HKIE THE HONG KONG INSTITUTION OF ENGINEERS 香港工程師學會

Performance enhancement of PRNU-based source identification for smart video surveillance

Sai Chung Law and Ngai Fong Law

Department of Electronic and Information Engineering, Faculty of Engineering, The Hong Kong Polytechnic University, Hong Kong, People's Republic of China

ABSTRACT

This paper introduces a current signal-based source verification (SSV) system for images in video surveillance networks including the cloud. Using a signal, the well-known photo-response non-uniformity (PRNU) can be used, which is unique and intrinsic in every digital image taken by a source camera, like fingerprints. The SSV system using PRNU has proved before to be useful for reliable video source identification in both network- and cloud-based video surveillance. However, in the era of smart living, security video systems have become part of the IoT devices which typically have limited resources such as low computation, power, storage and memory. To address these problems in the IoT applications, the effects of I-frames only and infra-red night scenes are studied as well as two proposed approaches for the SSV system. Then a hybrid version of the SSV scheme is further suggested, in combination with the best approach using averaged noise residues (for reduced false positive rate), and a recent technique using spatial domain averaged frames (for reduced computational complexity). The enhanced performance of the improved SSV system for smart video surveillance has been verified through tests.

KEYWORDS Photo-response non-uniformity (PRNU); source identification; IoT; cyber security; smart video surveillance
CONTACT Sai Chung Law s.c.law@connect.polyu.hk
Received 9 May 2022

1. Introduction

Video surveillance systems have been very popular for monitoring different areas within premises as well as other parts of a city. Security cameras are increasingly connected through network/Internet so that people can remotely access videos of live scenes via mobile apps or web-based applications. On one hand, such flexibility of access to such videos provides much convenience to people. On the other hand, the same system becomes more prone to cyber-attacks at the vulnerable points of networks. For example, the video data on reception or storage can be tampered with to forged data. The scenes displayed in a monitor may not be the actual scenes captured by the registered security cameras of the system (Farid, 2016; Redi et al., 2011). Serious security problems may occur, especially if the recorded scenes are for future reference or forensic purposes.

To validate the source of the received video, one may inspect the video to see if there are any visual discontinuities among consecutive frames. However, this method is unreliable and tedious to be performed. Alternatively, network-based methods are often used in existing systems for intrusion detection and video source verification. One can check if anyone has broken into a security system through the network and modified the video source. The sensor pattern noise called PRNU (Photo Response Non-Uniformity) has been used successfully for image source verification. Nevertheless, its use in video source verification in practical situations presents some challenges such as speed and reliability. With the video surveillance system becoming part of the Internet of Things (IoT), it is vital to study how the PRNU-based video source verification can be performed efficiently under limited resources in smart applications (Kitchin and Dodge, 2019).

The motivations of this research were (i) to extend the knowledge of digital image forensics techniques into the area of video application with comparable performance; and (ii) to address the problems of real-life applications of PRNU-based source identification schemes for smart video surveillance systems.

The objective of this paper is therefore to study the ways to enhance the performance of the PRNU-based video source identification. In particular, different challenges in extracting unique device characteristics and their robustness to various environmental factors such as day, night and infra-red scenes are investigated. The false positive rates and computational complexities of these methods are also studied. The rest of the paper is organised as follows. A background review is presented in Section 2. Section 3 explains the methods. Section 4 provides the experimental results. A discussion is given in Section 5. Finally, Section 6 concludes the paper.

2. Background

A typical video surveillance system consists of several components, such as a set of cameras, encoders/decoders, data storage devices, video management software, workstations and network equipment (Anthony, 2019). Popular video sizes for surveillance applications are

© 2022 The Hong Kong Institution of Engineers

1920x1080 (high definition), 1280x1080 and 720x1080. Security video formats generally employ H.264 and MPEG4. The following gives a brief review of a sensor pattern noise called PRNU (photo response non-uniformity) used to identify the source camera that produced an image or a video.

2.1. Review of PRNU

Sensor pattern noise is present in every image taken by a digital camera. This type of noise is intrinsic and unique in the image sensors (both CCD and CMOS) for a wide range of video and still-imaging devices, including smartphones, digital cameras, video cameras and even image scanners. The sensor pattern noise is present due to imperfections of pixels during manufacturing. PRNU (Photo Response Non-Uniformity) is a type of sensor pattern noise, so named because of the non-uniform sensitivity of pixels on exposure to uniform illumination. It is content-independent and exists in every digital image and video footage. Furthermore, the uniqueness of PRNU for different devices of even the same model and brand makes it useful as a sensor biometric for source identification and image or video forgery detection in the digital forensic context.

The image processing pipeline for digital video is broadly similar to that for digital still images, as shown in Figure 1. One significant difference between still and video images lies in the capture rate. Any video footage can be considered a sequence of still images captured and stored at a typical rate of 25 fps (frames per second) or 30 fps. In recent works, sensor pattern noise has proven to be useful and reliable for passive video forensic analysis (Pandey et al., 2016; Milani et al., 2012), especially for low-resolution videos from YouTube and smartphones.

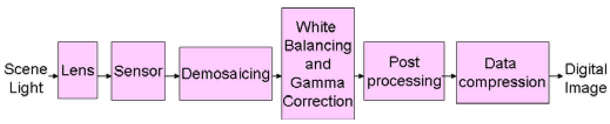


Figure 1. Schematic diagram of the image processing pipeline of digital cameras.

A basic image model (Lukas and Fridrich, 2006) can be expressed mathematically as,

$$Y = I + IK + N\,, \tag{1}$$

where Y denotes the output image taken by a camera, I is the original scene, K is the PRNU factor which is a multiplicative noise, and N models white Gaussian noise. By using the denoising operator, F, PRNU can be extracted from the residue noise, W, which is obtained by subtracting the denoised image, F(Y), from the output image, as indicated below in Equation (2).

$$W = Y - F(Y) = IK + N + I - F(Y)\,. \tag{2}$$

Hence, the residue noise, W, can be approximated as,

$$W = IK + \Xi\,, \tag{3}$$

where Ξ is the sum of all additive noise terms. The white Gaussian noise term, Ξ, is a random noise. To reduce the effect of Ξ, a maximum likelihood framework can be adopted to obtain the PRNU signal. Considering a group of d photos taken by the same camera, the sensor fingerprint or camera PRNU can be obtained by averaging. Hence,

$$K = \frac{\sum_{k=1}^{d} W_k I_k}{\sum_{k=1}^{d} {I_k}^2}\,, \tag{4}$$

where W_k and I_k are noise residue and image intensity of the k^{th} image taken by the camera, respectively. As PRNU is a type of multiplicative noise, the PRNU can be obtained with better results from a group of smooth images such as photos of a blue sky. On the other hand, PRNU information can be little or lost in dark photos as well as in saturated ones.

Using the camera PRNU K, the target photo or video frame can be tested to check if it comes from this digital camera or a video camera. The noise residue is extracted from the testing image through Equation (2). It is then matched with the camera PRNU through using correlation. The correlation is defined as,

$$corr(W_t, P_r) = \frac{(W_t - \overline{W}_t) \bullet (P_r - \overline{P}_r)}{\left\| W_t - \overline{W}_t \right\| \bullet \left\| P_r - \overline{P}_r \right\|}\,, \tag{5}$$

where t is the photo or video frame under test, W_t is the noise residue of t, and P_r is the camera PRNU obtained by using Equation (4). A large correlation value means that the noise residue of the test image matches the camera PRNU. Thus, it is highly probable that the testing image is taken by that camera device. In practice, a more stable detection statistic called the Peak to Correlation Energy (PCE) measure is used, rather than the correlation in Equation (5). PCE values over 30 can be used for source identification with a higher certainty.

The detection method can be formulated as a two-channel hypothesis testing, that is,

$$\begin{cases} H_0\text{: } t \text{ not taken by the reference camera} \\ H_1\text{: } t \text{ taken by the reference camera} \end{cases}. \tag{6}$$

Denoising is required in signal processing to separate the noise component from the original image scene data. Typically, Gaussian filters are used for denoising. Besides, wavelet filters and BM3D (Block-Matching and 3D) have been considered. In this work, BM3D filters (Dabov et al., 2007) were chosen based on their better performance in denoising operations over other state-of-the-art techniques.

2.2. Challenges of video source verification

While PRNU has much success in image source identification, its application in video source identification has various challenges. There are generally four different challenges.

2.2.1. Scene content effect

Scene content can seriously affect the accuracy of camera identification. Previous studies have shown that scene content can affect the quality of the extracted PRNU. Various research works have been performed to reduce the scene content effect (Chan et al., 2013; Shi et al., 2017).

There are mainly two factors related to this effect, namely (i) the image intensity and (ii) the image texture. The PRNU is a multiplicative factor in relation to intensity as shown in Equation (1). The PRNU characteristics can be detected easily with an increase in the image intensity. However, if the image intensity reaches the maximum level, or becomes saturated, the PRNU feature is lost completely in such a case. For instance, the sun image under a strong light intensity gives a zero noise residue, giving no sensor pattern noise for reliable use in forensic applications (Chan et al., 2013).

For texture/edge regions, the high-frequency components will be left in the noise residue after denoising, thus contaminating the extracted PRNU. Examples are details of a feature-rich image such as trees or building structures, which can be clearly left in the noise residue. The effect of the scene content on the correlation as in Equation (5) should be studied.

It follows that PRNU needs to be extracted from reliable regions of images (Shi et al., 2017). However, for video surveillance applications, image frames are highly correlated in scene content. Hence, certain regions would be unreliable for PRNU extraction in all frames collected from video surveillance applications. Different treatments are thus required in video surveillance applications, in contrast to those applied to digital still images.

2.2.2. Image compression effect

For enormous data transmission efficiency in a finite throughput channel, digital images including security videos need to undergo the image processing steps called quantisation and compression. Video images have significantly higher compression than still images. The effects of image compression on the extracted PRNU may vary with different frame types of GOP (Group of Pictures) of a video. In the tests, a loss-free and non-compressed image format such as BMP (Bit Map Pixel) converted from the original lossy MPEG video format was used. This helps prevent the second compression of images and reduces the adverse effects of compression on reliable PRNU extraction. However, this also leads to high computation complexity in operation.

2.2.3. Video stabilisation effect

Digital video cameras nowadays often have video stabilisation functions. This can, however, present challenges to PRNU extraction for camera source attribution. Misalignment between frames after the stabilisation can prevent a reliable estimation of PRNU fingerprints. Some recent works (Iuliani et al., 2019; Mandelli et al., 2020) have proposed solutions to counteract the effects of video stabilisation on source camera identification. This work focuses on video surveillance application. As the camera under testing is always in a still position, it apparently does not affect the performance of the proposed approaches.

2.2.4. Infra-red effect

Since video surveillance cameras operate 7x24 for security purposes, both day and night videos should be studied. Cameras can capture dark scenes even at 0 lux light intensity with the use of Infra-red (IR) lighting LEDs. Only light intensity components manifest in IR video images, while the colour component does not exist in these images. PRNU in colour images has been extensively researched, but not in IR images. Therefore, the effect of IR images taken by digital video surveillance cameras for tests is worth studying (Banerjee et al., 2018).

Apart from the above mentioned challenges, there is a recent non-uniqueness argument over PRNU in image-based source identification (Iuliani et al., 2021). In particular, it is suspected that some proprietary in-camera processing operations are similar among different cameras of both the same and other models. These similarities cause the same type of artefacts to be observed in the PRNU pattern, thus giving false device identification from images. However, as far as is known, no similar research work has been conducted in the area of video-based source identification.

3. Video source verification methods

In video source verification, the main concerns are the computational complexity and the false positive rate. There are two existing works addressing these issues. They are the SSV system (Law SC and Law NF, 2018) as shown in Figure 2, and the spatial domain averaged (SDA) frames technique (Taspinar et al., 2020) as shown in Figure 4.

The SSV system contains two parts: extraction of camera characteristics and source verification. In the first part, the security camera is registered, and its characteristics are extracted through analysing video signals taken by the camera. This step is needed so that features about the security camera can be kept in the SSV system for comparison purposes in later stage. In the second part, the live video signal is analysed to ascertain whether the video contains the particular features of the registered security camera. If the features are missing, it is likely that the video scene may not have come from the registered security camera. In this case, a warning can be issued.

The source verification is based on one frame in which its noise residue (test PRNU) is matched with the camera signature (reference PRNU). The quality of the noise residue is thus important. Otherwise, this may give rise to a high false positive rate. A good result can be obtained if the scene contains primarily smooth content. However, in video monitoring, it is possible that various types of scenes are being monitored. Thus, methods must be used to reduce the false positive rate. Figure 3 shows the diagrams of the two proposed approaches for improving the SSV system.

The first approach, SSV_1, as shown in Figure 3(a), requires consecutive M video frames to have low PCE values (less than a threshold *T*) before a warning is issued. In a typical GOP video structure with I, B and P frames, B and P frames depend on I frames for decoding. Thus, the accuracy of B and P frames may be lower than that of I frames in source verification due to the use of motion estimation and error coding. Hence SSV_1 requires a consecutive number of frames having low PCE values to trigger a warning.

The second approach, SSV_2, as shown in Figure 3(b), attempts to improve the quality of the noise residues. Rather than using one noise residue signal from one frame, noise residues are extracted from a group of frames instead. These noise residues are then averaged to produce the resultant noise residue for the source verification against the camera signature afterward.

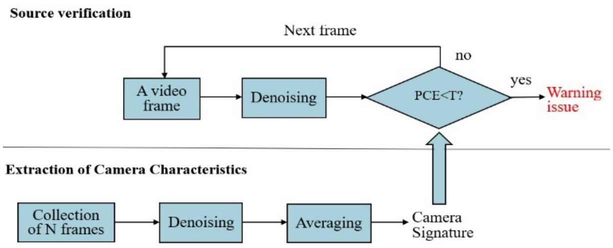
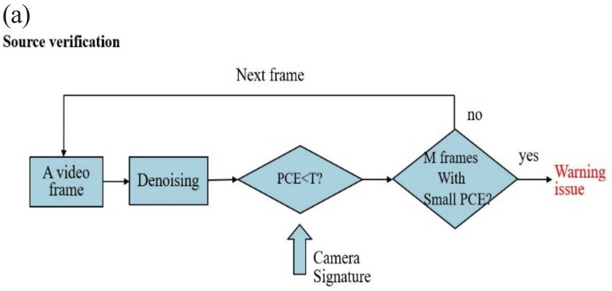


Figure 2. Existing signal-based source verification (SSV) system (Law SC and Law NF, 2018).

(a)



(b)

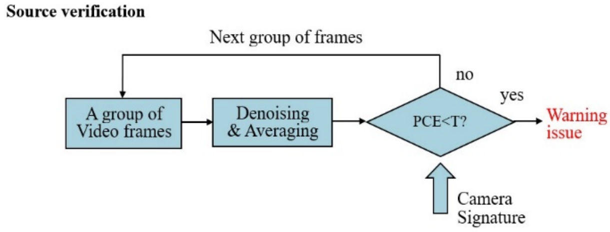


Figure 3. Two proposed approaches to improve the performance, (a) SSV_1; and (b) SSV_2.

Another technique used to reduce the computational complexity is to estimate a camera fingerprint via spatial domain averaged frames. Figure 4 shows the major steps of the SDA method. In this approach, the video frames are divided into subgroups in which frame averaging of each subgroup is performed to give SDA frames. Each SDA frame is then denoised and finally the camera fingerprint is obtained from these denoised frames. As denoising is performed on the SDA frames, the total number of denoising operations is reduced to save computation time.

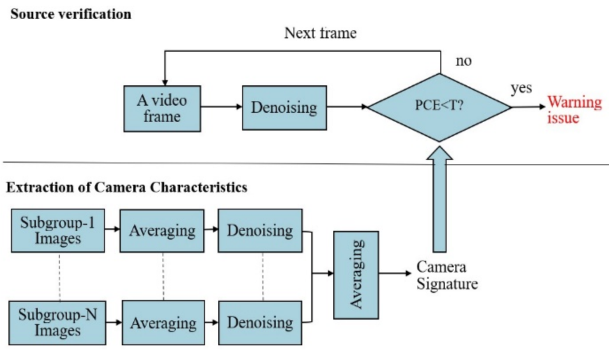


Figure 4. Steps of the SDA method using spatial domain averaged frames for fingerprint extraction.

4. Robustness under different scenarios

The data for study were collected from video images taken by two IP network CCTV cameras, namely the Wansview 630KC(White) and Wansview 630KC(Black). The two cameras are of the same model but only distinguishable based on their exterior body colour, White or Black. Video frames of these devices can be merged to form combined video sequences that contain different scenes for tests. The resolutions of the cameras can be either 480p or 720p. In general, 1280x720p is adequate for home security purposes, so mainly this resolution was chosen in this work. Matlab codes were used for the simulation tests and analysis of the results.

A CCTV camera system usually operates 24 hours a day. Therefore it is necessary to check its robustness under different scenarios, say, day and night. In summary, experiments and tests were conducted regarding the effects of infra-red night vision (Figure 5), normal daylight scene (Figure 6) and I-frames, as well as the computational complexities and false positive issues. The results and discussion can be found in the following.

Figure 5. The infra-red scene for tests.

Figure 6. A normal daylight scene for tests.

4.1. Study of a normal daylight scene

The performance of the SSV and SDA methods in regard to sunny daylight scene videos was tested (Figure 6). For the SDA method, each subgroup contained at least 30 non-interlacing images with a total of seven sub-groups. Seven SDA frames were then formed to obtain the PRNU fingerprint. Figure 7 and Figure 8 show the results. It can be seen that high PCE values in the range of 230 to 300 were obtained if the tested video signal came from the same camera as the reference PRNU. However, PCE values in between 0.5 and 1.2 were obtained if the tested video signal did not come from the same camera as the reference PRNU. For clarity, Figure 9 shows the combined video sequence from Figure 7 and Figure 8. A clear boundary and high contrast can be seen between camera matching and non-matching cases, with high PCE values for matching and low PCE values for non-matching. In terms of system performance, the more significant the difference of PCE values between matching and non-matching frames in the tests of an approach, the more accurate the approach. Similar results were obtained for the SSV method. Hence, it verified that both the SSV and SDA methods are effective for normal daylight scene video source verification.

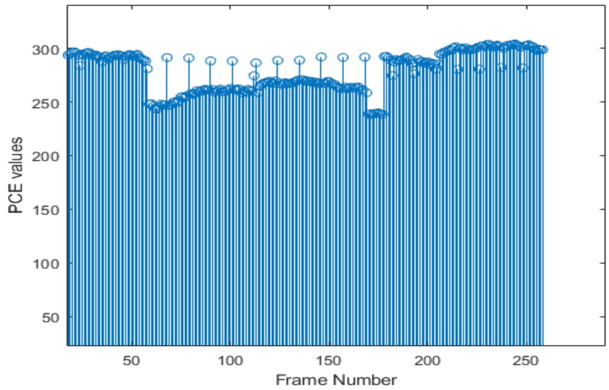


Figure 7. Plot of PCE value results for all frames of a sunny day scene taken by the Wansview (White), tested *vs.* Wansview (White) reference PRNU obtained by the SDA-frames approach.

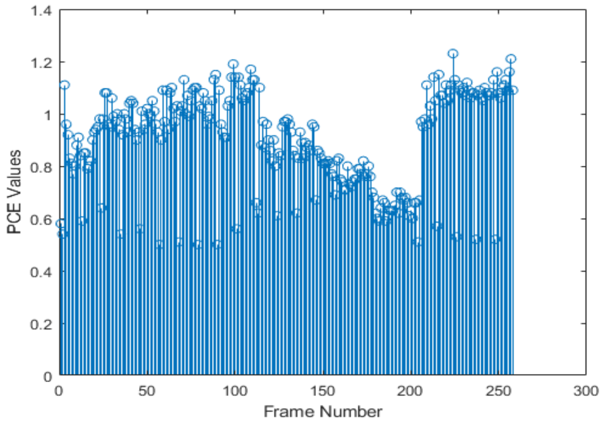


Figure 8. Plot of PCE value results for all frames of a sunny day scene taken by the Wansview (White), tested *vs.* Wansview (Black) camera reference PRNU obtained by the SDA-frames approach.

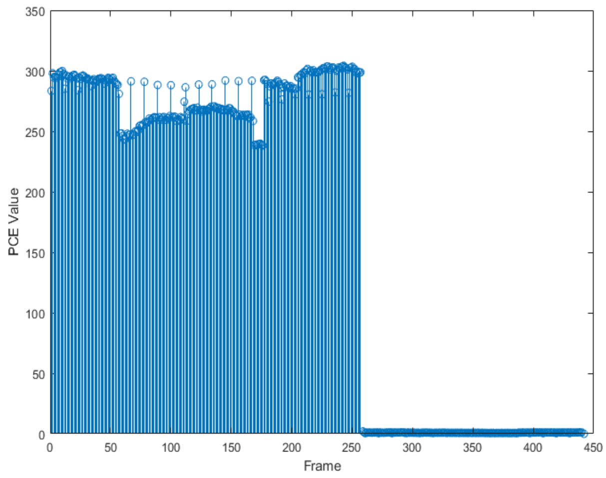


Figure 9. Combined video sequence from Figure 7 (Frame nos. 1 to 256) and Figure 8 (Frame nos. 257 to 447).

4.2. Study of infra-red (IR) night vision video images (size 1280x720)

Infra-red security cameras can capture video in low light and no light (0 Lux) areas. Under such a scenario, it is of interest to check if the SSV and SDA methods can still perform source camera verification. As far is known, no one has studied source camera verification in this case. The two Wansview cameras were used to take videos of a dark/night setting as shown in Figure 5 for the following tests.

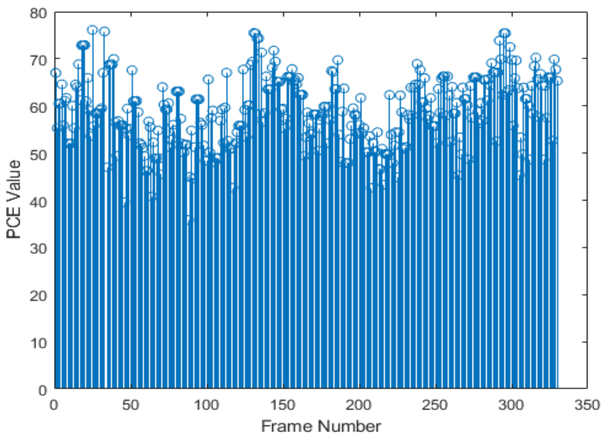


Figure 10. Plot of PCE value results with all frames of an IR scene taken by the Wansview 630KC(Black), tested *vs.* the reference PRNU of the same camera, with Average PCE Value=58.72.

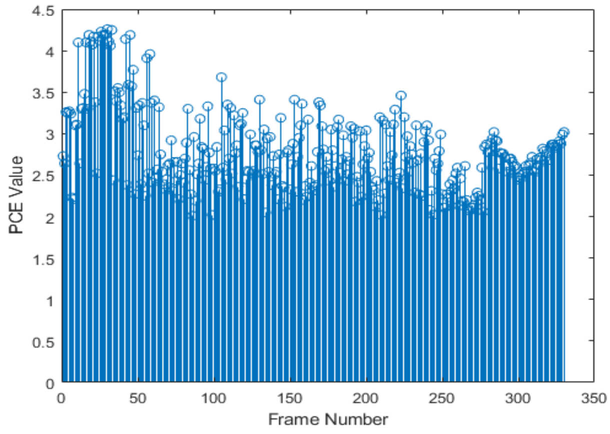


Figure 11. Plot of PCE value results for a similar IR video scene taken by the Wansview (White), tested *vs.* reference PRNU of another camera, Wansview (Black), with PCE_{AVE} Value=2.79.

Figure 10 shows promising results, with all obtained PCE values exceeding 40. It is suggested that this is because IR (infra-red) night vision images are in black-and-white, and so the light intensity component can be strongly detected. Figure 11 depicts the plot for all PCE values of about 2 to 4.3 obtained when the video frames were tested against the reference PRNU of the other camera (not the one taking the video frames under test).

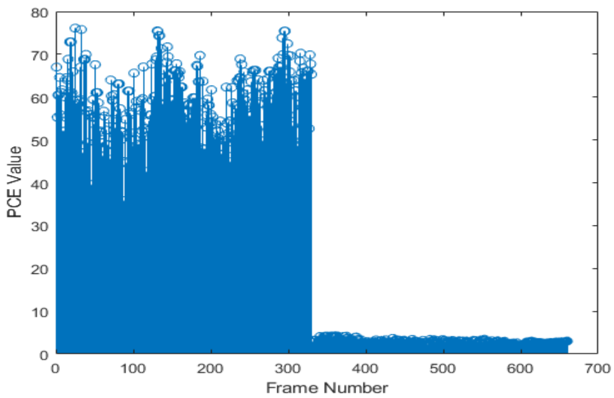


Figure 12. Combined video sequence of Figure 10 (Frame nos. 1 to 330) and Figure 11(Frame nos. 331 to 660).

The results of Figure 10 and Figure 11 can be combined into one graph as shown in Figure 12. It clearly demonstrates that the average and overall PCE values for all video frames taken by the Wansview 630KC(White), tested against the Wansview 630KC(Black) reference PRNU, are much lower than the respective PCE values for all video frames taken by the Wansview 630KC(Black), tested against the reference PRNU of the same camera, Wansview 630KC(Black).

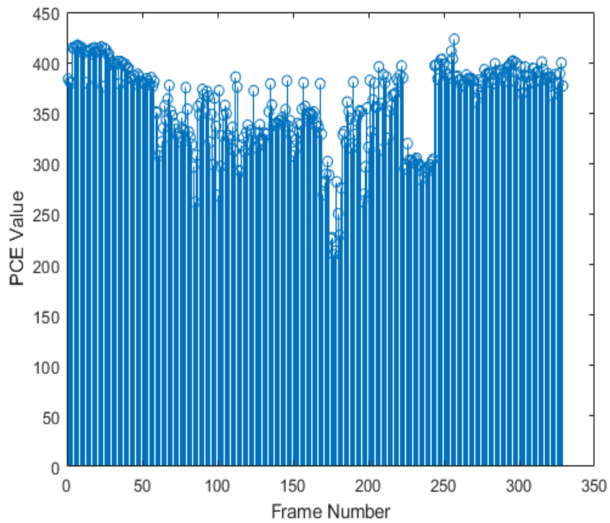


Figure 13. Plot of PCE value results with all frames of an IR scene taken by the Wansview (Black), tested vs. Wansview (Black) reference PRNU obtained by the SDA-frames approach on the same grouped IR video images as the test images.

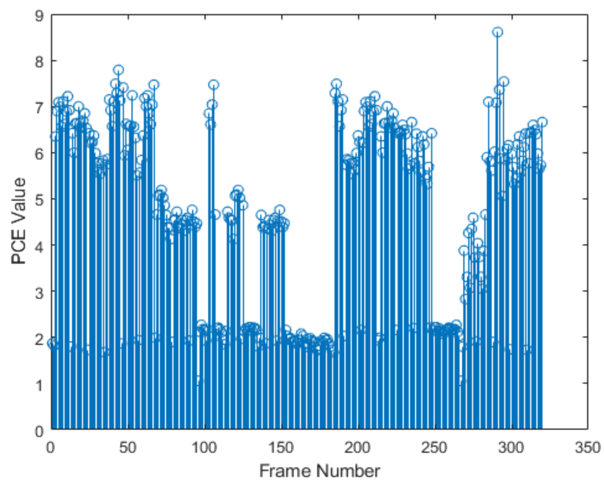


Figure 14. Plot of PCE value results for a similar IR video scene taken by the Wansview (White), tested vs. Wansview (Black) reference PRNU obtained by the SDA-frames approach on the same grouped IR video images as those for the reference PRNU in Figure 13.

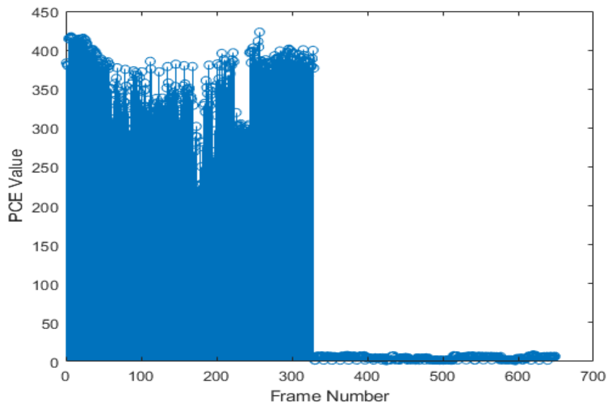


Figure 15. Combined video sequence of Figure 13 (Frame nos. 1 to 325) and Figure 14 (Frame nos. 326 to 650).

The tests on the IR frames were repeated using the SDA-frames method. From Figure 13, high PCE values in the range of 250 to 400 can be seen for the source matching case. Nevertheless, as shown in Figure 14, when the video frames of a similar IR scene but taken by another camera, the Wansview (White), were tested against the reference PRNU of the Wansview (Black), low PCE values between 2 and 9 for all test frames were obtained. Figure 15 shows a plot of the combined sequence for the results of Figure 13 and Figure 14.

Consequently, the results show that both proposed methods and SDA-frames method can verify the video source for infra-red (IR) night vision video images.

4.3. Study of using I-frames only of IR night video images (size 1280 x 720)

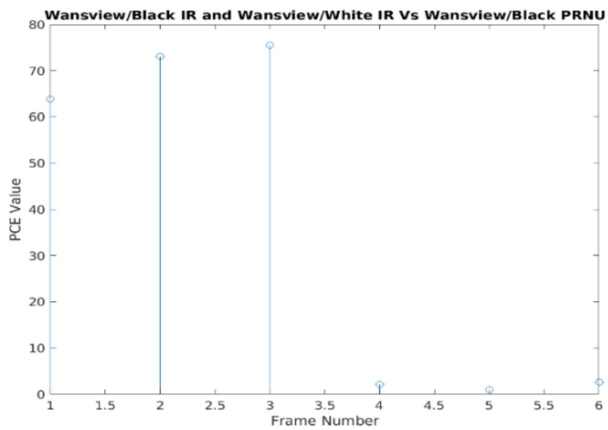


Figure 16. Plot of PCE values for a combined sequence of I-frames only of IR scenes taken by the Wansview (Black) with Frame Numbers 1-3, and Wansview (White) with Frame Numbers 4-6, tested against the reference PRNU of the Wansview (Black).

High computational complexities incur if one obtains the PCE values for all frames in a video signal for camera source identification. To reduce the computational complexity, the PCE value calculation can be restricted to only I-frames instead of all frames in the tests for video source identification. I-frames have less compression than B-frames and P-frames. Therefore, I-frames should have cleaner PRNU patterns than other frame types in GOP.

From the full IR video sequence of 330 frames as in Figure 10, 3 I-frames could be obtained. In Figure 16, with the reference PRNU taken by the Wansview 630KC(Black), PCE values of Frame nos. 1 to 3 in the sequence, taken by the Wansview 630KC(Black) are correspondingly 63.8, 73.0 and 75.5. On the other hand, with the same reference PRNU, PCE values for Frame nos. 4 to 6 in the remaining part of the sequence, taken by the Wansview 630KC(White) become 2.2, 1.0 and 2.6, respectively. The great disparity in the PCE values can thus be useful to verify the video source.

The results show that I-frames are sufficient for video source verification. However, the disadvantage of using I-frames only is that there are too few I-frames from a video for testing. This may lead to unreliable detection as it may take a while before the system notices a significant change (drop or rise) in the PCE value. In other words, false negative rates or false positive rates will increase. Thus, different choices of the solution may need to be considered.

4.4. Issues of false positives

As discussed in previous work (Law and Law, 2018), one has to deal with the false positive problem when the SSV system is used in practical surveillance applications. Two variants of the SSV system are proposed to specifically deal with the false positive problems. With the SSV_1 approach, M consecutive video frames (M=1, M=2, and M=3) from the normal daylight scene (Figure 6) were tested on two different IP video cameras (White, Black) of the same model, the Wansview 630KC. With the SSV_2 approach, the noise residues of a group of frames (denoted as *No*) were averaged to give a resultant noise residue for testing. Two different groups of frames with *No*=3 and *No*=5 were tested for the cameras indicated above.

Table 1 shows the test results. In SSV_1, decreasing false positive rates (FPR), and false negative rates (FNR) were found when M increases. In SSV_2, decreasing FPR and FNR were also found when the number of frames (*No*=5 from *No*=3) used to produce cleaner resultant noise residues for tests increases. This holds true for all source verification tests corresponding to each of these two video cameras. A significant improvement in identification accuracy of 43% (FPR reducing from 51% to 8% for the Wansview (Black) in SSV_1), was observed when M increases from 1 to 2. However, approach SSV_2 generally performs better than approach SSV_1 when comparing their results for both cameras, in terms of FPR and FNR.

Table 1. False Positive (FP) and False Negative (FN) values for tests with two proposed approaches of the SSV system.

		Basic SSV Scheme			Improved SSV Scheme	
		M=1	M=2	M=3	*No*=3	*No*=5
Wansview 630KC (White)	FP	43%	17%	5%	0%	0%
	FN	16%	7%	0%	3%	0%
Wansview 630KC (Black)	FP	54%	12%	4%	19.5%	2%
	FN	6%	3.5%	0%	5%	0%

The results demonstrate that the proposed approaches of the SSV system can verify the video source with a low false positive rate. Furthermore, the approach of averaging noise residues, SSV_2, gives better performance than the other approach, SSV_1, on video source identification with low false positive rates in the majority of cases. SSV_2 can provide better performance because the averaging operation can lessen the effect of the additive noise. As the noise residues resulting from the sensor is a weak signal, removing the additive noise can improve the robustness of the verification. When *No* is chosen to be 5 and the frame rate is 25 fps, it does not introduce much delay in verifying the video source.

5. Discussion

This section discusses how different aspects such as frame type and scene content can affect the performance of signal-based source verification. Through this analysis, implementation of a new method is proposed with examples of the practical applications arising from the whole study of this work.

5.1. Effect of frame types and image content

In the GOP (Group of Pictures) video frame structure, there are I frames, P frames and B frames. From the test results, spikes of PCE values occur periodically where I-frames may be detected. I-frames are coded independently from other frames and have less degree of compression.

This study shows that IR video scenes give even higher PCE values than those of normal daylight scenes, because of the strong light intensity (no chrominance) effect on IR video images. This agrees with the fact that PRNU is the noise residue modulated by a factor of light intensity of the image, as referred to in Equation (3). It follows that IR video images with a higher intensity of noise residues can provide greater accuracy of higher true positive rates (TPR), lower FPR and FNR, as compared to colour images in source identification.

By using I-frames only for tests, both false positive rates and computation complexity can be reduced. Therefore, the effectiveness and reliability of the SSV system can be further enhanced by using I-frames only for testing IR night video scenes. This can be vital, especially when the system is used to detect the source of a received IR night video scene in video surveillance applications, under the digital forensic works that require high accuracy and efficiency.

5.2. Proposed implementation of a new method

One crucial issue in the practical surveillance application is the computational complexity of different approaches. The computation time for fingerprint extraction via the SDA-frames approach has been found to be less than for the conventional approach using the single-thread approach (Farid, 2016). However, the proposed approaches (SSV_1 and SSV_2) address the computational issues in regard to source verification, not fingerprint extraction. Therefore, the computational complexity of the SSV approaches cannot be compared with the SDA-frames approach directly. In order to study the computational complexity of the SSV_2 approach, the PRNU detection algorithm for the SSV approach was run on the full length of the combined video sequence (Figure 12) with clocking of the time of code execution. Then this test was repeated for the SSV_2 approach with *No*=5 on the same sequence to obtain the total execution time. The simulation tests were done using Matlab 2020a on Windows 10 PC with 8 GB memory and an Intel Core i3-7100U @2.40GHz CPU.

Table 2 presents the test results. The difference between the two values of time is due to the reduction of time in testing fewer frames by the approach of averaging noise residues. It should be noted that the denoising of each test frame is still required by this SSV_2 approach, which can take up much of the computation time in the detection process. Despite that, the computational complexity for the proposed SSV_2 approach becomes less than that of the existing SSV approach, due to the implication that the computation time spent on the proposed approach has been reduced by about 25%, as compared to the existing approach.

Table 2. Comparison of execution times between the SSV and SSV_2 approaches.

	Execution Time (sec)	Time Per Frame (sec)
Improved Scheme (*No*=5)	2,574.3	3.9
Basic Scheme (All Frames)	3,457.6	5.2

Following the above discussion, a new implementation to enhance the video source verification performance further is proposed. The SSV_2 approach can complement the SDA-frames approach to form a hybrid, improved version of the existing SSV approach. The outline of operations for this new scheme is shown in Figure 17. In building the camera signature, the SDA method is used. In source verification, SSV_2 is chosen because it is found from the experimental works to be more reliable in regard to source verification than the other approaches, (i.e., SSV, and SSV_1), by giving cleaner noise residues. Besides, SSV_2 has the advantage of taking less computation time in PRNU detection of video test images, in comparison with other SSV approaches. The main benefits of this new method are the reduced false positive rates when using the SSV_2 approach, and the reduced computational complexity under the SDA-frames approach when building the camera signature.

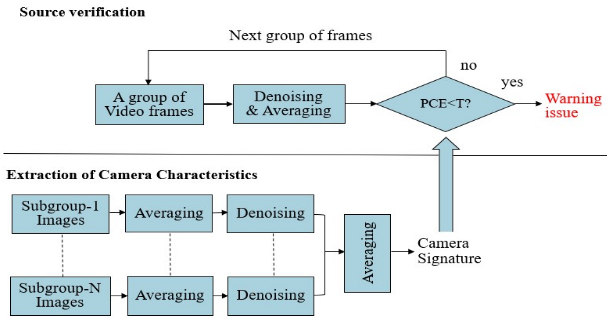


Figure 17. Outline of operations of the new method (SSV_2 and SDA).

6. Conclusions

This paper considers efficient and effective source camera verification for use in networks, IoT and cloud environments. In particular, the PRNU signal is extracted from the video frames for source identification. Different challenges and the robustness of the PRNU-based method to various environmental factors such as day/night and infra-red scenes have been studied. The study indicates that infra-red night vision setting produces accurate camera device verification. Besides, examining only I-frames rather than B- or P-frames can maintain highly accurate verification while reducing computational complexity. Because of the small number of I-frames in video signal, a problem is delay that may occur during device verification. This paper proposes two robust approaches to improve the existing signal-based source verification (SSV) system using PRNU. In particular, by considering the correlation pattern of a number of frames, the false positive rate can greatly be reduced. A comparative study was also performed with a recent technique called SDA-frames approach. Finally a new hybrid method is proposed by combining the best approach with reduced false detection rates, and the SDA-frames approach with reduced computation time. In conclusion, the main contribution of this work is the implementation of video source attribution methods which are robust but reliable with reduced computational complexity.

Acknowledgements

The authors would like to thank the Department of Electronic and Information Engineering, and Faculty of Engineering of The Hong Kong Polytechnic University for inspiring the development of this work.

Notes on contributors

Mr Sai Chung Law received his B.Eng. degree in Electronic Systems from the University of Ulster in the U.K., and M.Sc. degree in Electrical Engineering in The Hong Kong Polytechnic University. He is a practising electronics and information engineer with both professional experience in private and public sectors. His research interests include Internet of Things, information security, digital image processing and digital image forensic techniques for video applications. He is currently pursuing an engineering doctorate degree at The Hong Kong Polytechnic University. In particular, he has two international conference papers and one proceeding paper published on the areas of source camera identification and its applications. He is also a Member of the HKIE and a Chartered Engineer of the U.K.

Dr Ngai Fong Law received her B.Eng. (Hons) degree with first-class honours from the University of Auckland, New Zealand, in 1993 and Ph.D. degree from the University of Tasmania, Australia, in 1997. Her research interests include image forensics, machine learning, and image/signal processing. Dr Law has published over 80 research papers, 40 of which have appeared in international journals, such as IEEE Transactions on Image Processing and Pattern Recognition. She has also extended her study to various new applications, such as information security and bioinformatics. The former includes source camera identification and network traffic anomaly detection. The latter involves works on gene expression and DNA sequence analysis.

References

[1] Anthony C (2019). *Digital Video Surveillance and Security*. 2nd ed. Butterworth-Heinemann.

[2] Banerjee S and Ross A (2018). Impact of photometric transformations on PRNU estimation schemes: A case study using near infrared ocular images. *International Workshop on Biometrics and Forensics (IWBF)*, pp. 1-8.

[3] Chan LH, Law NF and Siu WC (2013). A confidence map and pixel-based weighted correlation for PRNU-based camera identification. *Digital Investigation*, 10(3), pp. 215-225.

[4] Dabov K, Fok A, Katkovnik V and Egiazarian K (2007). Image denoising by sparse 3D transform-domain collaborative filtering. *IEEE Transaction on Image Processing*, 16(18), pp. 2080-2095.

[5] Farid H (2016). *Photo Forensics*. United States: The MIT Press.

[6] Iuliani M, Fontani M, Shullani D and Piva A (2019). Hybrid reference-based video source identification. *Sensors*, 19(3, 649), pp. 1-19.

[7] Iuliani M, Fontani M and Piva A (2021). A leak in PRNU based source identification. Questioning fingerprint uniqueness. *IEEE Access*, 9, pp. 1-9.

[8] Kitchin R and Dodge M (2019). The (In)Security of smart cities: Vulnerabilities, risks, mitigation, and prevention. *Journal of Urban Technology*, 26(2), pp. 47-65.

[9] Law SC and Law NF (2018). PRNU-based source identification for network video surveillance system. *IEEE International Conference on Consumer Electronics (ICCE)*, pp.1-2.

[10] Lukas J and Fridrich J (2006). Digital camera identification from sensor pattern noise. *IEEE Transactions of Information Forensics and Security*, 1(2), pp. 205-214.

[11] Mandelli S, Bestagini P, Verdoliva L and Tubaro S (2020). Facing device attribution problem for stabilized video sequences. *IEEE Transactions of Information Forensics and Security*, 15, pp. 1-14.

[12] Milani S, Potani M, Bestagini P, Barni M, Piva A, Tagliasacchi M, and Tubaro S (2012). An overview on video forensics. *International Conference on Advanced Technologies for Signal and Image Processing*, pp. 1-18.

[13] Pandey RC, Singh SK and Shukla KK (2016). Passive forensics in image and video using noise features: A review. *Digital Investigation*, 19, pp. 1-28.

[14] Redi JA, Takak W and Dugelay JL (2011). Digital image forensics: A booklet for beginners. *Multimedia Tools and Applications*, 51(1), pp. 133-162.

[15] Shi C, Law NF, Leung HF and Siu WC (2017). A local variance based approach to alleviate the scene content interference for source camera identification. *Digital Investigation*, 22, pp. 74-87.

[16] Taspinar S, Mohanty M and Memon N (2020). Camera fingerprint extraction via spatial domain averaged frames. *IEEE Transactions of Information Forensics and Security*, 15(2), pp.3270-3282.