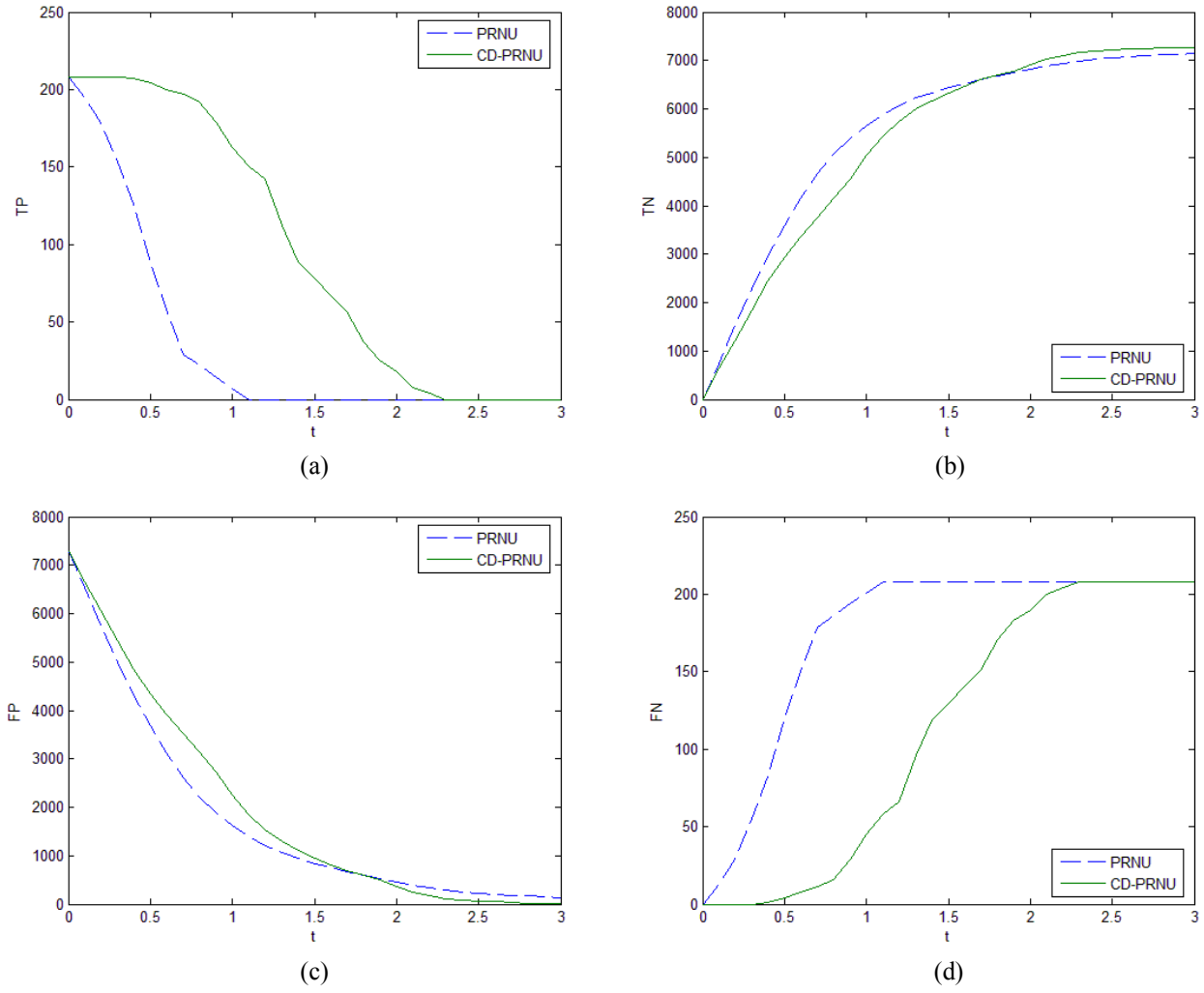


Figure 11. Authentication results on image *III*.3: Integrity verification performance of the PRNU and CD-PRNU in terms of a) TP, b) TN, c) FP and d) FN across a range of correlation threshold $T(t)$, with t varying from 0.0 to 3.0.

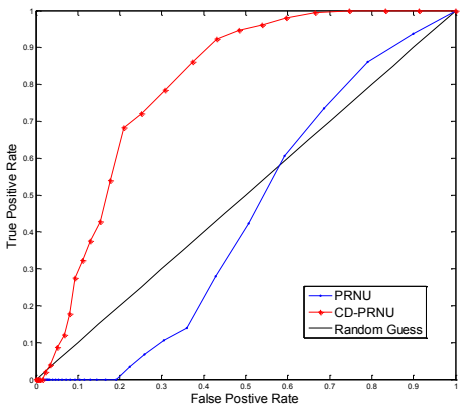


Figure 12. The ROC curve of Truth Positive Rate with respect to *False Positive Rate* of PRNU and CD-PRNU when authentication is performed on image *III*.3

Table 1. Cameras used in the experiments.

Symbol	Camera
C_1	Canon IXUS 850IS
C_2	Canon PowerShot A400
C_3	Canon IXY Digital 500
C_4	FujiFilm A602
C_5	Olympus FE210
C_6	Olympus C730

Table 2. Source camera identification rates using traditional PRNU and proposed CD-PRNU.

Block size	Methods	Identification rate of different cameras						
		C_1	C_2	C_3	C_4	C_5	C_6	***Total***
1536×2048	PRNU	0.9200	0.9600	0.9800	1.0000	0.9800	0.8000	**0.9400**
	CD-PRNU	0.9600	0.9600	0.9800	1.0000	1.0000	0.9600	**0.9767**
768 ×1024	PRNU	0.6800	0.8400	0.7200	1.0000	0.6200	0.7600	**0.7700**
	CD-PRNU	0.9400	0.9200	1.0000	1.0000	0.8200	0.9800	**0.9433**
384 × 512	PRNU	0.5000	0.7600	0.4600	0.9600	0.4200	0.6000	**0.6167**
	CD-PRNU	0.8400	0.8000	0.8400	0.9800	0.6800	0.8800	**0.8367**
192 × 256	PRNU	0.2200	0.6600	0.3200	0.7600	0.3000	0.3200	**0.4300**
	CD-PRNU	0.6000	0.6000	0.5800	0.8200	0.4600	0.5800	**0.6067**
96 ×128	PRNU	0.2600	0.4200	0.1600	0.5400	0.2200	0.3200	**0.3200**
	CD-PRNU	0.3000	0.4200	0.4800	0.6600	0.3200	0.5400	**0.4533**
48 × 64	PRNU	0.1400	0.4800	0.1600	0.3800	0.2000	0.1400	**0.2500**
	CD-PRNU	0.2400	0.4200	0.3000	0.6200	0.3600	0.2400	**0.3633**

Table 3. Number of manipulated and non-manipulated areas in each image (unit: block).

	Image I.3	Image II.3	Image III.3
Manipulated blocks	2346	358	208
Non-manipulated blocks	5142	7130	7280