Video Source Identification Algorithm Based on 3D Geometric Transformation

Jian Li[1], Yang Lv[1], Bin Ma[1*], Meihong Yang[2], Chunpeng Wang[1] and Yang Zheng[3]

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

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

[3]*University of Connecticut, Mansfield, Connecticut, 06269, USA*

Digital video has become one of the most preferred ways for people to share information. Considering people tend to release illegal information in anonymous way, the problem of video source identification attracts more and more attention as an important part of multimedia forensics. The Photo-Response Non-Uniformity (PRNU) based algorithm shows to be a promising solution for the problem of video source identification. However, it is necessary to make a geometric transformation for testing PRNU noise to align it with the reference noise, due to the effect of video stabilization. This paper analyzes the three-dimensional (3D) characteristics of camera jitters and studies how to estimate the parameters of 3D geometric transformation when aligning PRNU noises between reference and test. In the algorithm design, quaternion is used to transform PRNU noise image in 3D space, and 15 rotation axes of 3D space are estimated for experiments. 162 videos of 9 smart phones were tested. Most of the videos got a higher peak to correlation energy (PCE) value by using this algorithm, and showed better results when applied to videos with complex texture. The experiment part also records the geometric transformation parameters of the PRNU noise which need to map from image domain to video domain.

Keywords: Video forensics; Photo-Response Non-Uniformity (PRNU) noise; video source identification; video stabilization

1. INTRODUCTION

In the past decade, multimedia information security technology has attracted more and more attentions from researchers. There are various technologies proposed for the sake of data security like data hiding [1,2], watermarking [3] and digital forensics [4]. In the field of digital forensics, video source identification is still an unmet challenge.

---

*Bin Ma. Email: sddxmb@126.com
yang.zheng@uconn.edu

PRNU noise is an effective feature used to attribute a video to its capturing device. When the camera captures an image, the sensor of camera will leave PRNU noise on the image, which is caused primarily by the pixel non-uniformity of the camera sensor. With this feature, the reference PRNU noise is first estimated from a set of images taken by the camera as the fingerprint of the device, and then the fingerprint is associated with the testing PRNU noise of the anonymous video to determine whether the camera is the source device of the anonymous video.

The way to solve the problem of video source identification is to connect the media information with the device capturing it [4]. Lukas J et al. [5] proposed to use sensor pattern noise (SPN) to determine the source device. The non-uniformity in the silicon chip of the shooting equipment and

the defects in the manufacturing process of the sensor will lead to the non-uniformity of photodiode sensitivity to light, which is the cause of SPN in the video or picture. SPN has two characteristics. One is that different devices have different sensors (even the same brand and model of devices also have different sensors). So the SPN left by the sensor in the video or picture is inherent and unique, which can be used as the unique feature of connecting a video with an image or its source device; The other is that for still images, SPN is robust in compressed by JPEG [5,6], cut and scaled [7-10]. From these two characteristics, we can know that using SPN of image can solve the source identification problem of anonymous photo very well. However, due to the effect of video stabilization of current shooting equipment, anonymous video has encountered great obstacles in solving the source identification problem. Chen et al. [8] has referred to mathematical characteristics of still images, i.e., PRNU noise, to solve the source identification problem of media information. Mondaini et al. [11] used PRNU noise to identify whether there are forged frames in video frames. All these results prove the effectiveness of PRNU noise in solving the problem of source identification. Taspinar et al. [12] first attempt to use still images to get fingerprint of the device. The authors noted that the images taken by some devices without video stabilization can be associated with the unstable video captured by the device through simple scaling and clipping; Mandelli et al. [13] proposed to use particle swarm optimization algorithm to improve the accuracy of the reference PRNU noise of the device, the author's idea is mainly reflected in extracting reference PRNU noise from reference video. Due to the effect of video stabilization, each frame of video may come from different parts of the source device sensor, which will make the acquired reference PRNU noise deviate greatly from the real PRNU noise of the source device, so the device reference PRNU noise extracted from the image is more accurate than that extracted from the video; Li et al. [14] design a maximum-likelihood-estimation algorithm for extracting PRNU noise from partly decoded video frames; Iuliani et al. [15] proposed a hybrid algorithm to solve the source identification problem of video. Through a simple 2D geometric transformation of the test video frame, the PRNU noise on the video frame is aligned with the more accurate reference PRNU noise obtained from the image, which reduces the impact of video stabilization on solving the source identification problem of video.

After analyzing the slight shaking of hands when shooting video in real life, we notice that a shaking process is actually a performed in 3D space [16]. In another words, shooting video is to map the real 3D world to a mobile video frame, which is a 2D plane. The video stabilization is to reduce the impact of hand shake on the video content, so the video stabilization needs to reverse the 3D hand shake process of the 2D video [17], and then convert it into the 2D video frame process. In this case, in order to improve the success rate of solving the video source problem, we propose a video source identification algorithm based on 3D geometric transformation. The proposed algorithm mainly includes two contributions. One is to correct the PRNU noise of test video more accurately. In the algorithm, 3D geometric transformation mainly includes three operations: rotation, scaling and translation. The influence of rotation axis on correction accuracy is introduced for the first time. The other is that a specific solution is designed to reduce the complexity of the algorithm and ensure the effectiveness of the algorithm.

The rest of this paper are composed of the following parts: the second section will introduce state-of-the-art video source device identification algorithm based on PRNU noise, the third section will introduce the video source device identification algorithm based on 3D geometric transformation, the fourth section will introduce the video data set used in the experiment, record the 3D geometric transformation parameters corresponding to the connection of reference photos and test videos, and the experimental results of this algorithm are compared with the latest algorithm at this stage to verify the effectiveness and advantages of this algorithm. The summary of this paper will be given in the fifth section.

2. SOURCE DEVICE IDENTIFICATION ALGORITHM BASED ON PRNU NOISE

The main idea of source device identification algorithm based on PRNU noise is divided into three steps. The first step is to estimate the reference PRNU noise from the media information captured by the reference device. The second step estimates the testing PRNU noise of the test video by processing the test video. Finally, the normalized cross-correlation calculation (NCC) of the reference PRNU noise and the testing PRNU noise is performed to obtain the PCE value [18,19], which can be used as an evaluation index to solve the problem of video source identification. It lays a foundation for further determining the photographer's work [20,21]. As shown in Figure 1, the flow chart is the basic flow of source device identification algorithm based on PRNU noise.
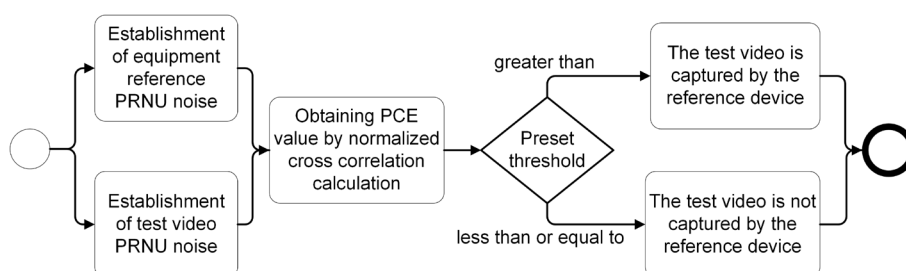
Figure 1 Basic flow of source device identification algorithm based on PRNU noise.

2.1 Estimation Method and Video Stabilization of PRNU Noise

PRNU noise is an inherent and unique feature of media information, and belongs to the pixel level zero mean multiplicative noise [5]. Because it is impossible to capture completely flat images that only contain PRNU noise, in the basic algorithm proposed by [22], N images of the same device is employed to estimate the reference PRNU noise of the device, and usually N > 50 for accuracy of estimation:

$$\tilde{K}_I = \frac{\sum_{i=1}^{N} W_I^{(i)} P^{(i)}}{\sum_{i=1}^{N} (P^{(i)})^2},$$

where $i \in \{1, 2, 3, \ldots, N\}$ represents the index value corresponding to a single image in the reference image group; $P^{(i)}$ represents the image i in the reference image group, $W_I^{(i)}$ represents the noise residual image corresponding to the reference image i. The calculation method is as follows:

$$W_I^{(i)} = P^{(i)} - F(P^{(i)}),$$

In this paper, we address the issue of the source of anonymous video, and a large part of the anonymous video is captured by amateur photographers. These photographers usually do not use professional stability auxiliary tool when capturing video. So these anonymous videos are often optimized by the video stabilization technology equipped in the capturing device for releasing slight hand shaking of photographers. The purpose of video stabilization is to improve the quality of video so that the content of each frame visually seems to move stably along a smooth path [23]. At the same time, video stabilization brings great adverse effects on solving the source problem of video. When video stabilization processes the video frame, according to the hand shake parameters fed back by the physical device of the device, it makes the corresponding geometric transformation to the frame, which results in the same coordinate position of the adjacent frames, the content is similar, but corresponding to different physical areas of the device sensor. Figure 2 shows the corresponding sensor position and scene content between the adjacent frames of video, which was captured by the device including the video stabilization technology and the device without the video's stabilization technology. Due to the influence of the video's stabilization, when estimating the PRNU noise of the test video, it is necessary to make the corresponding geometric transformation of the video frame to obtain more accurate testing PRNU noise.


(a)   (b)   (c)
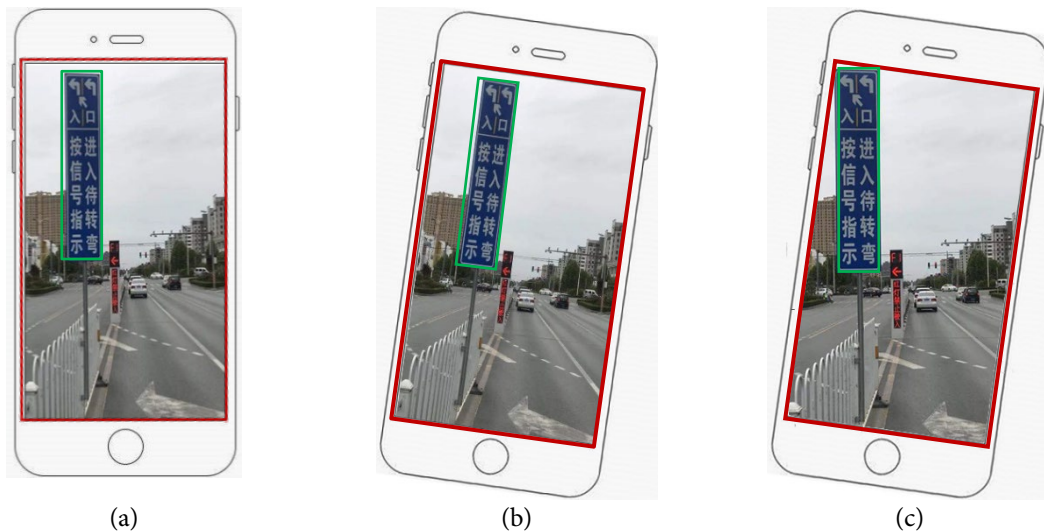
Figure 2 In the above three figures, the red box represents the position of the sensor, and the green box represents the position of the same scene content: (a) describes the scene content of frame I in the video; (b) describes the scene content of frame I+1 in the video which captured by the device without the video stabilization technology, when simulating the hand shake; (c) describes the scene content of frame I+1 in the video which captured by the device with the video stabilization technology, when simulating the hand shake.

2.2 Normalized Cross Correlation Calculation

In general, the resolution of anonymous video captured by a device can be up to 4K, while the resolution of the image captured by the same device is far greater than 4K. If we directly calculate the normalized cross-correlation between the reference PRNU noise and the testing PRNU noise, the obtained PCE value is not accurate, which will seriously affect the accuracy of solving the equipment source problem.

When capturing video, the sensor will be operated inside the device to make the sensitive part of the sensor adapt to the width-height ratio according to the requirement of video, and then the acquired image will be zoomed to the video's resolution. So, before we match the reference PRNU noise with the testing PRNU noise, we need to enlarge the testing PRNU noise from the video in equal proportion, and cut the reference PRNU noise from the image moderately. As shown in Figure 3, after a series of operations, the processed reference PRNU noise and the testing PRNU noise would be matched.
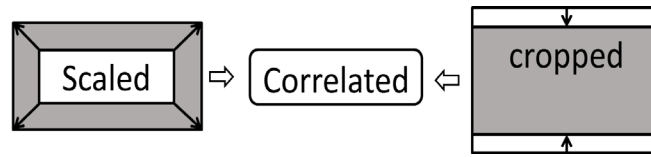
Figure 3 Refer to the pre operation before PRNU noise and testing PRNU noise match, the left is the testing PRNU noise amplification operation, and the right most is the reference PRNU noise cropping operation.

When the device reference PRNU noise and the test video PRNU noise can be basically aligned, we can use PCE value to measure the matching degree between the test video and the reference device [8], We set $m = \max(r_I, r_v)$, $n = \max(c_I, c_v)$, (r_I, c_I) is the length and width of the reference PRNU noise of the equipment after the right operation in Figure 3, respectively; (r_v, c_v) is the length and width of PRNU noise of test video after the left operation in Figure 3, respectively. PCE value is calculated in the matrix $(m \times n)$, assuming that the testing PRNU noise have been shifted (u, v) on the matrix, and then matched with the reference PRNU noise. When the displacement is (u_p, v_p), PCE value reaches the maximum. We define the correlation formula as follows:

$$R_{(u,v)} = correlate(K_{(i,j)}, K'_{(i-u,j-v)}),$$

Currently, the calculation formula of PCE is:

$$PCE(K, \ K') = \frac{R^2_{(u_p, v_p)}}{\frac{1}{m \cdot n - |\phi_p|} \cdot \sum_{u,v \notin \phi_p} R^2_{(u,v)}},$$

where ϕ_p is the small neighborhood of the peak of correlation.

When we solve the source problem of anonymous video, we can obtain reference PRNU noise $\tilde{K}_I$ of the device and test video noise residuals $\tilde{K}_v$ by the methods in Section 2.1 and Section 2.2. Normalized cross correlation calculation is carried out to obtain PCE value by $PCE(\tilde{K}_I, \tilde{K}_v)$. The PCE value is compared with the confidence threshold. If it is higher than the confidence threshold, the test video is determined to be captured by the reference device. There are two main reasons for using PCE value as the criterion to determine the attribution of anonymous video, one is the validity of PCE value, the other is that the PCE value is also effective when the parameters are transformed to a certain extent, and it has better robustness.

3. VIDEO SOURCE IDENTIFICATION ALGORITHM BASED ON 3D GEOMETRIC TRANSFORMATION

As we all know, every movement of people in real life, regardless of time dimension, are in 3D space, including hand shaking. The video stabilization of the capturing device mainly aims at the optimization and correction of video content when the hand shake amplitude is small. It relies on some sensors of the device to record the manual shake amplitude, direction and other parameters, which cannot be obtained in the case of only anonymous video, so we can only use brute force search to carry out these hand shake parameters. It is estimated that in [13-15], hand jitter is considered as a motion in 2D space, so when matching the device reference PRNU noise with the test video PRNU noise, only the test video PRNU noise is corrected in 2D space to align the device reference PRNU noise, which is obviously inaccurate. Hence we employ 3D geometric transformation to accurately estimate the movement between the reference and testing PRNU noises.

3.1 3D Geometric Transformation

The video stabilization of the device makes the same coordinate position between the video frames correspond to different positions of the device sensor, which makes the calculated PRNU noise of the test video deviate from the reference PRNU noise of the source device to a certain extent. Therefore, in order to obtain more accurate PRNU noise of the test video, we need to carry out geometric transformation of the noise residual image of the video frame as the PRNU of the test video frame noise:

$$\tilde{K}_v^{(i)} = T_{\alpha, Ra}^{(i)}(P^{(i)} - F(P^{(i)})),$$

where $P^{(i)}$ is the frame i of the video, $T_{\alpha, Ra}^{(i)}$ is the rotation matrix corresponding to the frame $P^{(i)}$, the rotation angle is α, and the corresponding rotation axis is Ra.

Rotation in 3D geometric transformation needs to determine two parameters: rotation angle and rotation axis. For the rotation angle, we refer to the data in [15], and set the rotation angle α range to [-1.5°, 1.5°]; there are infinite rotation axes in 3D space, so in order to reduce the complexity of the algorithm, we need to select a representative finite number of rotation axes. The external features of the mobile phone have obvious characteristics, the length is greater than the width, and the two are far greater than the thickness, so when selecting the rotating axis, we can regard the device as a 2D plane to simplify the selection process of the rotating axis.

For the anonymous video with the resolution of 1080×1920, we mark the three bisection points in the row direction and two endpoints in the column direction as the bisection points. As shown in Figure 4, we selected any two points to connect into a line as the rotation axis to be tested. After removing the parallel axis, only the following 15 vectors are left in the rotation axis to be tested: (-1080,640,0), (-1080,1280,0), (-1080,1920,0), (-540,640,0), (-540,1280,0), (-540,1920,0), (540,640,0), (540,1280,0), (540,1920,0), (1080,640,0), (1080,1280,0), (1080,1920,0), (1,0,0), (0,1,0), (0,0,1).

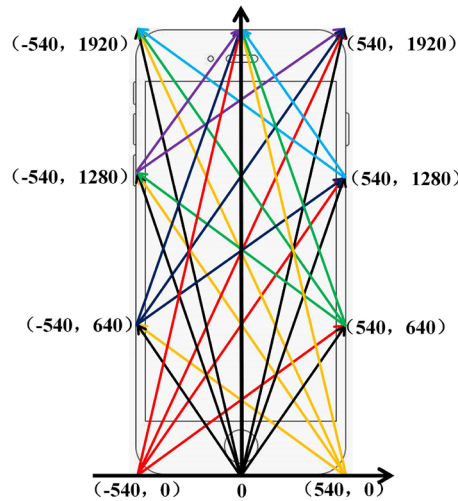Let the 15 vectors form the rotation axis vector set R.

Figure 4 Equidistant points are connected into a line as the axis of rotation.

3.2 Specific Solutions

After introducing the geometric transformation of 3D space, the complexity of the algorithm is greatly increased. We design the following scheme to reduce the complexity of the algorithm, while maintaining the effectiveness of the algorithm.

3.2.1 Quaternions

In order to make the algorithm spend less time and space, we choose to use quaternion to represent the geometric transformation of 3D space [24]. Quaternion is a simple hypercomplex number, which consists of a real part and three imaginary parts, which can be expressed as $Q = a + bi + cj + dk$, where a, b, c, d are all real numbers. For its own geometric meaning, it can be understood as a kind of rotation, wherein rotation i represents the positive rotation of x-axis to y-axis in the intersection plane of x-axis and y-axis, rotation j represents the positive rotation of z-axis to x-axis in the intersection plane of x-axis and z-axis, rotation k represents the positive rotation of y-axis to z-axis in the intersection plane of y-axis and z-axis, in addition that -i, -j, -k represents the reverse rotation of rotation i, j, k.

In this paper, we set the quaternion format of 3D rotation as follows:

$$Q(\alpha, Ra) = \cos(\alpha) + \sin(\alpha) \times Ra(x) \times i + \sin(\alpha) \times Ra(y) \times j + \sin(\alpha) \times Ra(z) \times k$$

where Ra is the vector corresponding to the normalized rotation axis, $Ra(n)$ represented the component of the vector Ra in the n axis direction, and $Ra \in R$.

There are three advantages of using quaternion to transform PRNU noise in three dimensions:

a) The 3D geometric transformation based on quaternion provides a spherical linear interpolation method, which cannot be provided by other 3D geometric transformation algorithms;

b) The quaternion cross multiplication can transform the angular displacement sequence into a single angular displacement, and it is obviously slower to do the same operation with the matrix;

c) The quaternion contains only four numbers, and the matrix uses nine numbers, which can reduce the time and space of the algorithm.

3.2.2 Block Matching

After solving the problem of rotation, we turn our attention to the matching process, hoping to reduce the complexity of the algorithm by optimizing the matching process. It is known that the existing algorithm is to translate the video frame after geometric transformation so that it can match the reference PRNU noise better. Because the resolution of the video frame can also reach 4K, the time complexity of the matching process is higher。 In view of the video frame can also see an image, we can learn from the image segmentation algorithm to improve the efficiency of this algorithm.

Block algorithm [25] is to divide an image into different parts according to certain rules, select the parts we are interested in for operation processing, maintaining the internal information of the image and ensure a certain time complexity; literature [7, 8] proposed that SPN is not sensitive to clipping, which shows that a certain degree of clipping operation for testing PRNU noise during matching does not affect the final result and, on the other hand, greatly improve the matching efficiency of the algorithm.

As mentioned in Section 2.2, when the PRNU noise of the test video matches the reference PRNU noise of the device, the reference PRNU noise needs to be partially clipped. As shown in Figure 5, in order to reduce the complexity of the algorithm, we use some still frames in the video to determine the part of the frame that needs to be clipped for the reference PRNU noise, because the corresponding clipping parameters of different frames in different videos of the same device are invariable.
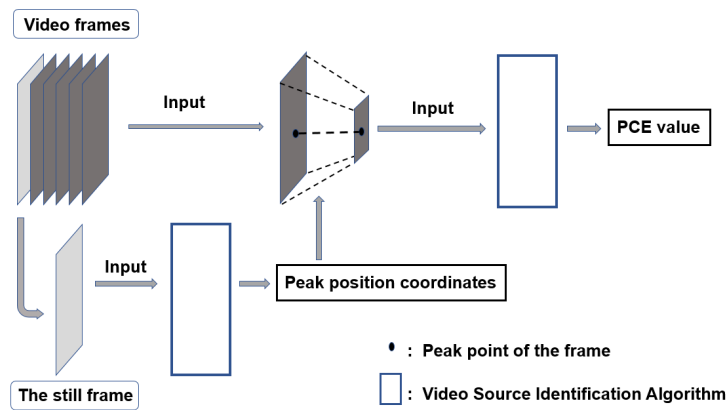
Figure 5 The above figure displays the basic process of block algorithm. In order to obtain the accurate peak position coordinates, we use the still frame to input into the algorithm in advance. Then a smaller block is obtained from the rest frame of the video according to the peak position. Finally, the smaller frame block is input into the algorithm to get the final PCE value.

4. EXPERIMENT

This section mainly displays and analyzes the experimental part. First, we introduce the environment and video database used in our experiment, and then illustrate and summarize the experimental process, data and results.

4.1 Database and Environment Introduction

In order to promote the development of multimedia forensics technology, video data set is proposed in reference [26]. We test the algorithm proposed in this paper on a data subset of this data set to verify the effectiveness and advantages of this algorithm. This subset contains 1064 smooth images and 162 videos captured by 9 devices. The hardware platform used in the experiment is inter (R) core (TM) i7-6700 CPU, frequency 3.41ghz, memory capacity 16GB, operating system 64 bits windows 10, simulation software MATLAB r2018b.

Specifically, first, we need to build a reference PRNU noise for each device as the fingerprint of the device. We select all available images captured by the device, and the content of these images is relatively smooth. It can be known from Eq. (1)

that when the number of images is greater than 50, the more images, the more accurate the data, so we need as many images as possible. For anonymous test video, we choose a device whose video resolution is equal to High Definition (1920 × 1080 pixels), and consider static (video labeled "still") and dynamic scenes (video labeled "pan rot" and "move"). In addition, in order to make the experiment more complete, the experimental data also includes videos with almost flat content (videos marked as "flat") and videos with obvious textures (videos marked "indoor" and "outdoor"). For clarity, we provide Table 1 that provides basic information about the devices used and the videos.

4.2 Estimation of Matching Parameters

It can be seen from chapter 2.2 that the testing PRNU noise extracted from the video frame needs to be matched with the reference PRNU noise after being amplified. Set the length and width of the reference PRNU as (R, c), and it is known that the same size of the testing PRNU noise and the video frame is (1080 × 1920):

$$scale\_\max = (\min(r/1080, c/1920))^{-1},$$

Table 1 Information of equipment and media data used in the experiment.

Device ID	Equipment Model	Number of flat images	Number of videos
D02	iPhone 4s	103	13
D05	iPhone 5c	113	19
D06	iPhone 6	149	17
D10	iPhone 4s	133	15
D14	iPhone 5c	130	19
D15	iPhone 6	110	18
D18	iPhone 5c	101	13
D20	iPad Mini	119	16
D34	iPhone 5	106	32

We use the method of violent search to select the most appropriate scale, and its range is [1, *scale*_max]. When we match the most appropriate scale in the experiment, we use the parameter t to narrow the scale, that is to say, the scale range is reduced to:

$$[best\_scale - t \times (scale\_\max - 1), \quad scale\_\max + t \times (scale\_\max - 1)]$$

and in our experiments, we set $t = 0.2$.

We use all the smooth images captured by each device to estimate the reference fingerprint of the device, and then use the smooth and still video captured by the device as the test video to estimate the clipping parameters. In order to reduce the time complexity of the algorithm, we divide the testing PRNU noise into 3×3 blocks, and take the middle block as the representative to match with the device reference PRNU noise. When the PCE value reaches the maximum value, the corresponding peak coordinate can be used as the reference cropping position for the next detection of other videos. Table 2 shows the peak position, scaling ratio and rotation angle of each device recorded during the test of 9 smartphones.

Table 2 Geometric transformation information mapped from image domain to video domain.

| Device ID	Cropping Position	Scaling Ratio	Rotation Angle	Axis of Rotation
D02	(988,1206)	[0.748,0.752]	-1.3°	[540 ,1920, 0]
D05	(963,1167)	[0.681,0.690]	-0.3°	[0 ,1920, 0]
D06	(968,1187)	[0.697,0.713]	-1.3°	[0 ,1920, 0]
D10	(976,1206)	[0.748,0.752]	1.5°	[0 ,1920, 0]
D14	(962,1167)	[0.683,0.689]	-1.3°	[0 ,1920, 0]
D15	(973,1172)	[0.703,0.708]	1.5°	[-540 ,1920, 0]
D18	(962,1167)	[0.684,0.689]	-1.5°	[-540 ,1920, 0]
D20	(750,904)	[0.813,0.816]	1.3°	[-540 ,1920, 0]
D34	(962,1168)	[0.684,0.691]	0.9°	[0,0,1]

4.3 Comparison of Experimental Results

The experimental design of this paper is divided into two parts: the intra class experiment and the inter class experiment. The intra class experiment uses the video captured by the reference device as the test video, and the inter class experiment uses the video captured by other devices as the test video. Obviously, if the algorithm is effective, a threshold value can be selected between the maximum PCE value of the intra class experiment and the inter class experiment. According to the threshold value, it can be determined whether the test video is captured by the reference device.

We experimented with 162 videos on 9 devices, including more than 100 videos in the inter-class experiment, almost all kinds of situations are fully considered.

The specific PCE value comparison results are shown in Figure 6. It can be clearly seen that for most videos, the PCE value of 3D geometric transformation is much larger than that of 2D geometric transformation.
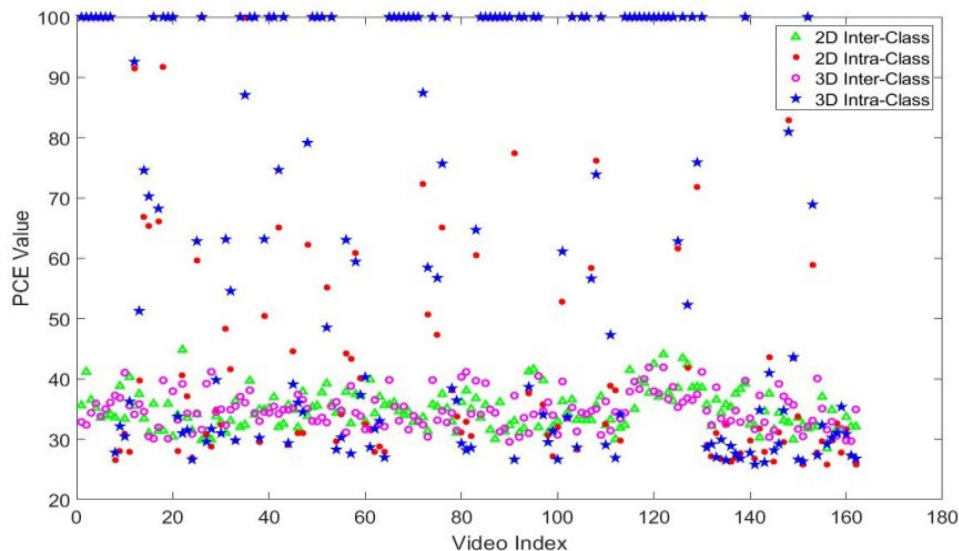
(a)
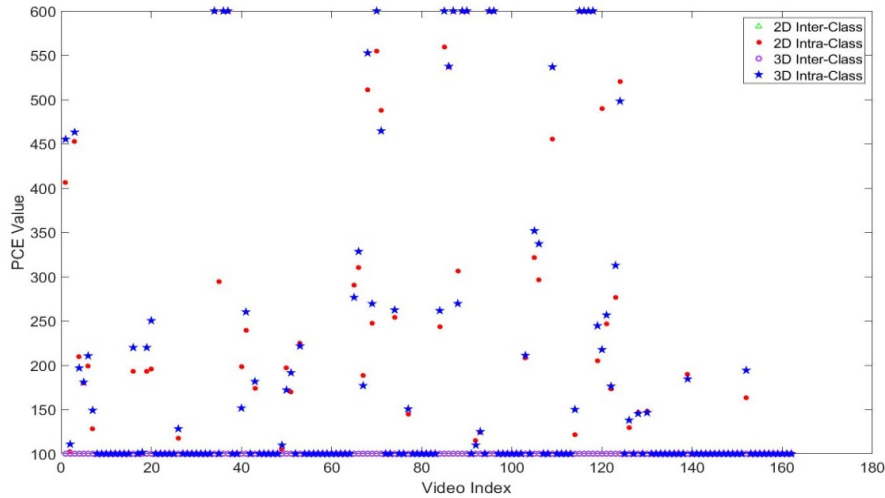
(b)

Figure 6: The points in figure (a) represent the maximum PCE value of each video with PCE value between [0,100] after 2D geometric transformation and 3D geometric transformation, respectively. The points in figure (b) represent the maximum PCE value of each video with PCE value between [100,600] after 2D geometric transformation and 3D geometric transformation, respectively.

We also consider the situation that anonymous video is attacked, for example, anonymous video is cropped. At this time, the video we can get may be any part of the original video frame, which will bring great challenges to the algorithm. In the experiment, we only select 10 frames in the video, and in order to compare with the algorithm of 2D geometric transformation, the selection of 10 frames is not random, but every 15 frames (excluding the first frame). The experimental results are shown in Table 3. The True Positive Rate (TPR) of algorithm based on 3D geometric transformation is slightly higher than that based on 2D geometric transformation, when the False Positive Rate (FPR) = 0.

Table 3 The TPR of the experience with different geometric transformation.

Test videos	Geometric transformation dimension	TPR	FPR
Stabilized	2D	0.537	0
Stabilized	3D	0.567	0

5. CONCLUSION

In this paper, we propose a video source device authentication algorithm based on 3D geometric transformation, which uses the reference PRNU noise from still image as the fingerprint of the device, and obtains the testing PRNU noise from the anonymous video frame. Before the two are matched, the testing PRNU is transformed in 3D space to improve the matching accuracy. To ensure the complexity of the algorithm, we estimate 15 rotation axes and use block algorithm on the PRNU noise image. Experiments on 162 stable videos of 9 kinds of devices show that the video source identification algorithm based on 3D geometric transformation is more effective than that based on 2D geometric transformation. In the next work, we will focus on further reducing the complexity of the algorithm, for example, by improving the efficiency of the PRNU noise matching process to reduce the complexity of the algorithm, etc.

REFERENCES

1. 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.

2. R. Das, M. Baykara and G. Tuna, "A novel approach to steganography: enhanced least significant bit substitution algorithm integrated with self-determining encryption feature," *Computer Systems Science and Engineering*, vol. 34, no. 1, pp. 23–32, 2019.

3. 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.

4. S. Milani, M. Fontani, P. Bestagini, M. Barni, A. Piva et al., "An overview on video forensics," in *Proc. 20th European Signal Processing Conference (EUSIPCO)*, Bucharest, 2012, pp. 1229-1233.

5. 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.

6. Y. Wu, W. Shang and J. Chen, "Anti-JPEG compression steganography based on the high tense region locating method," *Computers, Materials & Continua*, vol. 59, no. 1, pp. 199-214, 2019.

7. F. Bertini, R. Sharma, A. Iannı and D. Montesi, "Social media investigations using shared photos," in *The Int. Conf. on Computing Technology, Information Security and Risk Management (CTISRM2016)*, 2016.

8. 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. in *Proc. of SPIE - International Society for Optics and Photonics*, San Jose, CA, United States, pp. 6505, 2007.

9. 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 (Early Access)*, 2019.

10. K. Rosenfeld, H. T. Sencar *et al.,* "A study of the robustness of PRNU-based camera identification," in *Proc. SPIE-The International Society for Optical Engineering*, San Jose, CA, USA, 7254, 2009.

11. N. Mondaini, R. Caldelli, A. Piva, M. Barni and V. Cappellini, "Detection of malevolent changes in digital video for forensic applications," Security, Steganography, and Watermarking of Multimedia Contents IX. in Proc. of SPIE - International Society for Optics and Photonics, San Jose, CA, United States, pp. 6505, 2007.

12. S. Taspinar, M. Mohanty and N. Memon, "Source camera attribution using stabilized video," in *IEEE International Workshop on Information Forensics and Security (WIFS)*, IEEE, Abu Dhabi, United Arab Emirates, 2016.

13. S. Mandelli, P. Bestagini, L. Verdoliva and S. Tubaro, "Facing device attribution problem for stabilized video sequences," *IEEE Transactions on Information Forensics and Security*, 15, 14-27, 2019.

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

16. S. Akbarov and M. Mehdiyev, "The interface stress field in the elastic system consisting of the hollow cylinder and surrounding elastic medium under 3D non-axisymmetric forced vibration," *Computers, Materials & Continua*, vol. 54, no. 1, pp. 61-81, 2018.

17. F. Liu, M. Gleicher, H. Jin and A. Agarwala, "Content-preserving warps for 3D video stabilization," *ACM Transactions on Graphics*, vol. 28, no. 3, pp. 1–9, 2009.

18. M. Goljan, J. Fridrich and T. Filler, "Large scale test of sensor fingerprint camera identification," in *Media Forensics and Security International Society for Optics and Photonics*, San Jose, CA, USA, 2009.

19. S. S. Li, T. J. Cui and J. Liu, "Research on the clustering analysis and similarity in factor space," *Computer Systems Science and Engineering*, vol. 33, no. 5, pp. 397-404, 2018.

20. C. Yu, Q. M. Hong, D. Z. Yao and H. Jin, "Tensor-based user trajectory mining," *Computer Systems Science and Engineering*, vol. 33, no. 2, pp. 87-94, 2018.

21. W. Y. Liu, Z. J. Liu, L. Wang, B. B. Li and N. Jing, "Human movement detection and gait periodicity analysis via channel state information," *Computer Systems Science and Engineering*, vol. 33, no. 2, pp. 137-147, 2018.

22. M. Chen, J. Fridrich, M. Goljan *et al.,* "Determining image origin and integrity using sensor noise," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 1, pp. 74-90, 2008.

23. M. Grundmann, V. Kwatra and I. Essa, "Cascaded camera motion estimation, rolling shutter detection, and camera shake detection for video stabilization," U.S. Patent 9,374,532[P]. 2016-6-21.

24. J. C. Hart, G. K. Francis and L. H. Kauffman, "Visualizing quaternion rotation," *ACM Transactions on Graphics*, vol. 13, no. 3, pp. 256-276, 1994.

25. 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.

26. D. Shullani, M. Fontani, M. Iuliani *et al.,* "VISION: a video and image dataset for source identification," EURASIP Journal on Information Security, vol. 2017, 1, pp. 15, 2017.