Three Dimensional Denoising Filter For Effective Source Smartphone Video Identification and Verification

ASHREF LAWGALY*[1], FOUAD KHELIFI[1], AHMED BOURIDANE[2], SOMAYA AL-MAADDEED[3], YOUNES AKBARI[3]

[1]Department of Computer and Information Sciences, Northumbria University, Newcastle upon Tyne, UK
[2]Department of Computer Engineering, University of Sharjah, United Arab Emirates
[3]Department of Computer Science and Engineering, Qatar University, Doha, Qatar.

ABSTRACT

The field of digital image and video forensics has recently seen significant advances and has attracted attention from a growing number of researchers given the availability of imaging functionalities in most current multimedia devices at no cost including smartphones and tablets. Photo response non-uniformity (PRNU) noise is a sensor pattern noise characterizing the imaging device. However, estimating the PRNU from smartphone videos can be a challenging process because of the lossy compression that digital videos normally undergo for various reasons in addition to other non-unique noise components that interfere with the video data. This paper presents a new filtering technique for PRNU estimation based on the three-dimensional discrete wavelet transform followed by a 3D wiener filter. The rationale is that the 3D filter can filter out the compression artifacts along the temporal dimension in a more effective way than simple averaging. Experimental results on a public video dataset captured by various smartphone devices have shown a significant gain obtained with the proposed approach over the well-known two-dimensional wavelet-based Wiener approach.

CCS CONCEPTS

Security and privacy→Software and application security.

KEYWORDS

Photo Response Non-Uniformity Noise, Source Smartphone Identification, Video Forensics, 3D Denoising, 3D Wavelets & Wiener Filter.

1. INTRODUCTION

Nowadays, many organizations and individuals use digital devices in everyday life due to their undeniable advantages. A prime example of such device is smartphone, which includes a camera for taking good quality images /videos. As a result, many videos are commonly shared through the internet without applying any authentication system. This may cause serious problems, particularly in circumstances where the videos are an important component of the decision-making process for example, child pornography and movie piracy. Motivated by this, the present work investigates the performance of estimating Photo Response Non-Uniformity (PRNU) for smartphone videos and developing new denoising approach to improve the performance of digital source video identification. The PRNU noise is a sensor pattern noise characterizing the imaging device and it has been broadly used in the literature for image authentication and source camera identification. The abundant information that the PRNU carries in terms of the frequency content makes it unique, and therefore suitable for identifying the source camera and detecting forgeries in digital images. However, the PRNU estimation procedure is faced with the presence of image-dependent information as well as other non-unique noise components [1]. The field of image forensics is involved with image authentication, integrity verification and Source Camera Identification (SCI) by processing digital images[1]. Nevertheless, video forensics is involved with video recorder identification and video authentication using digital videos. Over the past decade, a large number of attempts to extract features which characterize the digital device using the Photo Response Non-Uniformity noise (PRNU) obtained from digital images was proposed. It is important to emphasize the fact that the PRNU characterise imperfections caused by the manufacturing process due to the lack of homogeneity of the silicon area in the imaging sensor

[2]. The unique noise caused by sensor imperfections is a weak signal of the same dimensions as the output of the image or video, indicated here by $K \in \mathcal{R}^{\mathcal{W}\times\mathcal{V}}$, where $\mathcal{W} \times \mathcal{V}$ represent the dimension of the sensor. Despite the fact that the sensor can be different from one device to another, the final digital image output can be expressed as [3],[4].

$$J = J^0 + J^0K + \Theta \quad (1)$$

Where J^0 denotes to the original video frame, J^0K is the PRNU term and Θ a random noise factor. In the literature, there has been a rising body of research based on PRNU to identify the source of digital images. A technique to estimate the PRNU-pattern was proposed by [3]. The residual signal r_i is calculated by denoising an image J_i using wavelet-based de-noising filter. Next the residual signal is obtained from an image J_i as $r_i = J_i - F(J_i)$ where the $F(J_i)$ is the de-denoised image. The PRNU, K, is the result of averaging N of the residual signal, where N refer to refers to the number of images used to estimate the PRNU. Another PRNU estimation method relying on Maximum Likelihood Estimator (MLE) for SCI was presented by [4]. In this system, the K is calculated as:

$$K = \frac{\sum_{i=1}^{N} r_i J_i}{\sum_{i=1}^{N} (J_i)^2} \quad (2)$$

In [5], an improved locally adaptive DCT Filter followed by a weighted averaging [6] to exploit the content of images carrying the PRNU efficiently was proposed . Although numerous of PRNU estimation techniques were developed for digital images [3-12] , less research has been conducted towards the forensic analysis of videos. Chen et al. [13] were the first authors to extend their PRNU technique [3] from an image to video and demonstrated that PRNU can be utilized in order to link a video to its source camcorder effectively. In this method, the PRNUs were estimated from both (training and query) video clips using MLE as shown in (2). Then, the peak-to-correlation energy (PCE) is utilized as measurement to detect the presence of PRNU. The key theory behind PCE is to take into account the correlation between the PRNU and shifted versions of the noise residue to lessen the similarity which may be present between the PRNU of a specific digital device and the noise residue of an image taken by a different camera. The PCE measure is defined in [4] and [13] as:

$$PCE(x,y) = \frac{C_{xy}^2(0,0)}{\frac{1}{\omega\times v - |A|}\sum_{m1,m2\notin A} C_{xy}^2(m1,m2)} \quad (3)$$

where A is a small neighbor area of size 11 × 11 around the central point at (0,0), $|A|$ is the number of pixels in A, and Cxy ($m1$, $m2$) represents the circular cross-correlation. Additional approach that considers only the non-textured frames in estimating the PRNU was proposed in [14]. This proposed technique applied several texture measures which was obtained from the Grey Level Cooccurrence Matrix (GLCM) to split the frames into the textured space and non-textured space using an unsupervised learning process. In [15] confidence weight PRNU based on image gradient magnitudes is proposed in order to improve PRNU estimation and evaluate the impact of video content on the performance of Chen et al.[13]. In [16] the video frames are resized to 512×512 using bilinear interpolation and the PRNU is extracted only from the green channel which is the noisiest channel among the RGB video. Existing video coding standards such as MPEG, H264 uses three types of video frames, which are intra-coded frame (I-frame), predictive coded frame (P-frame), and bi-predictive coded frame (B-frames) [15]. Chuang et al. [17] and Goljan et al. [18] analysed the video compression impact on PRUN estimation in the compressed domain. [17] reported that extracting the PRNU from I-frames is more reliable than P-frames and B-frames. In [19], a PRNU-based technique for out-of-camera stabilized videos, such as rotation and cropping processing was proposed. In this technique also 50 I-frames are extracted from each video to estimate the PRNU. A smartphone device may lead to rotate the video 180 degrees while recording videos with rolling 180 degrees. The authors In [20] are focused on effect of cameras rolling ,whether videos with several rolling degrees, 0, 90, 180, and 270 degrees, can affect the PRNU analysis or not. In [21], a hybrid methodology that utilizes both videos and still images were proposed to estimate the PRNU. In this technique, the PRNUs were estimated from still images obtained by the source device, while the query PRNU is estimated from the video and subsequently linked with the reference to verify the possible match. In [22] , the authors analysed some factors such as compression, resolution and length of the video, which could influence a decrease of the PRNU's correlation value in videos. In [23] , the minimum average correlation energy (MACE) filter [24] was introduced to reduce effect of heavily compressed in low-resolution videos. In this technique, the reference PRNU was extracted from number of videos, and then the MACE filter was applied for the reference PRNU to reduce impact of noises on normalized cross-correlation (NCC). Although there was previous research provided to enhance the PRNU estimation for source smartphone identification, an effective method that taken into consideration the frame content is still lacking.

Moreover, existing techniques that takes into an account the effect of lossy compression on the estimation of PRNU in the compressed domain requires full access to the right decoder in order to have separate I-frames at the estimation of the PRNU. This may not be useful in real scenario since a considerable number of video codecs applied in smartphones and released with different versions as standalone applications. This work addresses the problem of source smartphone video identification based on PRNU estimation at the filtering stage. The traditional approaches [13], [14], [15], [16] and [17] estimating the PRNU in digital videos using well-known two dimensional wavelet-based Wiener filter [25]. In the rest of the paper, this filter is referred to as the "2D-WWF". Compared to images, videos include highly temporal correlation between frames and the lossy compression artifacts can be very similar in the temporal dimension between adjacent frames[26]. Therefore, the 2D-WWF may not be an efficient approach for video denoising as it relies on simple averaging of the extracted noise residuals. In this work, a new filtering method based three-dimensional wavelet-based Wiener filter (3D-WWF) for efficient PRNU estimation is proposed. Experimental results on a video dataset [27], acquired by various smartphones, have shown a significant gain obtained with the proposed 3D-WWF over the 2D-WWF using different sizes of frames. The rest of this paper is structured as follows; section 2 describes the proposed method. Experimental results and analysis are provided in section 3. A conclusion is drawn in section 4.

2. PROPOSED PRNU ESTIMATION APPROACH

The justification behind the proposed 3D filter is that the PRNU is hard to be estimated in digital videos due to the lossy compression nature in which digital videos are stored, distortions that mainly occur in textured and edged regions. Because the imperfections in the silicon area of the imaging device remain unchanged while a video is being recorded, the video frames should contain an identical PRNU regardless of their different contents. However, because the compression artifacts are highly correlated across the consecutive frames [26] , they can't be removed effectively with the traditional technique by a simple averaging of the 2D noise residuals. In the 3D transform, the filter considers these artifacts as a 3D noise. Also, 3D DWT is used since the low-frequency and high frequency coefficients of the 3D DWT could be analysed to obtain more statistical properties of correlation among several successive frames [26]. These procedures lead to reduce efficiently the effect of compression in PRNU estimation. Figure. 1 demonstrates the proposed 3D-WWF for source smartphone video verification and identification system. First, to estimate the smartphone PRNU, videos are filtered using the proposed 3D-WWF. Then for each video, the residual signal is calculated by subtracting the original 3D video from the 3D denoised version. Next the PRNU is estimated using (3) and saved in a database to be used later for verification and identification. At matching stage, the PRNU for query video is estimated by applying the same process that has been used in PRNU extracting stage. For identification, the query video PRNU will be compared to all available PRNUs using the PCE similarity measure as shown in (3). The maximum PCE value is corresponding to the smartphone which has been used to record the video. In smartphone verification, however, the similarity between the PRNU and the query video PRNU of a certain smartphone is compared to a particular threshold in order to verify whether the video is recorded by that smartphone. The 3D filtering technique will be discussed in more detail in the next subsections.

2.1 Three-Dimensional Wavelet Transform

DWT has been applied in image compression standard JPEG2000 [26]. Also, DWT has played a significant role in the data analyzing and de-noising of MR images [28]. The DWT convert a finite energy signal to the scaled frequency domain. The orthogonal 1D DWT could be written as

$$X(t) = \sum_{k\epsilon Z} u_{j_{0,k}} \phi_{j_{0,k}}(t) + \sum_{j=\infty}^{j_0} \sum_{k\epsilon Z} w_{j,k}\, \psi_{j,k}(t), \qquad (4)$$

Where $\phi_{j_{0,k}}(t) = 2^j\, \phi(2^j - k)$ refers to the scaling, $\psi_{j,k}(t) = 2^{j/2}\, \psi(2^j t - k)$ is the wavelet function. Also the inner products $u_{j,k} = \langle X, \phi_{j,k} \rangle$ and $w_{j,k} = \langle X, \psi_{j,k} \rangle$ are the scaling and wavelet coefficients, respectively [29]. The 3D DWT is a rapidly developing research area, and it has been utilized in many fields such as seismology, biomedicine, material science, remote sensing, etc. [28]. The 3D DWT structure can be implemented as separate products of 1D DWT by applying 1D DWT in the x, y and z directions. The Video is firstly filtered along the x-direction which led to obtain sub-bands L and H. Next, filters are utilized along to y- direction, producing four decomposed sub-volumes: LL, LH, HL and HH. After that every sub-volume is filtered along the z-dimension, resulting in

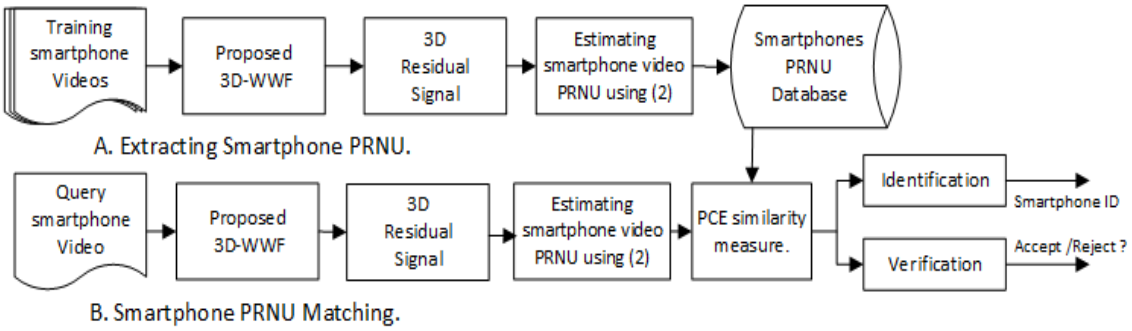

Figure 1. High-level of proposed source Video smartphone identification and verification system.

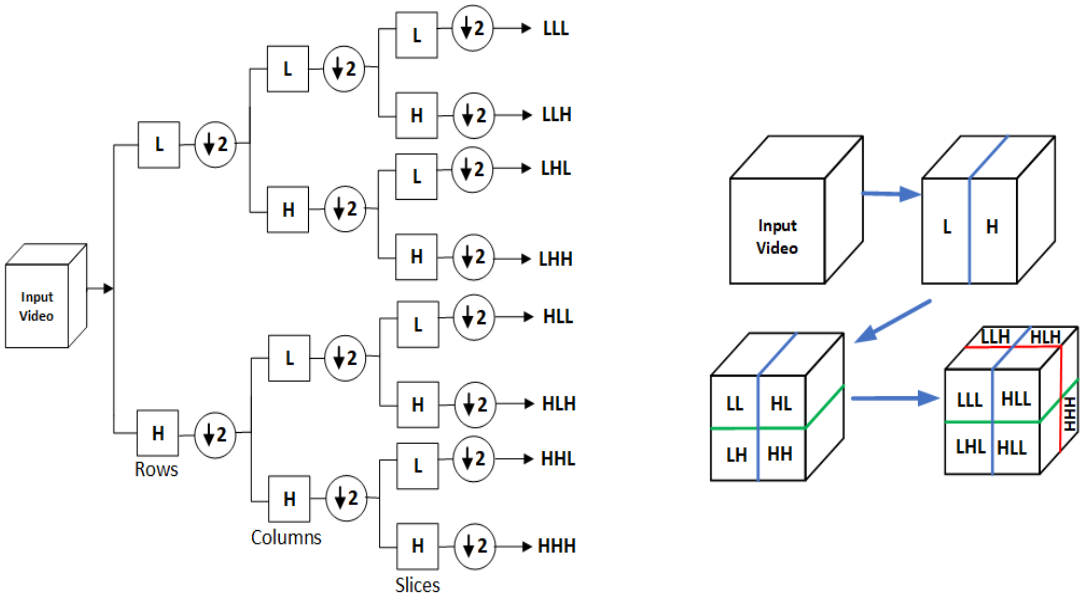

Figure 2. Separable 3D Wavelet decomposition.

eight sub-volumes: LLL, LLH, LHL, LHH, HLL, HLH, HHL and HHH as shown in figure 2.[28],[30]. The 3D DWT structure has advantages in analyzing changes of spatial and temporal information simultaneously [26]. Recently, the 3D DWT has been used in video watermarking, video coding and video denoising. Video and image denoising are crucial process as an initial stage in various recognition, analysis, and detection systems, which take visual input. Some Magnetic Resonance Imaging (MRI) and Computed Tomography (CT), utilize a set of consecutively captured images, which can be treated as 3D images, such images frequently including larger redundancy than single 2D images. The results show that filtering images as 3D data is more effective than using 2D filters over every sub-image being an element of them [31]. The authors in [29] reported that 3D wavelet improved the denoising results in comparison with 2D wavelet measured in Peak Signal to Noise Ratio (PSNR) and visually. The work [29] compering 2D and 3D version of three different filtering methods and the results have shown that the 3D versions always outperform the 2D ones based on the PSNR.

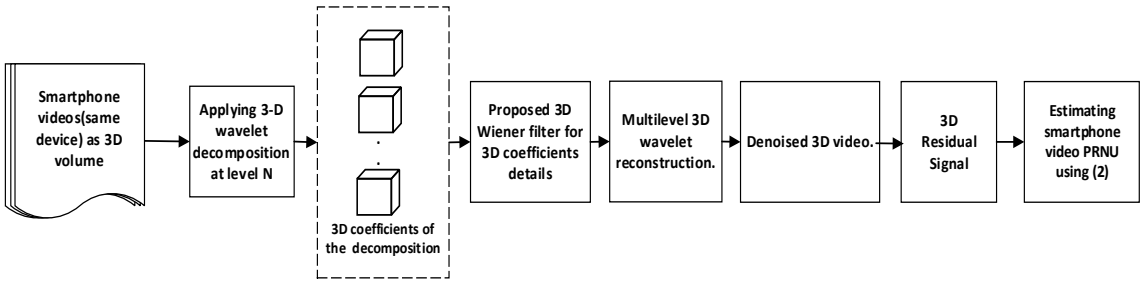

Figure 3. proposed 3D-WWF for Video smartphone identification and verification system.

2.2 Three-dimensional wavelet-based Wiener filter (3D-WWF) for PRNU estimation.

One of the commonly applied filters against image noise is the Wiener filter, which can be used for estimating uncontaminated signal by minimizing the mean square error between the estimated and the uncontaminated signal in a statistical sense[32]. Most of traditional methods [13], [14], [15], [16] and [17] applied the well-known two dimensional wavelet-based Wiener filter [25]. With aim to reduce the effect lossy compression on the PRNU estimation , this work use an extended version of [25] that treating video as 3D volume rather than applying 2D filter over each video frame. Figure 3 illustrates the main components of the proposed 3D filtering method. Firstly, a 3D video is decomposed at specific level in each direction (row, column, and slices). This process has been done by applying 1D DWT in each direction. At the first decomposition level, eight sub-bands are obtained (LLL, LLH, LHL, LHH, HLL, HLH, HHL and HHH $\in \mathcal{T}$), where $\mathcal{T}$ is the index set of the wavelet coefficients that depends on the decomposition level. It is worth mentioning that each video is decomposed with multiple levels which is equal to 4. The 3D fourth-level wavelet decomposition of the 3D video with the 8-tap Daubechies quadrature mirror filter is computed. Next, the 3D denoised wavelet coefficients are achieved utilizing the proposed 3D Wiener filter. This filter can be calculated for each sub-band as:

$$LLH_W(x,y,z) = LLH \quad (x,y,z)\frac{\hat{\sigma}^2(x,y,z)}{\hat{\sigma}^2(x,y,z)+\sigma_0^2} \quad (5)$$

Let the coordinates for each sub-band over the horizontals, verticals and slices be respectively denoted by x, y, z . σ_0^2 is the variance of white Gaussian noise (AWGN) and $\hat{\sigma}^2(x,y,z)$ refers to the estimated local variance of the 3D coefficients for the noise-free video. The maximum a posterior (MPA) estimation is applied to obtain the local variance:

$$\hat{\sigma}^2{}_q(x,y,z) = \max\left(0, \frac{1}{q^3}\sum_{(x,y,z)\in B_q} LLH^2(\mathrm{x,y,z}) - \sigma_0^2\right) (6)$$

Where $q \times q \times q$ is the size of the 3D mask B_q around (x,y,z). In [3] the authors was suggested to set $q \in \{3,5,7,9\}$. After that the minimum value of the four variances as shown in (7) is applied in (5).

$$\hat{\sigma}^2(x,y,z) = \min\left(\hat{\sigma}^2{}_3(x,y,z), \hat{\sigma}^2{}_5(x,y,z), \hat{\sigma}^2{}_7(x,y,z), \hat{\sigma}^2{}_9(x,y,z)\right) (7)$$

The value of σ_0^2 can slightly affect the performance of the filter in PRNU extraction and it has been suggested that the value of σ_0 to be between 2 and 5 [3]. The denoised 3D video is then obtained by applying 3-D wavelet reconstruction (inverse of wavelet transform) on the de-noised coefficients. Next, the 3D residual signal is calculated by subtracting the original 3D video from the 3D denoised version. Finally, the PRNU is estimated using (3). Once the PRNU is estimated for every single smartphone device, the above procedure is applied for the query video (See figure 1).

3. EXPERIMENTAL RESULTS

In this section, several experiments have been conducted to evaluate the performance of the proposed system. The evaluation has been conducted using the Video Authentication and Camera Identification Database (Video-ACID) [27]. Table 1 shows the twelve smartphones which have been used in our experiments. It is worth mentioning that this experimental contains videos from 7 different smartphones manufacturers, also, some of these videos are recorded by same brand device such as Motorola Moto E, Nokia 6.1 and Sony Xperia L1, letter A and B are used to differentiate between them. In this work, the proposed 3D wavelet Wiener filter (3D-WWF) is compared with the well-known wavelet-based Wiener filter(2D-WWF) [24] as has been used

in [13], [14], [15], [16] and [17]. Each PRNU for both estimated from 50 videos recorded by the same sensor, while the remaining videos are used in the testing stage. The estimation of PRNU has been performed by considering cropped blocks from the frame with different sizes, i.e., 512 × 512 and 720 × 720. The blocks are cropped from the centre of the full-size frame without affecting their content. For fair comparison, each smartphone PRNU for both approaches (the 2D-WWF Vs proposed 3D-WWF) is estimated using MLE as shown in (2) and the similarity between two PRNUs are calculated using PCE as shown in (3). Also, same parameters values have been used in both approaches for example the value of σ_0 and the decomposition level are set to be 3 and 4 respectively for both methods. The performance of the proposed filter is examined in two different aspects, i.e., source identification and source verification.

Table 1. Smartphones used in the experimental.

Smartphone name	Symbol	number of videos
Huawei Mate SE	M01	257
Kodak EKTRA	M02	239
LG X Charge	M03	234
Motorola Moto E (A)	M04	251
Motorola Moto E (B)	M05	227
Nokia 6.1 (A)	M06	234
Nokia 6.1 (B)	M07	242
Samsung Galaxy J7 Pro	M08	169
Samsung Galaxy S3	M09	230
Samsung Galaxy S5	M10	257
Sony Xperia L1 (A)	M11	233
Sony Xperia L1 (B)	M12	237

3.1 Source smartphone identification:

The objective of this scenario is to identify the smartphone used to record the video. Here, it is supposed that the video is recorded by one of the available smartphones. Accordingly, a query video is assigned to a specific smartphone if the corresponding PRNU provides the highest PCE. Table 2 demonstrates the false negative rate (FNR) for each smartphone using a frame size of 720×720. A clear improvement is demonstrated in most of smartphones for instance the FNR has been reduced from 62.02% to 3.80%and 36.36 % to 2.14 %. Furthermore, as shown in table 2 the proposed 3D-WWF lead a reduction in the overall FNR from 19.11% to 7.82%. Table 3 shows the false positive rate (FPR) for each smartphone using both approaches (2D-WWF vs. the proposed). As can be seen, a significant enhancement is obtained using the proposed 3D-WWF in ten smartphones out of twelve. Also, the proposed 3D-WWF lead to reduce the overall FPR to more than 50% less as shown in table 3. In addition to this, the accuracy for each smartphone is calculated as shown in (8) where TP, TN, FP, and FN are referred to the number of true positive, true negative, false positive and false negative respectively. The accuracy rate is defined as the proportion of videos which are correctly found to have been recorded or not to have been recorded by a given smartphone device. This contains the correct acceptance of videos recorded by the smartphone as well as the correct reject of recorded that were not acquired by the smartphone. As shown in table 4, the proposed technique leads to increase in the accuracy for most of the smartphones. For instance, in M11 and M12 the accuracy has been increased compared with 2D-WWF by about 3% regardless the frame size. Moreover, an increase in accuracy is shown in M03 of up to 5%. Additionally, with frame size 512×512 the overall accuracy is increased by more than 1% and by 2 % when the frame size is equal to 720×720. (See table 4).

$$Accurary = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

3.2 Source smartphone verification:

In this scenario, the task of the forensic analyst is to confirm whether a smartphone has been acquired video evidence by a particular threshold. This threshold is representing the least possible similarity among the reference PRNU of a smartphone and the PRNU of a video acquired by the same device. In this section, the Receiver Operating Characteristics (ROC) curve is utilised to examine the performance of the proposed filter. Twelve smartphones are used to compute the PCE values of similarity between each smartphone PRNU and the PRNU of video obtained by the same smartphone, On the other hand, the PCE values of similarity between every smartphone PRNU and the PRNU of videos recorded by different smartphones have been considered. Based on the previous procedures, the false positive rate and false negative rate for each threshold value is computed and then ROC curve is drawn. The ROC curve performance of the proposed 3D-WWF along with the traditional 2D-WWF is shown in figure 4 and figure 5. The ROC curve shows that the proposed method performs better than 2D-WWF approach. This is true for all frame sizes.

Table 2. FNR (%) for each smartphone using the traditional 2D-WWF and proposed 3D-WWF.

frame size	methods	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12	overall FNR
720×720	2D-WWF	28.02	3.17	69.02	17.41	13.56	4.89	4.17	0.84	7.78	18.36	25.68	36.36	19.11
	Proposed 3D-WWF	30.92	3.17	3.80	18.41	3.95	3.26	0.00	1.86	4.44	15.46	6.56	2.14	7.82

Table 3. FPR (%) for each smartphone using the traditional 2D-WWF and proposed 3D-WWF.

frame size	methods	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12	overall FPR
720×720	2D-WWF	0.35	4.60	0.35	1.10	1.52	1.14	1.04	1.63	0.89	3.89	6.76	0.79	2.00
	Proposed 3D-WWF	0.10	2.47	1.04	0.55	0.64	0.94	0.05	1.05	0.49	0.65	0.30	0.84	0.76

Table 4. Accuracy (%) for each smartphone using the traditional 2D-WWF and proposed 3D-WWF.

frame size	methods	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12	overall Accuracy
512×512	2D-WWF	97.47	96.56	93.89	98.42	98.14	99.14	99.82	98.64	98.96	97.01	92.49	94.98	97.13
	Proposed 3D-WWF	97.65	96.70	99.23	98.37	99.59	99.00	99.86	99.50	99.19	98.24	98.78	98.05	98.68
720×720	2D-WWF	97.06	95.52	93.94	97.42	97.51	98.55	98.69	98.42	98.55	94.75	91.67	96.20	96.52
	Proposed 3D-WWF	97.01	97.47	98.73	97.83	99.10	98.87	99.95	98.91	99.19	97.96	99.19	99.05	98.60

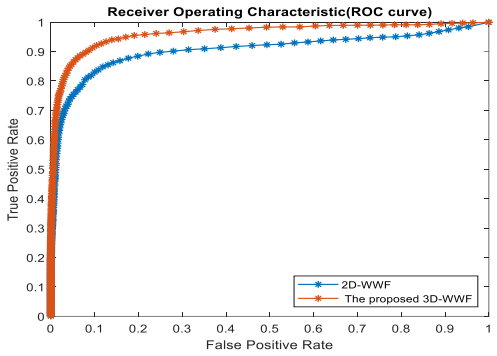

Figure 4. Overall ROC curve for 12 smartphones, frame size 512×512.

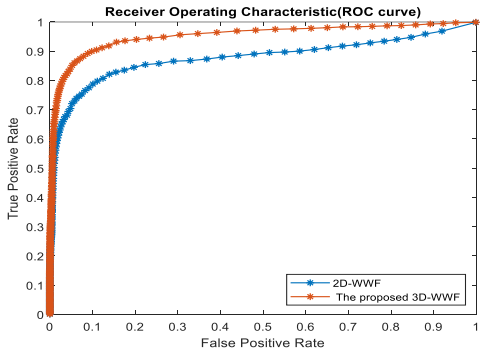

Figure 5. Overall ROC curve for 12 smartphones, frame size 720×720.

4. CONCLUSION

In this paper, an effective 3D filtering approach for source smartphone identification and verification has been proposed. In the traditional approach, the noise residual signals are extracted from video frames through the 2D wavelet Wiener filter and then averaged to estimate the PRNU. However, lossy video compression can lead to strong temporal correlation of artifacts between neighboring frames then averaged to estimate the PRNU. However, lossy video compression can lead to strong temporal correlation of artifacts between neighboring frames Therefore, the averaging process does not filter out the undesirable noise due to lossy compression artifacts across the temporal dimension. The rationale behind the proposed approach is that the denoising process should consider both temporal and spatial

dimensions to reduce the effect of lossy compression artifacts. This is achieved via a new 3D wavelet Wiener filter operating in the 3D wavelet domain (3D-WWF). An experimental evaluation covering two application scenarios in smartphone video forensics has shown the superiority of the proposed 3D-WWF over the traditional 2D-WWF

ACKNOWLEDGMENT

This work was supported by NPRP grant # NPRP12S-0312-190332 from the Qatar National Research Fund (a member of the Qatar Foundation). The statements made herein are solely the responsibility of the authors.

REFERENCES

[1] J. Lukás, J. Fridrich, and M. Goljan, "Digital" bullet scratches" for images," in *IEEE International Conference on Image Processing 2005*, 2005, vol. 3: IEEE, pp. III-65.

[2] J. R. Janesick, *Scientific charge-coupled devices*. SPIE press, 2001.

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

[4] M. Chen, J. Fridrich, M. Goljan, and J. Lukás, "Determining image origin and integrity using sensor noise," *IEEE Transactions on information forensics and security*, vol. 3, no. 1, pp. 74-90, 2008.

[5] A. Lawgaly and F. Khelifi, "Sensor pattern noise estimation based on improved locally adaptive DCT filtering and weighted averaging for source camera identification and verification," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 2, pp. 392-404, 2016.

[6] A. Lawgaly, F. Khelifi, and A. Bouridane, "Weighted averaging-based sensor pattern noise estimation for source camera identification," in *2014 IEEE International Conference on Image Processing (ICIP)*, 2014: IEEE, pp. 5357-5361.

[7] X. Kang, Y. Li, Z. Qu, and J. Huang, "Enhancing source camera identification performance with a camera reference phase sensor pattern noise," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 393-402, 2011.

[8] A. Lawgaly, F. Khelifi, and A. Bouridane, "Image sharpening for efficient source camera identification based on sensor pattern noise estimation," in *2013 Fourth International Conference on Emerging Security Technologies*, 2013: IEEE, pp. 113-116.

[9] M. Al-Ani, F. Khelifi, A. Lawgaly, and A. Bouridane, "A novel image filtering approach for sensor fingerprint estimation in source camera identification," in *2015 12th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 2015: IEEE, pp. 1-5.

[10] F. Ahmed, F. Khelifi, A. Lawgaly, and A. Bouridane, "Comparative analysis of a deep convolutional neural network for source camera identification," in *2019 IEEE 12th International Conference on Global Security, Safety and Sustainability (ICGS3)*, 2019: IEEE, pp. 1-6.

[11] M. H. Al Banna, M. A. Haider, M. J. Al Nahian, M. M. Islam, K. A. Taher, and M. S. Kaiser, "Camera model identification using deep CNN and transfer learning approach," in *2019 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST)*, 2019: IEEE, pp. 626-630.

[12] K. Bolouri, A. Azmoodeh, A. Dehghantanha, and M. Firouzmand, "Internet of things camera identification algorithm based on sensor pattern noise using color filter array and wavelet transform," in *Handbook of Big Data and IoT Security*: Springer, 2019, pp. 211-223.

[13] M. Chen, J. Fridrich, M. Goljan, and J. Lukáš, "Source digital camcorder identification using sensor photo response non-uniformity," in *Security, steganography, and watermarking of multimedia contents IX*, 2007, vol. 6505: International Society for Optics and Photonics, p. 65051G.

[14] A. Lawgaly, F. Khelifi, A. Bouridane, and S. Al-Maaddeed, "Sensor Pattern Noise Estimation using Non-textured Video Frames For Efficient Source Smartphone Identification and Verification," in *2021 International Conference on Computing, Electronics & Communications Engineering (iCCECE)*, 2021: IEEE, pp. 19-24.

[15] S. McCloskey, "Confidence weighting for sensor fingerprinting," in *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2008: IEEE, pp. 1-6.

[16] M. Al-Athamneh, F. Kurugollu, D. Crookes, and M. Farid, "Digital video source identification based on green-channel photo response non-uniformity (G-PRNU)," 2016.

[17] W.-H. Chuang, H. Su, and M. Wu, "Exploring compression effects for improved source camera identification using strongly compressed video," in *2011 18th IEEE International Conference on Image Processing*, 2011: IEEE, pp. 1953-1956.

[18] M. Goljan, M. Chen, P. Comesaña, and J. Fridrich, "Effect of compression on sensor-fingerprint based

camera identification," *Electronic Imaging*, vol. 2016, no. 8, pp. 1-10, 2016.

[19] S. Taspinar, M. Mohanty, and N. Memon, "Source camera attribution using stabilized video," in *2016 IEEE International Workshop on Information Forensics and Security (WIFS)*, 2016: IEEE, pp. 1-6.

[20] W.-C. Yang, J. Jiang, and C.-H. Chen, "A fast source camera identification and verification method based on PRNU analysis for use in video forensic investigations," *Multimedia Tools and Applications*, vol. 80, no. 5, pp. 6617-6638, 2021.

[21] M. Iuliani, M. Fontani, D. Shullani, and A. Piva, "Hybrid reference-based video source identification," *Sensors*, vol. 19, no. 3, p. 649, 2019.

[22] L. de Roos and Z. Geradts, "Factors that Influence PRNU-Based Camera-Identification via Videos," *Journal of Imaging*, vol. 7, no. 1, p. 8, 2021.

[23] D.-K. Hyun, C.-H. Choi, and H.-K. Lee, "Camcorder identification for heavily compressed low resolution videos," in *Computer Science and Convergence*: Springer, 2012, pp. 695-701.

[24] A. Mahalanobis, B. V. Kumar, and D. Casasent, "Minimum average correlation energy filters," *Applied Optics*, vol. 26, no. 17, pp. 3633-3640, 1987.

[25] M. K. Mihcak, I. Kozintsev, and K. Ramchandran, "Spatially adaptive statistical modeling of wavelet image coefficients and its application to denoising," in *1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No. 99CH36258)*, 1999, vol. 6: IEEE, pp. 3253-3256.

[26] Z. Li and G. Liu, "Video scene analysis in 3D wavelet transform domain," *Multimedia Tools and Applications*, vol. 56, no. 3, pp. 419-437, 2012.

[27] B. C. Hosler, X. Zhao, O. Mayer, C. Chen, J. A. Shackleford, and M. C. Stamm, "The video authentication and camera identification database: A new database for video forensics," *IEEE Access*, vol. 7, pp. 76937-76948, 2019.

[28] A. Procházka, L. Gráfová, O. Vyšata, and N. Caregroup, "Three-dimensional wavelet transform in multi-dimensional biomedical volume processing," in *Proc. of the IASTED International Conference on Graphics and Virtual Reality, Cambridge*, 2011, vol. 263, p. 268.

[29] B. Rasti, J. R. Sveinsson, M. O. Ulfarsson, and J. A. Benediktsson, "Hyperspectral image denoising using 3D wavelets," in *2012 IEEE International Geoscience and Remote Sensing Symposium*, 2012: IEEE, pp. 1349-1352.

[30] R. M. Jiang and D. Crookes, "FPGA implementation of 3D discrete wavelet transform for real-time medical imaging," in *2007 18th European Conference on Circuit Theory and Design*, 2007: IEEE, pp. 519-522.

[31] L. Petkova and I. Draganov, "Noise Adaptive Wiener Filtering of Images," in *2020 55th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST)*, 2020: IEEE, pp. 177-180.

[32] N. Eriksson-Bylund, M. Ressner, and H. Knutsson, "Reverberation Reduction Using 3D Wiener Filtering," in *SSBA 2003 Symposium on Image Analysis, Stockholm, Sweden, 6-7 mars 2003*, 2003.

[28] A. Procházka, L. Gráfová, O. Vyšata, and N. Caregroup, "Three-dimensional wavelet transform in multi-dimensional biomedical volume processing," in *Proc. of the IASTED International Conference on Graphics and Virtual Reality, Cambridge*, 2011, vol. 263, p. 268.

[29] B. Rasti, J. R. Sveinsson, M. O. Ulfarsson, and J. A. Benediktsson, "Hyperspectral image denoising using 3D wavelets," in *2012 IEEE International Geoscience and Remote Sensing Symposium*, 2012: IEEE, pp. 1349-1352.

[30] R. M. Jiang and D. Crookes, "FPGA implementation of 3D discrete wavelet transform for real-time medical imaging," in *2007 18th European Conference on Circuit Theory and Design*, 2007: IEEE, pp. 519-522.

[31] L. Petkova and I. Draganov, "Noise Adaptive Wiener Filtering of Images," in *2020 55th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST)*, 2020: IEEE, pp. 177-180.

[32] N. Eriksson-Bylund, M. Ressner, and H. Knutsson, "Reverberation Reduction Using 3D Wiener Filtering," in *SSBA 2003 Symposium on Image Analysis, Stockholm, Sweden, 6-7 mars 2003*, 2003.