

PRNU Extraction from Stabilized Video: A Patch Maybe Better than a Bunch

Bin Ma¹, Yuanyuan Hu¹, Jian Li^{1,*}, Chunpeng Wang¹, Meihong Yang² and Yang Zheng³

¹Qilu University of Technology (Shandong Academy of Sciences), Shandong Provincial Key Laboratory of Computer Networks, Jinan, 250353, China

²Shandong Computer Science Center, Shandong Provincial Key Laboratory of Computer Networks, Jinan, 250014, China

³University of Connecticut, Mansfield, CT, 06269, USA

*Corresponding Author: Jian Li. Email: ljian20@gmail.com

Received: 01 September 2020; Accepted: 16 October 2020

Abstract: This paper presents an algorithm to solve the problem of Photo-Response Non-Uniformity (PRNU) noise facing stabilized video. The stabilized video undergoes in-camera processing like rolling shutter correction. Thus, misalignment exists between the PRNU noises in the adjacent frames owing to the global and local frame registration performed by the in-camera processing. The misalignment makes the reference PRNU noise and the test PRNU noise unable to extract and match accurately. We design a computing method of maximum likelihood estimation algorithm for extracting the PRNU noise from stabilized video frames. Besides, unlike most prior arts tending to match the PRNU noise in whole frame, we propose a new patch-based matching strategy, aiming at reducing the influence from misalignment of frame the PRNU noise. After extracting the reference PRNU noise and the test PRNU noise, this paper adopts the reference and the test PRNU overlapping patch-based matching. It is different from the traditional matching method. This paper conducts different experiments on 224 stabilized videos taken by 13 smartphones in the VISION database. The area under curve of the algorithm proposed in this paper is 0.841, which is significantly higher than 0.805 of the whole frame matching in the traditional algorithm. Experimental results show good performance and effectiveness the proposed strategy by comparing with the prior arts.

Keywords: Photo-response non-uniformity (PRNU); stabilized video; frames; maximum likelihood estimation; patch

1 Introduction

In the past decades, some important technologies such as digital watermarking [1,2], data hiding [3,4] and multimedia forensics have emerged to strengthen the protection of multimedia data and to combat illegal video transmission. Among them, multimedia forensics technology mainly refers to source camera identification (SCI). SCI is a method to tie an image or video with a certain possibility to its capturing device. Most of the existing methods are based on Photo Response Non-uniformity (PRNU) noise extraction and matching techniques [5]. The PRNU noise stems from the imperfect manufacturing



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

process of camera sensor and exists in all its captured medias. As the PRNU noise is unique to its camera and it is stable for a long time, we can take it as a fingerprint for SCI.

In the literature, Lucáš et al. [5] proposed for the first time to extract sensor pattern noise by filtering a large number of images taken by the same camera. Chen et al. [6] employed a mathematical method to estimate the PRNU noise from video clips. After that a large number of SCI schemes based the PRNU noise exist and some problems involved in this technique are solved as well [7]. Recently, Darvis et al. [8] used local PRNU matching method to SCI of HDR image. Taspinar et al. [9] proposed to extract the PRNU from spatial domain averaged (SDA) to reduce complexity and not lose performance.

In addition, with the widespread application of video, Galdi et al. [10] found through experiments that extending the traditional method of extracting the PRNU noise from images to video does not achieve ideal results. There are two main problems.

One problem is that video files have more data than images and are mostly recompressed to save storage space. Recompression will cause severe degradation of the extracted the PRNU noise. How to extract reliable PRNU noise from compressed video is addressed in [11–19]. Chuang et al. [12] proposed that when estimating PRNU noise, only key frames (I frames) were used, and motion compensation frames (P, B frames) were excluded. This is because the main information of the video file is concentrated in the I frames. And the P and B frames are not helpful in extracting PRNU noise. Li et al. [14] proposed to extract PRNU noise from partially decoded video frames. This method had improved the accuracy and efficiency of extraction. Amerini et al. [16] compared the existing algorithms for extracting the PRNU noise from videos and proposed an efficient comprehensive algorithm for videos uploaded to social media. In order to protect data integrity in network applications, Wang et al. [19] proposed an effective dual-chaining watermark scheme, called DCW.

Second, in-camera functions like video stabilization for the unconscious jitter reduction [17] can introduce misalignment to the PRNU noise frames [20]. As a result, it is uneasy to correctly estimate the PRNU noise by means of statistical method. To solve this problem, Höglund et al. [21] proposed to compensate the translation between the PRNU noises. Taspinar et al. [20] proposed to divide the video frame into two parts and associate its PRNU noise. The video is unstabilized based on the peak correlation energy (PCE) value higher than the threshold. Then, obtained the translation of each frame by an exhaustive method. Iuliani et al. [22] proposed to extract the reference PRNU from the image. Mandelli et al. [23] used particle swarm optimization technology, the PCE value of the reference PRNU noise and the test PRNU noise reaches the maximum and is greater than the threshold. The corresponding parameter is the geometric transformation parameter of aligning the two kinds of noises.

Considering that to estimate reliable the PRNU noise from stabilized video clips is still not well solved at the current stage, this paper proposes a new PRNU noise-matching algorithm for SCI regarding video. We first extract the PRNU noise from video clips, like the traditional methods. Then, the two PRNU noise are segmented into patches for matching. It is different to perform block processing on the image to select the part of interest [24,25]. The motivation of this matching method is from the theoretical analysis on the influence from image/frame registration introduced by in-camera functions about video stabilization. Experimental results show that the proposed method can achieve higher accuracy than the traditional ones that use a bunch of frames or one frame for extraction and correlation.

The rest of this paper is organized as follows. Firstly, relevant background will be further introduced in Section II. And then Section III gives the algorithm proposed in this paper. Extensive results will be discussed in Section IV, followed by the conclusion in Section V.

2 Related Works and Background

This section will introduce the traditional mathematical estimation method regarding video PRNU noise and in-camera video stabilization functions. Then we demonstrate the inaccuracy of PRNU noise extraction and matching due to the in-camera stabilization functions.

Considering that each frame can be regarded as an image taken by the same camera, the traditional method derives the estimation of PRNU noise \tilde{K} as follows:

$$\tilde{K} = \frac{\sum_{i=1}^n W^{(i)} I^{(i)}}{(I^{(i)})^2} \quad (1)$$

where $W^{(i)}$ is the noise residual extracted from $I^{(i)}$, $W^{(i)} = I^{(i)} - F(I^{(i)})$ being $F(I^{(i)})$ a denoised version of $I^{(i)}$, computed as suggested in Reference [6].

Nowadays, most smartphone cameras employ a so-called rolling shutter technology to output each rows of the pixel sensor array from top to bottom sequentially. The patent embodiments in Reference [26] points out the effect of smartphone rolling shutter, as shown in Fig. 1. The rolling shutter scans line by line from top to bottom. When there is relative movement between the subject and the smartphone camera, the subject will distortion and jag blur on the edge. Resulting frames suffering from rolling shutter distortions are often actual scene unwanted.

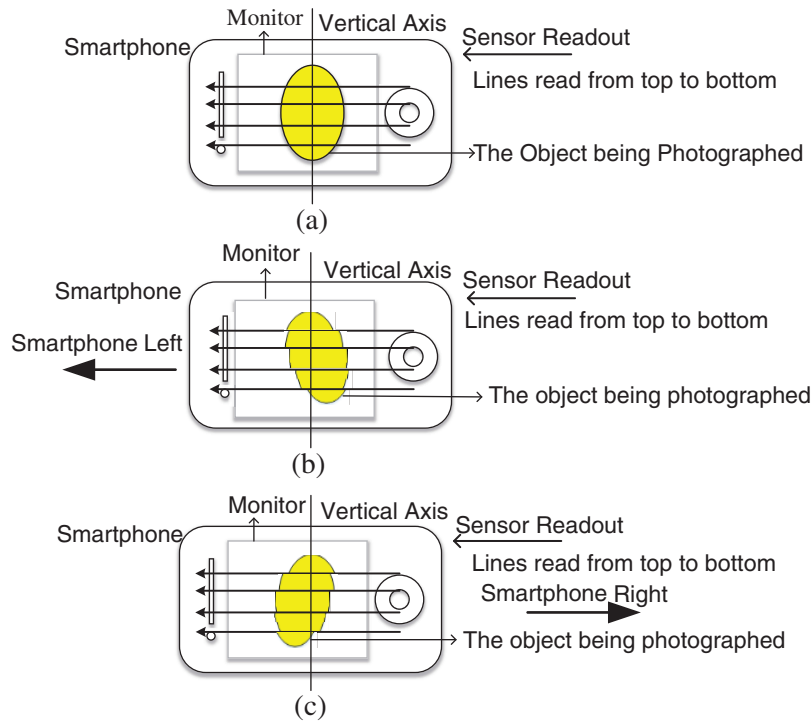


Figure 1: The influences of rolling shutter [26]. The rolling shutter scans line by line from top to bottom. Figs. 1(b)–1(c) assumes that during the rolling shutter scanning, the smartphone moves to the left (right), resulting in the lower part of the subject (the part scanned after the rolling shutter) is offset from the upper part (the part scanned first by the rolling shutter) and jagged blur on the edge

It is necessary to reduce the effects of rolling shutter distortion via appropriate perspective transformation during video capture. Individual registration process should be performed for different parts of a video frame. The two-dimensional perspective transformation matrix is independently applied to each part of each frame, and the corrected frame segments of each frame are composed into a corrected frame [26]. The PRNU noise is the inherent noise of the sensor. After geometric transformation of different parts of a video frame, the same pixel position is taken by different parts of the sensor array. This results in misaligned the PRNU noise in different parts of a video frame. Moreover, the offset of each part is different according to the two-dimensional perspective change matrix. The local or global offset of the PRNU noise makes SCI difficult.

3 PRNU Noise Extractions and Matching for Stabilized Video

Given a stabilized video file, which is common for us facing forensics job. The stabilized video suffers in-camera processing like rolling shutter correction. Each frame undergoes different global and local geometrical transformations. As a result, global and local misalignment exists among the PRNU noise contained in the video frames. Therefore, we design a computing method of maximum likelihood estimation algorithm for extracting PRNU noise and propose a new overlapping patch-based matching strategy. In other words, we match the PRNU noise in patches. The purpose of this is to reduce the impact of local or global stabilization on the PRNU noise matching. Fig. 2 shows the overall flowchart of the proposed scheme.

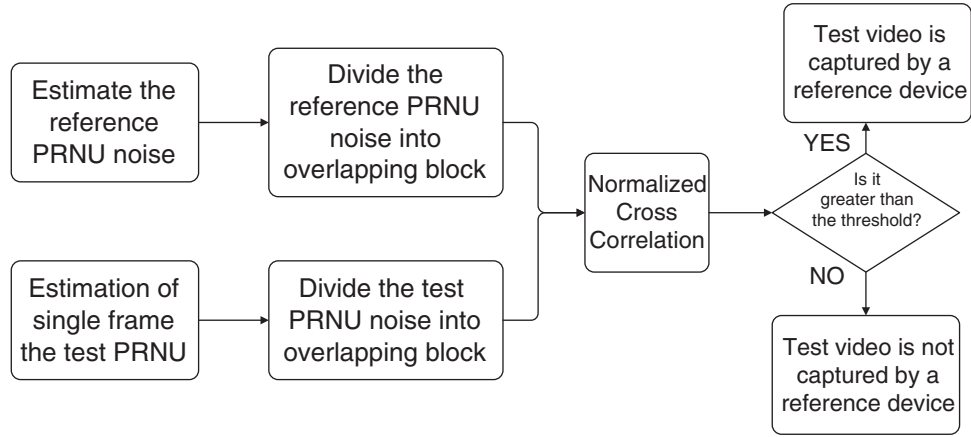


Figure 2: Scheme flowchart

3.1 PRNU Noise Extraction from Stabilized Video Clip

In some forensic scenarios, we may be unable to access the capturing device. As a consequence, the reference PRNU noise cannot be obtained from images taken by the device. But the reference PRNU noise can only be extracted from a number of obtained stabilized videos. In this case, neither the test PRNU noise nor the reference PRNU noise can be reliably estimated. There are global and local misalignments between PRNU noises contained in the video frames.

Let $\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_n\}$, $\vec{f}_i \in N^{x \times y}$ represents a set of frames coming from the video clips captured by the same smartphone. Each frame corresponds to a geometric transformation T_i^θ introduced by in-camera rolling shutter correction. Moreover, please notice that T_i^θ not a global geometrical transformation for the whole frame. Instead, it is changed for different part of a frame. Considering that camera motion may be a cyclical process, so the same geometric transformation exists in this process. Furthermore, there are a

large number of frames in a video clip, so there may be frames with the same geometric transformation. Moreover, classify the frames with the same geometric transformation into one group. We can obtain a sequence of frame groups $\{1, 2, \dots, m\}$, each group represented by k_1, k_2, \dots, k_m . The matrices $T_{k_1}, T_{k_2}, \dots, T_{k_m}$ are the geometric transformations to which each group of frames is subjected.

According to the prior art [27], it is suitable to use the method of maximum likelihood estimation to obtain the PRNU noise \tilde{K} . However, each frame undergoes a different local geometric transformation, resulting in individual pixels being misaligned within the frame. Therefore, the maximum likelihood estimation model extended from image to stabilized video can no longer effectively represent the PRNU noise extracted by grouping or single frame in traditional algorithms. As shown in Eq. (2), the maximum likelihood estimation model in this paper is decomposed to make it accurately represent the PRNU noise of traditional algorithm group estimation. Using a large number of frames in a video clip. Eq. (2) assumes that frames with the same correction transform are put in the same group, instead of the numerator in the original Eq. (1), mathematically,

$$\tilde{K} = \frac{\sum_{i=1}^n W^{(i)} I^{(i)}}{(I^{(i)})^2}$$

$$\tilde{K} = \frac{\sum_{i \in k_1} W_{k_1}^{(i)} I_{k_1}^{(i)}}{\sum_{i=1}^n (I^{(i)})^2} + \frac{\sum_{i \in k_2} W_{k_2}^{(i)} I_{k_2}^{(i)}}{\sum_{i=1}^n (I^{(i)})^2} + \dots + \frac{\sum_{i \in k_m} W_{k_m}^{(i)} I_{k_m}^{(i)}}{\sum_{i=1}^n (I^{(i)})^2}, \quad (2)$$

$$\tilde{K} = \frac{\sum_{i \in k_1} W_{k_1}^{(i)} I_{k_1}^{(i)}}{\sum_{i \in k_1} (I^{(i)})^2} \cdot \frac{\sum_{i \in k_1} (I^{(i)})^2}{\sum_{i=1}^n (I^{(i)})^2} + \dots + \frac{\sum_{i \in k_m} W_{k_m}^{(i)} I_{k_m}^{(i)}}{\sum_{i \in k_m} (I^{(i)})^2} \cdot \frac{\sum_{i \in k_m} (I^{(i)})^2}{\sum_{i=1}^n (I^{(i)})^2}. \quad (3)$$

Define,

$$P_{k_j} = \frac{\sum_{i \in k_j} (I^{(i)})^2}{\sum_{i=1}^n (I^{(i)})^2}, \quad (4)$$

where i is the index of the frames and $j \in \{1, 2, \dots, m, \dots\}$ indicates the index of the groups. Namely, frames in group j have the same geometric transformation introduced by rolling shutter correction. Eq. (4) shows the proportion of frames in each group to the total number of frames, indicating a probability. As long as $j > 0$ when the geometric transformation of some frames is the same, $P_{k_j} < 1$. Hence, the PRNU noise of all frames is,

$$\begin{aligned} \tilde{K} &= T_{k_1}(\tilde{K}_{k_1}) \cdot P_{k_1} + T_{k_2}(\tilde{K}_{k_2}) \cdot P_{k_2} + \dots + T_{k_m}(\tilde{K}_{k_m}) \cdot P_{k_m} \\ &= \sum_{j \in \{1, 2, \dots, m\}} P_{k_j} \cdot T_{k_j}(\tilde{K}_{k_j}) \end{aligned} \quad (5)$$

T_{k_j} is the correction transformation suffered by the j group PRNU noise. Furthermore, the test PRNU noise of each frame can be expressed as:

$$\tilde{K}_{testi} = T_i(\tilde{K}_i), \quad (6)$$

T_i represents the geometric transformation of the PRNU noise in frame \vec{f}_i

Because rolling shutter correction applies a two-dimensional perspective transformation matrix independently to different patches within the frame. Therefore, there is the misalignment of PRNU noise of different patches within the frame. In this light, decompose frame \vec{f}_i into strips $\{1, 2, \dots, l\}$. r_1, r_2, \dots, r_l represent the index of patches, the size of each patch is $\left\lceil \frac{x}{l} \right\rceil \cdot y$. $x \cdot y$ represents the video resolution. Each PRNU noise patch corresponds to a geometric transformation $T_{r_1}, T_{r_2}, \dots, T_{r_l}$. The PRNU noise of each patch in the frame is:

$$\tilde{K}_{test \mathcal{J}_i} = T_{r_i}(\tilde{K}_{r_i}), \quad (7)$$

where r_i represents the r_i th patch in the \vec{f}_i frame. $\tilde{K}_{test \mathcal{J}_i}$ indicates the PRNU noise after patching.

3.2 The Reference and Test PRNU Noise Matching

The similarity between the reference PRNU noise and the test PRNU noise measures by PCE value to determine whether it comes from the same camera. In order to calculate the PCE values on two-dimensional matrices, it is necessary to normalized cross correlation (NCC) the reference PRNU noise and the test PRNU noise. According to Eq. (5), in addition to misalignment exists between adjacent frame PRNU noises, there will be the proportion of the number of frames in each group to the total number of frames when PRNU noise is calculated by grouping. If use the traditional method, group extraction the test PRNU noise and then match the reference PRNU noise, for example, Eq. (8).

$$C(u, v) = corr(\tilde{K}_{ref}, \tilde{K}_{test}) = \sum_{j=1}^m P_{k_j}^2 corr(T_{k_j}(\tilde{K}_{k_j}), T_{k_j}(\tilde{K}_{k_j})). \quad (8)$$

There will be P_{k_i} when each group is matched. Because of $P_{k_i} < 1$, the similarity between the test PRNU noise and the reference PRNU noise will be weakened. The phenomenon of mismatching will occur. In addition, the false rejection rate will be improved. If the test PRNU noise of each frame matches the reference PRNU noise, namely, there is only one P_{k_i} from the reference PRNU noise at this time. If the PRNU noise of the frame belongs to the group transformation, the matching is successful, according to Eq. (9).

$$C(u, v) = corr(\tilde{K}_{ref}, \tilde{K}_{test}) = \sum_{j=1, i \in k_j}^m P_{k_j} corr(T_{k_j}(\tilde{K}_{k_j}), T_i(\tilde{K}_i)). \quad (9)$$

However, considering that the rolling shutter is commonly used in smartphones, Eq. (9) may have errors in the whole frame matching the reference PRNU noise with the test PRNU noise. Since the registration introduced by the rolling shutter correction is performed on the within frame. The PRNU noise has different offset and local misalignment in each frame. Taking into account the impact of the rolling shutter correction, we propose a new overlapping patch-based matching strategy. We put the PRNU noise from the whole frame to the patch. The reason for overlapping at this stage is that it is difficult to determine which of the PRNU noise intra-frame misalignment caused by the rolling shutter correction. We reduce the error as much as possible by overlapping the patches. As Eq. (10),

$$C(u, v) = \text{corr}(\tilde{K}_{ref-r_i}, \tilde{K}_{test-r_i}) = \sum_{j=1, i \in k_j, r_i \in i}^m P_{k_j} \text{corr}(T_{r_i}(\tilde{K}_{ref-r_i}), T_{r_i}(\tilde{K}_{tes-r_i})), \quad (10)$$

where $\tilde{K}_{ref-r_i}, \tilde{K}_{tes-r_i}$ indicate the reference PRNU noise and test PRNU noise for the corresponding patch. The PRNU noise for overlapping patch matching may be accurate. Because it reduces the influence of the PRNU noise local misalignment. It is easier to judge whether the test video comes from the reference smartphone.

A single peak is used to determine whether the video is taken by this camera. As long as the PCE value of a frame is greater than the threshold value, the video is considered to be taken from a reference camera.

$$PCE(\tilde{K}_{ref-r_i}, \tilde{K}_{test-r_i}) = \frac{\text{sign}(C) \cdot C(u_{peak}, v_{peak})^2}{\frac{1}{\left\lceil \frac{x}{7} \right\rceil \cdot y - N_{peak}} \sum_{u, v \notin N_{peak}} C(u, v)^2}. \quad (11)$$

Eq. (11) shows that after the reference PRNU noise and the test PRNU noise divided into overlapping patches, the similarity of the corresponding patches measured by PCE value. (u_{peak}, v_{peak}) represents peak coordinates, and N_{peak} represents peak neighborhood. After PRNU noise divided, the matching time complexity reduced.

4 Experimental Results

This section presents the experimental results of the proposed algorithm. First, we describe the used database. Then we demonstrate of the performance of our proposed patch matching method, via the comparison between the proposed algorithm with the prior arts.

4.1 Experiment Setup

The experiment in this paper is executes on an Intel® Core™ i7-8700 CPU with a frequency of 3.20GHz. The patch matching-based SCI scheme implements on Windows 10 (64) platform using MATLAB R2015b. All the videos used are from the VISION [28] database. Specifically, 13 smartphones, such as iPhone, Sony, OnePlus, etc. Each video is approximately one minute with a resolution of 1080×1920 , as shown in Tab. 1. Each device randomly selected about 60 videos as inter-class tests. Using Receiver Operating Characteristic Curve (ROC) and Area Under Curve (AUC) to show performance of the proposed algorithm and other algorithms.

Table 1: The list of our smartphone used

Smartphone No.	Smartphone name	No. of videos	Video resolution
D02	Iphone4s	13	1920×1080
D05	Iphone5c	19	1920×1080
D06	Iphone6	17	1920×1080
D10	Iphone4s	15	1920×1080
D12	Sony	19	1920×1080
D14	Iphone5c	19	1920×1080
D15	Iphone6	18	1920×1080

(Continued)

Table 1 (continued).			
Smartphone No.	Smartphone name	No. of videos	Video resolution
D18	Iphone5c	13	1920×1080
D19	Iphone6Plus	17	1920×1080
D25	OnePlus_A3000	19	1920×1080
D29	Iphone5	19	1920×1080
D32	OnePlus_A3003	19	1920×1080
D34	Iphone5	32	1920×1080

4.2 Comparison with Prior Arts

First, we tried to show the splitter size best suited to PRNU noise matching. According to the patents embodiment regarding rolling shutter correction, in most cases, video frames are divided into 25, 32, 60, 100 and 180 rows, for registration 44, 34, 18, 11 and 6 patches. Moreover, each patch is overlapped and taken half of the number of rows, each performing 107, 67, 35, 21, 12 matches. Calculate the correct rate and false alarm based on the matching results of different segmentation methods. The AUC of different patch methods show in Tab. 2. The AUC has little difference between 100 rows and 60 rows, but with time complexity considerations, 100 rows are better, as shown in Fig. 3.

Table 2: AUC of different branches

Different branches	100×1980	180×1980	60×1980	32×1980	25×1980
AUC	0.854	0.823	0.841	0.717	0.786

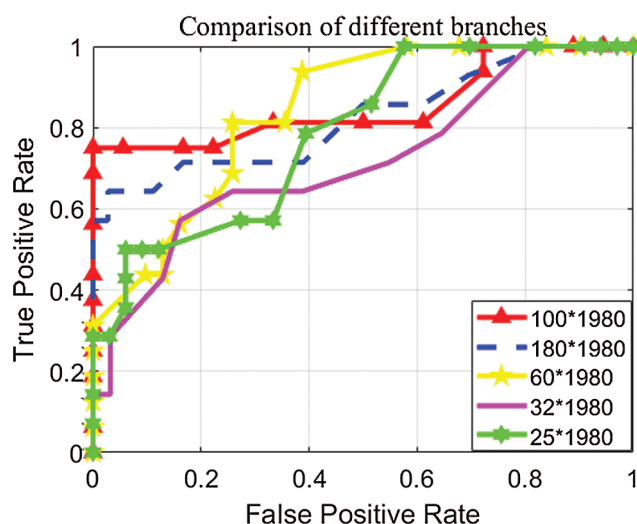


Figure 3: ROC curve of different overlapping patches. The largest AUC at 100×1980 , the best performance for different patches

According to the experimental comparison of the matching results of different segmentation methods, it is determined that the matching effect is better when the reference PRNU noise and the test PRNU noise are each divided into 100×1920 patches. The local geometric transformation may stabilize the video frames with a size of about 100×1920 per patches. Therefore, the PRNU noise patching may reduce the effect of local misalignment. The reference PRNU noise and the test PRNU noise can be matched according to the corresponding patch. Therefore, this paper proposes to match the PRNU noise patches after extracting the reference PRNU noise and the test PRNU noise. The ROC curve is compared with the prior arts based on the whole frame matching. The first prior art is to extract the test PRNU noise from a single frame without any processing. The second is that [12,20,28] only decodes the key frames in the video to extract the PRNU noise. The third is taking all video frames to extract PRNU noise. Each group contains the same number of frames to extract the test PRNU noise, and matches with the reference PRNU noise separately.

In order to prove the validity of the algorithm, we try to use intra-class and inter-class testing of videos from different smartphones of the same brand model to avoid the contingency and to prove the accuracy of the algorithm. As shown in Figs. 4(a)–4(d), for Iphone6, in the whole frame-based method, the maximum AUC is 0.799. The AUC of the PRNU noise overlapping patch matching in this paper is 0.819. The proposed method of patch matching has good performance. ROC curve of other smartphones in the database, such as Iphone 4S, Iphone 5C and Iphone 6plus, is shown in Figs. 4(b)–4(d).

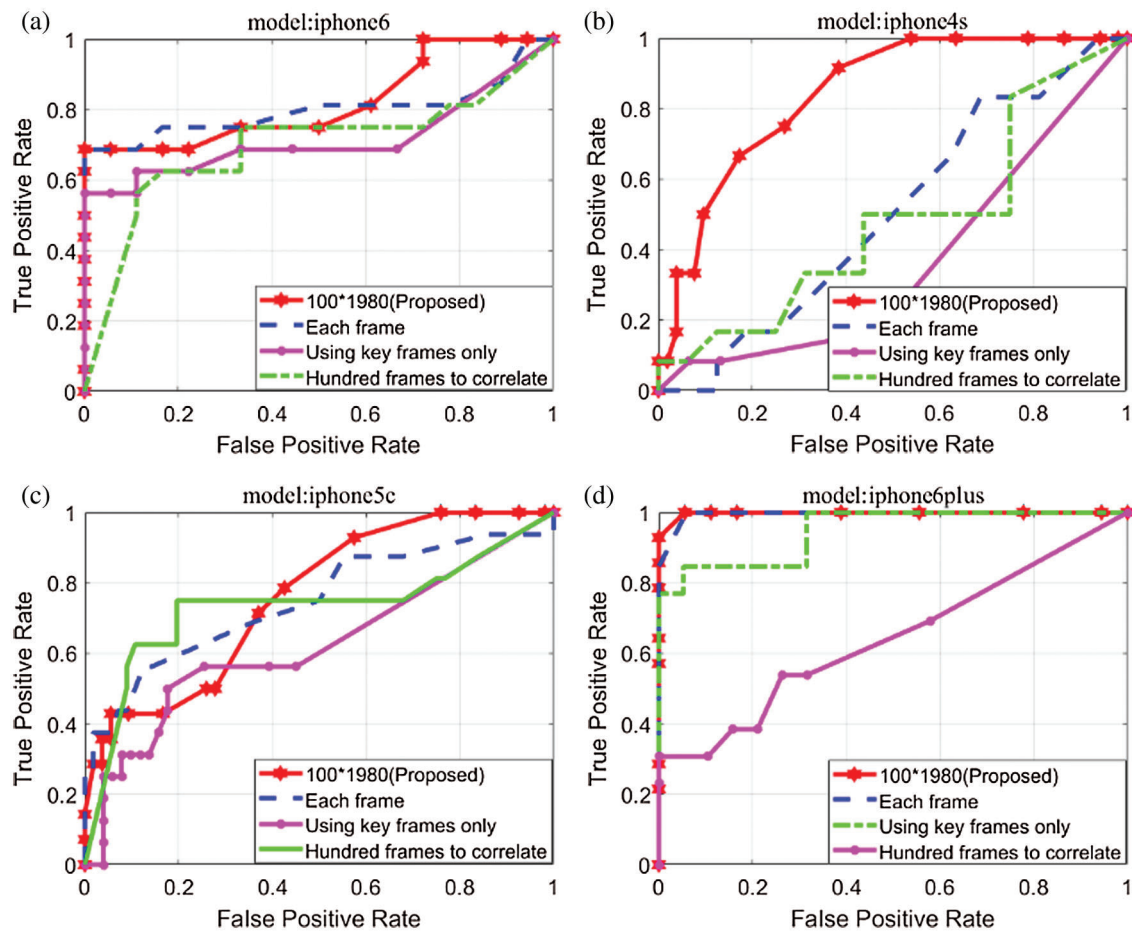


Figure 4: ROC curves of different smartphone

Fig. 4(a) iPhone 6, the algorithm is divided into 100×1980 AUC = 0.819. Compared with the experimental whole frame AUC = 0.799. Only the key frames in the video are decoded to extract AUC = 0.715 for the test PRNU noise. And AUC = 0.696 for extracting and matching the PRNU noise from video frames by grouping processing. It can be seen that the performance of this algorithm is better.

Then we further examine the performance of our proposed overlapping patch-based matching strategy algorithm by comparing with the above three prior arts through the overall ROC curve. We divide the two PRNU noises into 100×1920 sizes. Perform intra-class and inter-class experiments of 13 smartphones in the database according to different methods. And calculate the ROC curves of all smartphones. ROC curves of 13 smartphones of each method are averaged to measure the overall performance of the algorithm. As shown in Fig. 5, the AUC area of the algorithm proposed in this paper is 0.841. The maximum AUC based on the whole frame algorithm is only 0.805. Due to the combined effect of global and local geometrical transformations, the patch matching algorithm in this paper has a higher accuracy than the whole frame matching method. The overall ROC curve is shown in Fig. 5.

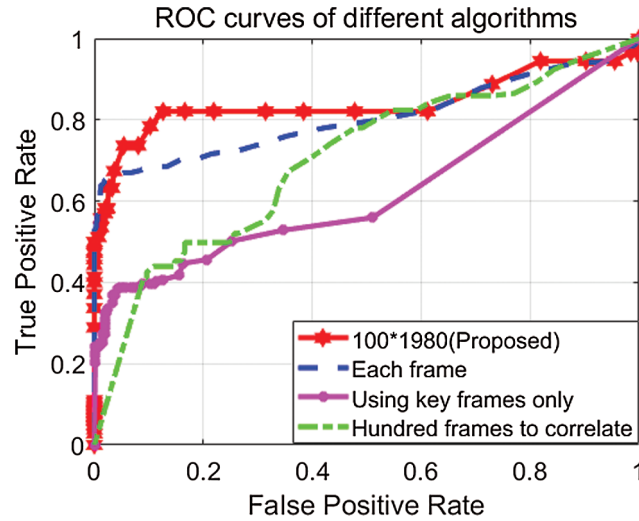


Figure 5: ROC Curves of Different Algorithms, experimental results show that the AUC = 0.841 of PRNU noise matched by the proposed algorithm. In contrast experiment, the whole frame matching PRNU noise AUC = 0.805. AUC = 0.622 using only key frames. The PRNU noise group accumulation matching once AUC = 0.702

Due to the combined effect of global and local geometrical transformations, experimental results show good performance and effectiveness the proposed strategy by comparing with the prior arts.

5 Conclusion

This paper has proposed a PRNU noise extraction and correlation algorithm for stabilized videos captured by the smartphone. We have two contributions. First, we update the mathematical model of PRNU noise based on the effects of frame registration introduced by in-camera processing. Therefore, it is more accurate to match the PRNU noise of a stabilized video than the PRNU mathematical model that extends from image to video. Second, when each frame undergoes a different global and local geometric transformation, we design a matching algorithm for PRNU noise. The method of adopting overlapping patch for the first time is better than the traditional method of matching whole PRNU noise. Moreover, determine the applicability of the algorithm. The experimental campaign is conducted on an available

dataset composed by almost 224 stabilized video sequences coming from smartphone. Experimental results demonstrate that the proposed computing method has good performance for stabilized video in-camera processing like rolling shutter correction. In the future, we plan to extend to our work both to reduce the error rate and to improve efficiency.

Acknowledgement: We are thankful to the reviewers for their useful and constructive suggestions which improved this paper very well.

Funding Statement: This research was funded by the National Natural Science Foundation of China (61872203 and 61802212), the Shandong Provincial Natural Science Foundation (ZR2019BF017), Major Scientific and Technological Innovation Projects of Shandong Province (2019JZZY010127, 2019JZZY010132, and 2019JZZY010201), Jinan City “20 universities” Funding Projects Introducing Innovation Team Program (2019GXRC031), Plan of Youth Innovation Team Development of colleges and universities in Shandong Province (SD2019-161), and the Project of Shandong Province Higher Educational Science and Technology Program (J18KA331).

Conflicts of Interest: All authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] C. Wang, X. Wang, Z. Xia and C. Zhang, “Ternary radial harmonic Fourier moments based robust stereo image zero-watermarking algorithm,” *Information Sciences*, vol. 470, pp. 109–120, 2019.
- [2] N. Jayashree and R. S. Bhuvaneswaran, “A robust image watermarking scheme using Z-Transform, discrete wavelet transform and bidiagonal singular value decomposition,” *Computers, Materials & Continua*, vol. 58, no. 1, pp. 263–285, 2019.
- [3] B. Ma and Y. Q. Shi, “A reversible data hiding scheme based on code division multiplexing,” *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 9, pp. 1914–1927, 2017.
- [4] D. Xiao, J. Liang, Q. Ma, Y. Xiang and Y. Zhang, “High capacity data hiding in encrypted image based on compressive sensing for nonequivalent resources,” *Computers, Materials & Continua*, vol. 58, no. 1, pp. 1–13, 2019.
- [5] J. Lucáš, 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.
- [6] M. Chen, J. Fridrich, M. Goljan and J. Lucáš, “Source digital camcorder identification using sensor photo response non-uniformity,” *Security, Steganography, and Watermarking of Multimedia Contents IX*, International Society for Optics and Photonics, San Jose, CA, United States, vol. 6505, 2007.
- [7] M. Goljan and J. Fridrich, “Camera identification from cropped and scaled images,” *Security, Forensics, Steganography, and Watermarking of Multimedia Contents X*, International Society for Optics and Photonics, San Jose, CA, USA, vol. 6819, 2008.
- [8] M. Darvish, M. Hosseini and M. Goljan, “Camera identification from HDR images,” in *Proc. of the ACM Workshop on Information Hiding and Multimedia Security*, Paris, France, pp. 69–76, 2019.
- [9] S. Taspinar, M. Mohanty and N. Memon, “Camera fingerprint extraction via spatial domain averaged frames,” *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 3270–3282, 2020.
- [10] C. Galdi, F. Hartung and J. L. Dugelay, “Videos versus still images: Asymmetric sensor pattern noise comparison on mobile phones,” *Electronic Imaging, Media Watermarking, Security, and Forensics*, vol. 2017, no. 7, pp. 100–103, 2017.
- [11] W. Van Houten and Z. Geradts, “Source video camera identification for multiply compressed videos originating from YouTube,” *Digital Investigation*, vol. 6, no. 1–2, pp. 48–60, 2009.

- [12] W. H. Chuang, H. Su and M. Wu, "Exploring compression effects for improved source a camera identification using strongly compressed video," in *IEEE International Conference on Image Processing (ICIP)*, Brussels, Belgium, pp. 1953–1956, 2011.
- [13] D. K. Hyun, C. H. Choi and H. K. Lee, "Camcorder identification for heavily compressed low resolution videos," *Computer Science and Convergence*, Dordrecht, pp. 695–701, 2012.
- [14] J. Li, B. Ma and C. Wang, "Extraction of PRNU noise from partly decoded video," *Journal of Visual Communication and Image Representation*, vol. 57, pp. 183–191, 2018.
- [15] E. Altinisik, K. Tasdemir and H. T. Sencar, "Extracting PRNU noise from H.264 coded videos," *European Signal Processing Conference (EUSIPCO)*, Rome, Italy, pp. 1367–1371, 2018.
- [16] I. Amerini, R. Caldelli, A. D. Mastio, A. D. Fuccia, C. Molinari *et al.*, "Dealing with video source identification in social networks," *Signal Processing: Image Communication*, vol. 57, pp. 1–7, 2017.
- [17] D. J. Thivent, G. E. Williams, J. Zhou, R. L. Baer, R. Toft *et al.*, "Combined optical and electronic image stabilization," U.S. Patent, no. 9,596,411, 2017.
- [18] C. Wang, X. Wang, Z. Xia, B. Ma and Y. Q. Shi, "Image description with polar harmonic Fourier moments," *IEEE Transactions on Circuits and Systems for Video Technology*, 2019.
- [19] B. Wang, W. Kong, W. Li and N. N. Xiong, "A dual-chaining watermark scheme for data integrity protection in Internet of Things," *Computers, Materials & Continua*, vol. 58, no. 3, pp. 679–695, 2019.
- [20] S. Taspinar, M. Mohanty and N. Memon, "Source camera attribution using stabilized video," *IEEE Transactions on Information Forensics and Security (WIFS)*, Abu Dhabi, United Arab Emirates, pp. 1–6, 2016.
- [21] T. H. öglund, P. Brolund and K. Norell, "Identifying camcorders using noise patterns from video clips recorded with image stabilization," *International Symposium on Image and Signal Processing and Analysis (ISPA)*, Dubrovnik, Croatia, IEEE, pp. 668–671, 2011.
- [22] M. Iuliani, M. Fontani, D. Shullani and A. Piva, "Hybrid reference-based video source identification," *Sensors*, vol. 19, no. 3, pp. 649, 2019.
- [23] S. Mandelli, P. Bestagini, L. Verdoliva and S. Tubaro, "Facing device attribution problem for stabilized video sequences," *IEEE Transactions on Information Forensics and Security*, 2019.
- [24] Q. Mo, H. Yao, F. Cao, Z. Chang and C. Qin, "Reversible data hiding in encrypted image based on block classification permutation," *Computers, Materials & Continua*, vol. 59, no. 1, pp. 119–133, 2019.
- [25] Z. Liu, B. Xiang, Y. Song, H. Lu and Q. Liu, "An improved unsupervised image segmentation method based on multi-objective particle, swarm optimization clustering algorithm," *Computers, Materials & Continua*, vol. 58, no. 2, pp. 451–461, 2019.
- [26] J. Zhou, "Rolling shutter reduction based on motion sensors," U.S. Patent, no. 8,786,716, 2014.
- [27] M. Chen, J. Fridrich, M. Goljan and J. Lukáš, "Determining image origin and integrity using sensor noise," *IEEE Transactions on Information Forensics and Security (WIFS)*, vol. 3, no. 1, pp. 74–90, 2008.
- [28] D. Shullani, M. Fontani, M. Iuliani, O. Al Shaya and A. Piva, "VISION: A video and image dataset for source identification," *EURASIP Journal on Information Security*, vol. 2017, pp. 15, 2017.