

Three-dimensional Region Forgery Detection and Localization in Videos

Xuan Hau Nguyen^{*1,2}, Yongjian Hu, Muhmmad Ahmad Amin and Khan Gohar Hayat

¹Research Centre of Multimedia Information Security Detection and Intelligent Processing,
School of Electronics and Information Engineering,
South China University of Technology, Guangzhou 510640, P.R.China.
Email: nguyenxuanhau@tic.edu.vn, eejhu@scut.edu.cn, ahmad.242@live.com, g.hayat@yahoo.com

Van Thinh Le

²Faculty Electronics of and Informatics Engineering
Mien Trung Industrial and Trade College, Phu Yen 620000, Vietnam
Email: levanthinh@tic.edu.vn

Dinh Tu Truong

Faculty of Information Technology
³Ton Duc Thang University, Ho Chi Minh 700000, Vietnam
Email: truongdinhtu@tdtu.edu.vn

Received: 29 September 2019; Accepted: 28 October 2019; Published: 08 December 2019

Abstract—Nowadays, with the extensive use of cameras in many areas of life, every day millions of videos are uploaded on the internet. In addition, with rapidly developing video editing software applications, it has become easier to forge any video. These software applications have made it challenging to detect forged videos, especially with forged videos have duplication of three-dimensional (3-D) regions. Recently, there has been increased interest in detecting forged videos, but there are very limited studies to detect forged videos which were duplicated 3-D regions. So, our research focused on this weakness and proposed a new method, which can be used for detecting and locating 3-D duplicated regions in videos based on the phase-correlation of 3-D regions residual more efficiently. To evaluate the efficiency of the proposed method, we experimented with two realistic datasets VFDD-3D and REWIND-3D. The results of the experiments proved that the proposed method is efficient and robust for detecting small 3-D regions duplication and frame sequences duplication, especially localization of duplication forgery in videos has shown impressive results.

Index Terms—Passive forensics, three-dimensional regions duplication, video forensics, video forgery detection, video authenticity

I. INTRODUCTION

Nowadays, smartphone, camcorder, and security camera are extensively used in many areas of daily life such as traffic lights, offices, houses, dormitories and lot more which are monitored by cameras. Besides that,

video content editing software like Video Editor, Adobe Photoshop, Window Movie Maker, and Adobe After Effect also are easily available. They support a lot of methods for editing video content easily, and anyone can edit video contents at their willing even edited content contrast with original content, which leads to seeing is no longer believing [1]. In addition, an authentic video gives strong evidence more than an authentic image in court. Therefore, video forensic proves that video authenticity becomes an urgent requirement today. So, nowadays, video forensic has become a hot topic of interest among researchers around the world.

The video forensic methods are divided into active and passive methods. Active methods use given information such as Watermarking or Signature, which is inserted into the video, then this information is checked. If it does not change, that video is authentic otherwise forged. Meanwhile, passive methods only analyze video content to find traces of forgery. Now, most of the videos do not usually insert given information, so the passive methods have become a hot topic that has attracted researchers. The passive methods can be classified into three categories which are Video Inter-Frame Forgery, Double/Multiple Compression, and Region Tampering [1].

Throughout state-of-the-art methods, 3-D region forgery detection and localization of forgery were not an area of interest to many researchers in the past, but nowadays with the steady development of video editing applications, 3-D regions in videos are copied then pasted to other position and edited the brightness, geometry, and similar things have become easy. These 3-D regions can be small 3-D areas inside frame sequences as shown in Fig. 1, in the four-right side pictures of this figure, a

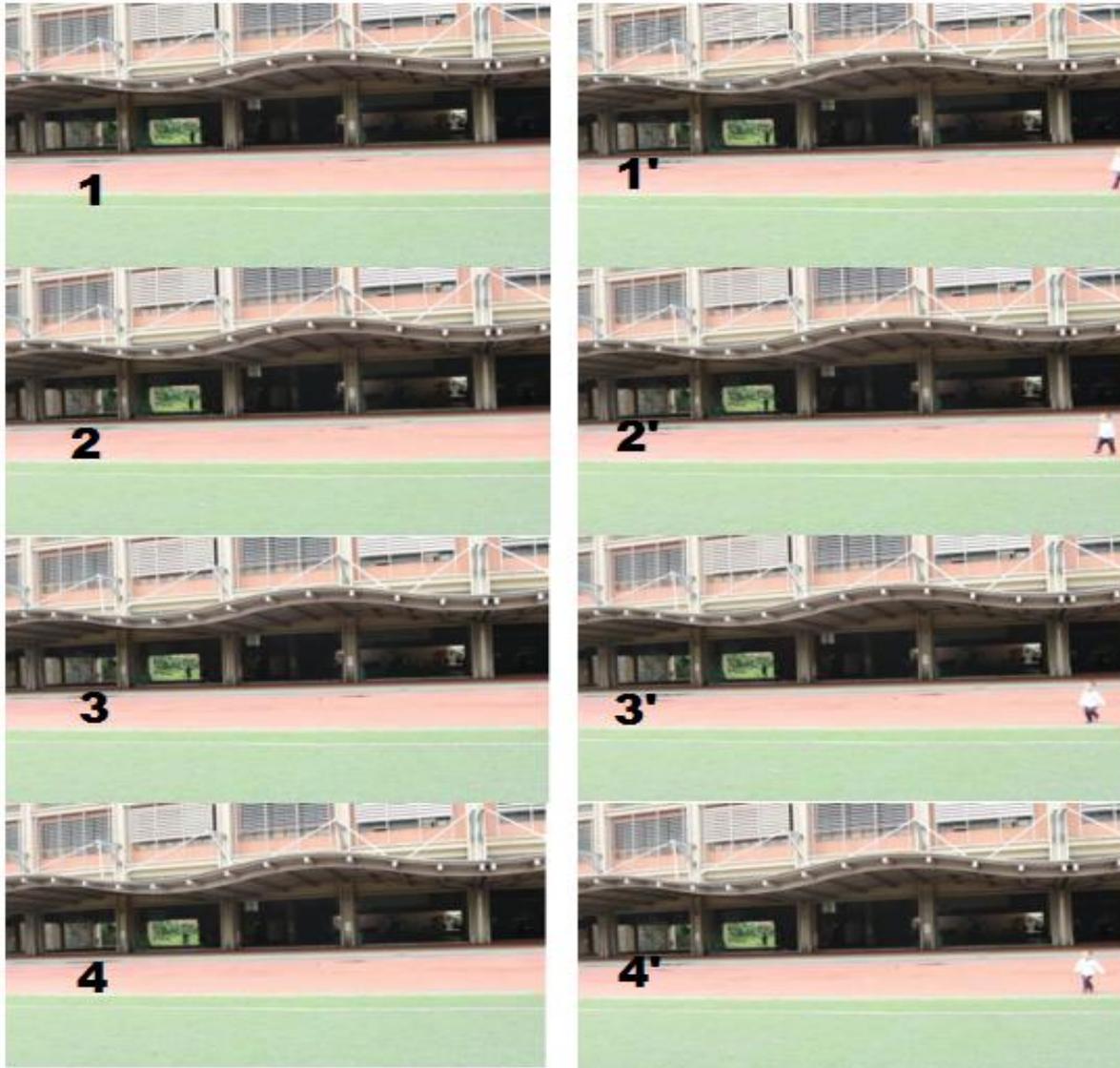


Fig. 1. Four left pictures (1 to 4) are taken from an original video; four right pictures (1' to 4') are taken from a forged video

small 3-D region covering the child be copied/pasted across consecutive frames. Besides, this 3-D region also can be full-frame sequences as shown in Fig. 2, the four consecutive frames on the right side were copied/pasted to hide men in the original video. These have become a popular tampering method used in video tampering and it is very difficult to detect. So, it has been urging us to study for finding algorithms to detect this kind of video forgery. Until now, Insufficient research work has been done about duplication of small 3-D regions in videos. As an example, in [2] the authors used a correlation between fixed regions of frame sequence and full video, which causes immense computational complexity and detection accuracy isn't high when 3-D region duplication is small. So, in this study, we propose a method for efficiently detecting duplication of 3-D regions in videos.

The proposed algorithm uses phase-correlation of frame residual for detecting duplication of 3-D regions and localization in videos. The proposed algorithm has shown improved accuracy when compared with other schemes on the same datasets. To evaluate the

performance of the algorithm, we have built our test dataset basing on the VFDD1.0 dataset from our lab [3] which contain more than 640 videos got from ground truth with 12 different cameras. In this dataset, the videos have high-resolution diversity from [404 x 720 pixels] to [1080 x 1920 pixels], by using 10 different cameras and these videos format is almost compression format (MP4). Besides that, we have used and expanded the REWIND dataset which is the popular dataset for detecting small 3-D regions duplication available online until now [4], but this dataset has some shortcoming such as low-resolution, captured from only 3 cameras, in one video which contains duck can be spliced rather than copy-move because the original and forged video have a difference of scenery.

The rest of the paper is organized as follows. In section II, we discuss related methods, and in section III we describe the theory and proposed method. Experimental results are given in section IV. Result analysis and comparison are given in section V, and section VI contains conclusions and future directions.

II. RELATED WORK

The passive methods are divided into three categories which are Video Inter-Frame Forgery, Double/Multiple Compression, and Region Tampering [1]. The methods of Video Inter-Frame Forgery detect manipulated frames such as Frame Insertion, Frame Deletion, Frame Duplication, and Frame Shuffling, which have been researched by many researches. As an example, in [5], the authors used differences in correlation coefficients of grey values between sequential frames to detect frame deletion and frame insertion. In [6], the authors calculated the spatial and temporal correlation to detect duplication of frame sequences. In [7], they used the correlation of SVD features of frame sequences to detect duplication of frame sequences. In [8], the authors used the correlation coefficient of sequences of frames DCT means. In [9, 10], used optical flow to detect frame insertion and frame deletion. In [11], the authors used Zernike opponent chromaticity moments to detect frame insertion, frame deletion, and duplication of frame sequences.

Meanwhile, the second methods detect traces left by double/multiple compressions. Because a video is usually saved in compression after recording, so a forged video is compressed again after editing. These methods depend on the codec of video, so it is a little attractive to the researchers. Having some studies of this method, such as [12] used statistic pattern in the distribution of DCT

coefficients to exploit double compression of frame I. In [13], the authors estimated pixel value in GOP. If there is an error between the actual value and the calculated value, the video may be compressed several times. In [14], Markov statistics of compression, noise is used to detect double compression, and in [15], with assuming that the second compression has same parameters with the first time of compression at which most of the methods fail, and the basic idea is, when a frame is recompressed with the same quantization matrix again. The number of different DCT coefficients between the sequential two versions will monotonically decrease.

The third methods, the Region Forgery Detection methods, exploit abnormal pixel values in a video to detect video forgery. Typical methods as follow, In [16] calculated motion vector for detecting object deletion in videos. In [17] used temporal noise to detect two-dimensional regions copy-move. In [18, 19], the authors used features (Histogram of Oriented Gradients - HOG, Scale Invariant Feature Transform -SIFT) to detect two-dimensional regions copy-move. In [2] residual of consecutive frames is used for detecting copy a two or three-dimensional region and paste it in many times. Among these approaches, the method is used in [2], which is most close to us but it split the sequence into fixed three-dimensional regions then calculating correlation value between regions. So, sometimes this leads not found forgeries in the video.



Fig. 2. Four left pictures (1 to 4) are taken from an original video; four right pictures (1' to 4') are taken from a forged video

III. PROPOSED METHOD

In this paper, we proposed a method to detect one of the usual kinds of video tampering to hide or duplicate objects and actions in videos. That is, 3-D regions are

copied then pasted into another location in the same video. This 3-D region can be a full-frames sequence or a small 3-D region inside of a frames sequence, as shown in Fig. 3.

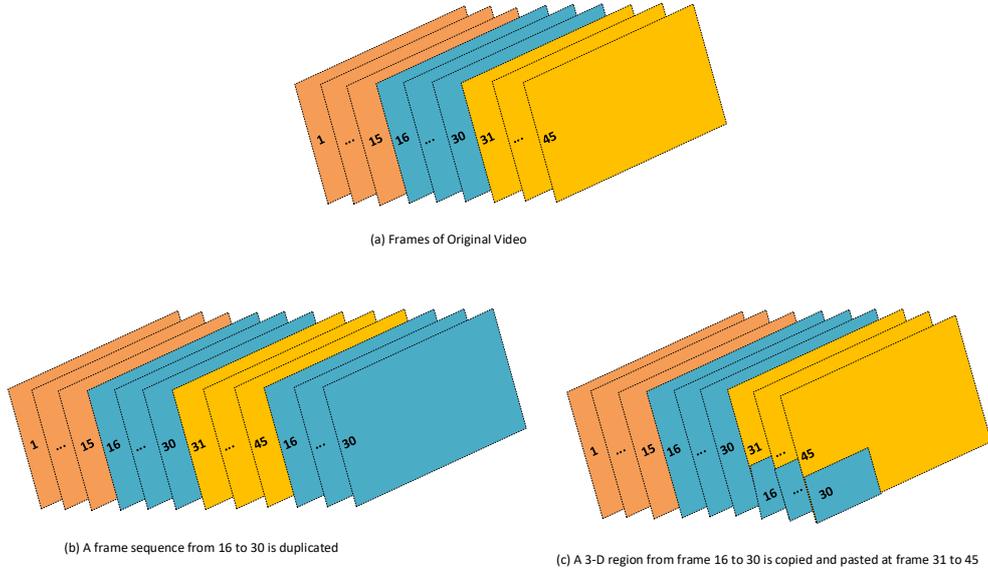


Fig. 3. 3-D regions duplication in video

In **Fig. 3**. A video has been presented by consecutive frames sequence. Fig. 3(a) is an original video with 45 frames. Fig. 3(b) the frames sequence from frame 16 to 30 is copied and placed next to the 45th frame in the same video. Fig. 3(c) a 3-D region (blue) is copied and pasted at frame 31 to 45 in the same video. Notably, this 3-D region shape can be arbitrary, but for easy understanding, we represented it as a rectangular shape. These kinds of attack are usually implemented, and there are efficient to hide or duplicate objects and actions in a video.

Nowadays, with the robust development of the software industry, video editing applications support powerful video editing. Therefore, duplication of 3-D regions becomes easier. Also, the pasted 3-D region could be changed in brightness or geometry. 3-D regions duplication becomes popularly and effectively in cases duplicating object, hiding objects or actions in videos. So, it is crucial to propose methods for 3-D regions duplication forgery detection in videos effectively.

Previous researches proved conclusively phase-correlation in calculating the correlation [2, 6]. In this research, we have researched then proposed the algorithm using phase-correlation of frame residual to detect duplication of 3-D regions. Main steps of the algorithm are given as follows:

The frames of suspected video are extracted and converted into grayscale. Therefore, the video after extracting and converting is denoted as follows:

Step 1: Extract and convert video into grayscale.

Let $V = \{v_{i,j}^t\}$ is a video extracted and converted to grayscale.

Where,

$t \in [1, T - 1]$, T is the number of frames in the video.
 $i \in [1, I]$, I is the height of the frame.
 $j \in [1, J]$, J is the width of the frame.

Step 2: Calculate

$$R = r_{(i,j)}^t \quad (1)$$

Where, $r_{(i,j)}^t = v_{(i,j)}^{(t+1)} - v_{(i,j)}^t$, $r_{(i,j)}^t$ is a frame residual is the difference between two adjacent frames.

Here, using R instead V , because R eliminates the linear effects (e.g. brightness change), and to enhance a pixel-to-pixel difference ratio when performing Fourier transformations in the following steps.

Step 3: Divide video become 3-D regions, as shown in Fig. 4 and calculate Fourier transformation F (3-D region) of all 3-D regions.

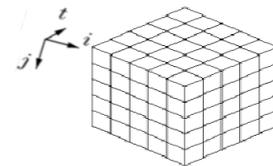


Fig. 4. A video is divided into 3-D regions

Step 4: Calculate the maximum of phase-correlation of two 3-D regions.

$$K = \max(P(B_n, B_m)) = F^{-1} \left(\frac{F(B_n)F^*(B_m)}{\|F(B_n)F^*(B_m)\|} \right) \quad (2)$$

Where, B_n and B_m are 3D regions, the size of which is determined in experiments. $F(B_n)$, $F(B_m)$ are the Fourier transforms of 3D regions B_n and B_m , $F^*(B_m)$ is

complex conjugate of $F(B_m)$, $\| \quad \|$ is complex magnitude and F^{-1} Inverse Fourier transforms.

Step 5: For concluding B_n, B_m regions are suspected duplicate or not; we performed the following operations as shown in Fig. 5:

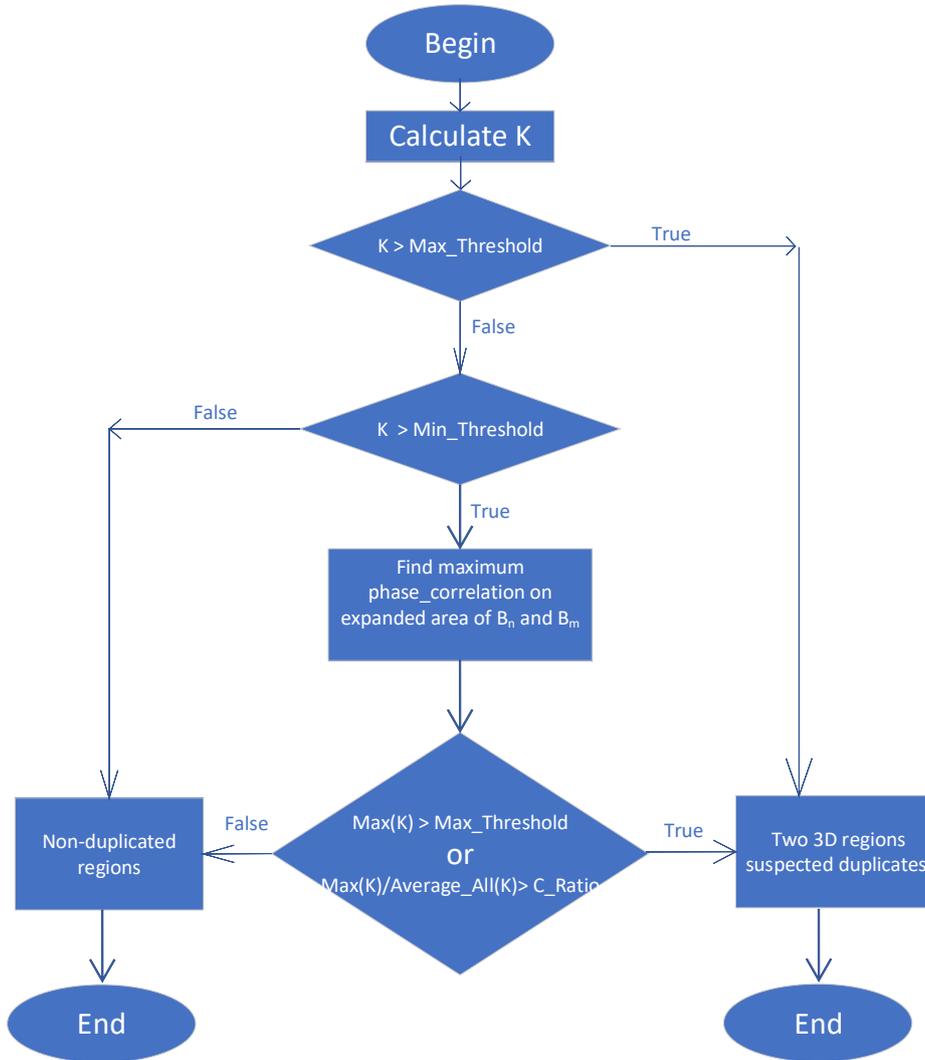


Fig. 5. Operations identify 3D suspected regions

K is the maximum value of phase-correlation between B_n and B_m .

$Max_Threshold$ is an above threshold value of phase-correlation. If K is greater than $Max_Threshold$, B_n and B_m are suspected duplicates. $Max_Threshold$ value at least is 0.55 in our experiments.

$Min_Threshold$ is a low threshold value of phase-correlation. If K is less than or equal to $Min_Threshold$, B_n and B_m are not duplicated. $Min_Threshold$ is a small value nearly 0, and in our experiments, $Min_Threshold$ is 0.3.

If $K > Min_Threshold$ and $K \leq Max_Threshold$, B_n and B_m are also possible duplicated regions. So, need to find the maximum of phase-correlation of B_n and B_m regions on expanded areas (by striding 3 pixels) denoted $Max(K)$.

The expanded areas are extended each of the direction of these 3D regions by one-half the size of the regions.

$Average_All(K)$ is an average value of all K of the video.

C_Ratio is a threshold value of a ratio between $Max(K)$ and $Average_All(K)$. C_Ratio is usually a large value to indicate abnormal differences compared between this phase-correlation value and average value phase-correlation of the video. Therefore, the ratio between $Max(K)$ and $Average_All(K)$ is greater than C_Ratio ; this 3-D regions pair is suspected duplicate otherwise not. In our experiments, C_Ratio is 20.

Step 6: Determine and localize the duplicated areas.

Each suspected regions pair is tiled into $8 \times 8 \times 10$ overlapping (by striding 1 pixel) blocks, and the

calculating correlation coefficient between each block pair in these two duplicated regions respectively by vectoring the blocks then using the equation (3). Any blocks whose correlation coefficient is greater than 0.9 (in our experiments) are considered duplicated each other otherwise not.

$$C(\vec{x}, \vec{y}) = \frac{\sum_i (x_i - \mu_x)(y_i - \mu_y)}{\sqrt{\sum_i (x_i - \mu_x)^2 \sum_i (y_i - \mu_y)^2}} \quad (3)$$

Where x_i and y_i are the i^{th} elements of \vec{x} and \vec{y} , and μ_x and μ_y are the respective means of \vec{x} and \vec{y} .

To remove the blocks do not belong to the duplicated area, we used the connectivity of four between blocks which were duplicated.

IV. EXPERIMENTAL RESULTS AND COMPARISON

To evaluate the effectiveness of the proposed algorithm, experiments are performed on VFDD-3D and REWIND-3D datasets. In the following, we will describe these datasets and experimental results.

1. Dataset

Nowadays, video forensic is a hot topic in which researchers are interested. Recently, there are many algorithms have been proposed. But to evaluate the efficiency of algorithms, most of which were only based on a database of authors themselves. So, until now there are only available a few datasets on the internet as follows: The SUFA dataset contains 150 original videos, but only five videos are forgery videos by duplicating small 2-D regions, and it has low-resolution videos captured from only three cameras[4]. The REWIND dataset [20] contains 20 videos (10 original, 10 forged), forgery videos by duplicating small 3-D regions which have been extended from some original videos in SUFA, and these videos also have low-resolution only 240 x 320 pixels. In [21], the authors built a VTD dataset, but original videos were taken from the internet, so the authenticity of the videos are not determined. Recently, our lab has designed a VFDD dataset for evaluating video forensic methods. The VFDD2.0 has contained 1500 original videos gotten from ground truth with 27 different cameras, and 695 forged videos which have made from some different kinds of forgery methods and published at [3] and this database will be modified extended larger and more complete, especially forgery videos in the next time. In the following, we will present datasets which were prepared and used to evaluate the proposed algorithm. One thing is noticeable here that, it is very difficult for the human eye to visualize any forgery in forged videos.

a. The first dataset (VFDD_3D Dataset)

We took original and some their forged videos which are forged by duplicating 3-D regions from the VFDD dataset. Moreover we had tried to tamper some original

videos by duplicating 3-D regions of this dataset to build VFDD_3D dataset in which the most of videos have high-resolution diversity from [404 x 720 pixels] to [1080 x 1920 pixels] and from 10 different cameras, specifically as follows: Totally, we took 50 original videos from VFDD2.0, then processing them as follows: separating these 50 original videos into two parts. The first part contains 50 videos including 25 original videos and 25 forged videos, therein the 25 forged videos were forged from these 25 original videos by duplicating small 3-D regions with different sizes of the 3-D region. This part dataset was used for detection evaluation on small 3-D regions duplication. Similarly, the second part contains 50 videos including 25 original videos and 25 forged videos, therein the 25 forged videos were forged from these 25 original videos by duplicating full-frame sequences. This second part dataset was used for detection evaluation on full-frame sequences duplication. Detail of this dataset in the supplementary data section of this paper.

b. The second dataset (REWIND_3D Dataset)

Because the REWIND dataset has not forged videos by duplicating frame sequences, so we extended this dataset to build REWIND_3D as follows: this dataset has two parts. The first part was taken from REWIND containing 20 videos (10 original videos, 10 forged videos) used for evaluating duplication of small 3-D regions. The second part was extended from 10 original videos by duplicating frame sequences. They were used for evaluating duplication of frame sequences (10 original videos, 10 forged videos).

Detail of this dataset in the supplementary data section of this paper, all of the dataset is published online at [22].

2. Experiments and Comparison

To evaluate the efficiency of the proposed algorithm, we have run the proposed algorithm with different sizes of the 3-D region. Fig. 6 illustrates the result of implementing the algorithm, Fig. 6(a) shows the result of implementing on an original video, and Fig.6(b) shows the result of implementing on a forged video. In addition, to evaluate the accuracy of forgery positions which are found out by the algorithm in videos, we have got the results of all pairs of 3D regions in the video then check accuracy whether the classification has regions duplication or not. There is a notation here when we divided a video into 3-D regions, and we have eliminated 3-D regions which contain all most 255 or 0 values. Because the pixels in these regions were not discriminated when captured by cameras.

In this section includes subsection 2.1, we present the result of the proposed algorithm on the VFDD_3D dataset and subsection 2.2 present the result of the proposed algorithm on REWIND_3D dataset.

2.1 Evaluating the Proposed Algorithm on the VFDD_3D Dataset

a. Detection of Small 3-D Regions Duplication

The first part of the VFDD_3D dataset contains videos with high-resolutions diversity, compression format MP4 and the forged videos were tampered by small 3-D regions forgery. So, it was used to evaluate the detection of small 3-D regions duplication. It was performed with different sizes of the 3-D region for assessing the effectiveness of the algorithm on different sizes of 3-D

regions. The results are shown in Fig. 7, these results showing that the too small or large of 3-D region sizes has given the bad result and both of this situation the detection accuracy of the algorithm is decreased. So, selection of 3-D regional size is neither large nor small to make increase accuracy of the algorithm. From these

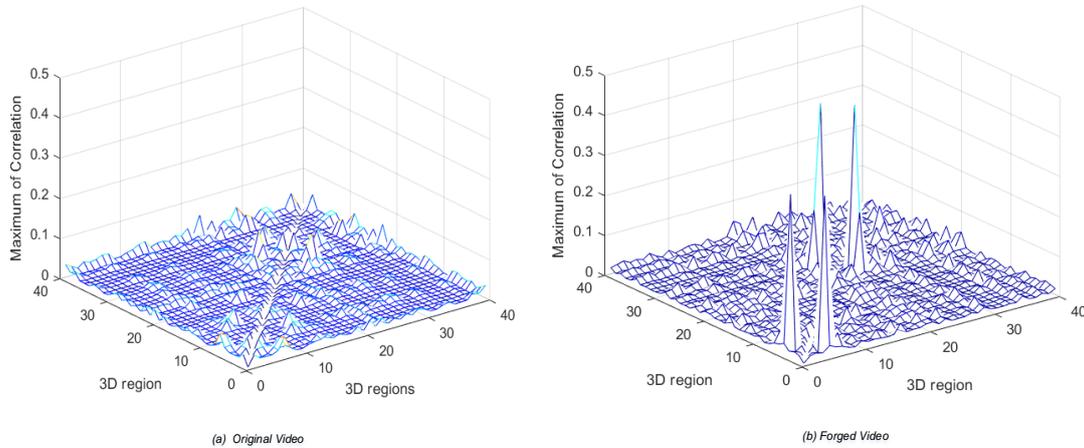


Fig. 6. Phase-Correlation of two 3-D regions in a video

experiments, we suggest selecting the 3-D regional size (Height x Width x Time) as 80 x 128 x 30 when detecting small 3-D regions duplication on this dataset. With 3-D regional size as 80 x 128 x 30, the algorithm has given the best result is the red ROC curve in Fig. 7.

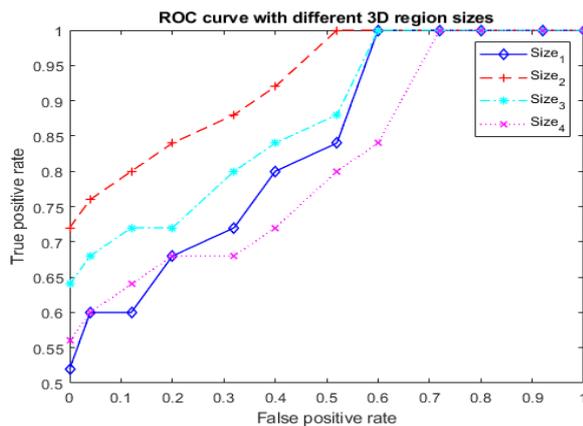


Fig. 7. ROC curves with different 3-D region size as follow: Size₁ is 36 x 64 x 30, Size₂ is 80 x 128 x 30, Size₃ is 120 x 180 x 30, and Size₄ is height of frame x width of frame x 30

In addition, to comparing the efficiency of the proposed algorithm for detecting small 3-D regions duplication with different methods, we have simulated other methods for detecting small 3-D region duplication from [6], [8] and [2] on this part dataset and the results are shown in Fig. 8.

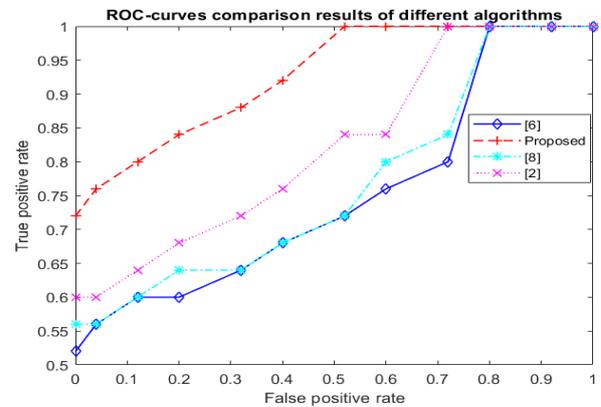


Fig. 8. Comparison [2] with, [6] and [8] schemes for 3-D region duplication on the first part of VFDD-3D dataset

b. Detection of Frame Sequences Duplication

The second part of the VFDD_3D dataset contains videos with high-resolutions diversity, compression format MP4, and the forged videos were tampered by frames sequence duplication. So, it was used to evaluate the effectiveness of detecting frame sequences duplication. It was also performed with different 3-D region sizes for assessing the effectiveness of the algorithm when detecting frames sequence duplication, and the results are shown in Fig. 9, which states that the larger in spatial of 3-D region sizes, the higher accurate detection. So, with 3-D regional size as frame size x 30, the algorithm has given the best result is the magenta ROC curve in Fig. 9.

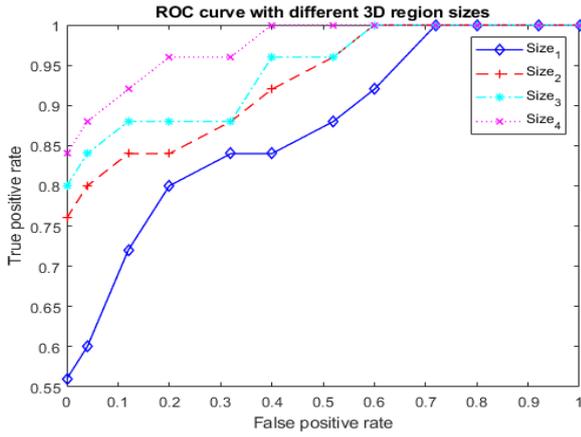


Fig. 9. ROC curves with different 3-D region size as follow: Size1 is 36 x 64 x 30, Size2 is 80 x 128 x 30, Size3 is 120 x 180 x 30, and Size4 is frame size x 30

To compare the proposed algorithm with other methods on detecting frames sequence duplication, we have simulated methods from [5], [6] and [8] for detecting frames sequence duplication and the results are shown in Fig. 10.

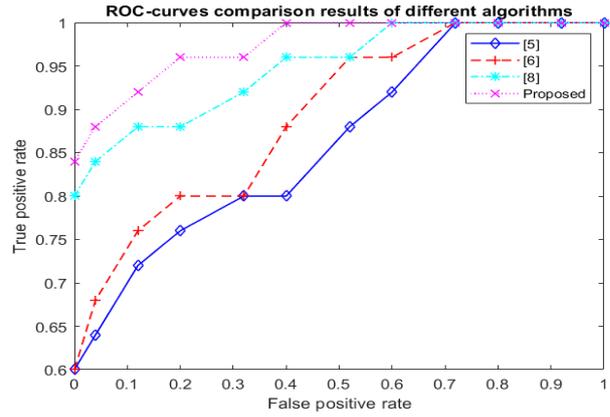


Fig. 10. Comparison with [5], [6] and [8] scheme for frames sequence duplication on the second part of VFDD_3D dataset

c. Localization

To evaluate the detection accuracy of forgery localizations which are found out by the algorithm in the VFDD_3D dataset. We depended on the identified result as duplication or not by the algorithm on all pairs of 3-D regions in the VFDD_3D dataset to evaluate the effectiveness of locating 3-D regions duplication. The following results were performed with 3-D regional size as 80 x 128 x 30 on the first part and width of frame x height of frame x 30 on the second part of VFDD_3D. The results of the algorithm are shown in Table 1.

Table 1. The result for locating of the algorithm on the VFDD_3D dataset

Video	The first part of VFDD_3D				The second part of VFDD_3D			
	Original Videos		Forged Videos		Original Videos		Forged Videos	
	Pair Number of True Negative	Pair Number of False Positive	Pair Number of True Negative and True Positive	Pair Number of False Positive and False Negative	Pair Number of True Negative	Pair Number of False Positive	Pair Number of True Negative and True Positive	Pair Number of False Positive and False Negative
01	1620	0	1620	0	7	0	7	0
02	1418	0	1418	0	10	0	10	0
03	170	0	170	0	8	0	8	0
04	256	0	256	0	6	0	6	0
05	170	0	170	0	8	0	8	0
06	284	0	284	0	13	0	13	0
07	170	0	170	0	7	0	7	0
08	170	0	170	0	8	0	8	0
09	1215	0	1209	6	10	0	10	0
10	227	0	227	0	5	1	6	0
11	1418	0	1418	0	7	0	7	0
12	366	3	369	0	7	0	6	1
13	369	0	369	0	11	0	11	0
14	341	0	341	0	7	0	7	0
15	312	1	312	0	10	0	10	0
16	1215	1	1215	0	12	0	12	0
17	2226	0	2224	2	9	0	9	0
18	625	0	625	0	8	0	8	0
19	170	0	170	0	9	0	9	0
20	256	0	256	0	7	0	7	0
21	256	0	256	0	18	0	17	0
22	270	0	270	0	14	0	14	0
23	426	0	426	0	9	0	9	0
24	284	0	284	0	6	0	5	1
25	282	2	284	0	15	0	15	0

The result for locating of 3-D regions duplication on Table 1 shows the detection accuracy as 99.95% on the first part and 99.35% on the second part of VFDD_3D dataset.

2.2 Evaluating the Proposed Algorithm on the REWIND_3D Dataset

a. Detection of Small 3-D Regions Duplication

Oppositely with VDFD_3D dataset, the first part of the REWIND_3D dataset contains videos with low-resolution, the forged videos have tampered with small 3-D regions duplication from [20]. So, it was used to evaluate the detection of small 3-D regions duplication. These experiments were performed with different sizes of 3-D region for assessing the efficiency of the algorithm on different sizes of 3-D region. The results are shown in Fig. 11. With 3-D regional size as 48 x 64 x 30 has the best result, which was a red ROC curve in Fig. 11.

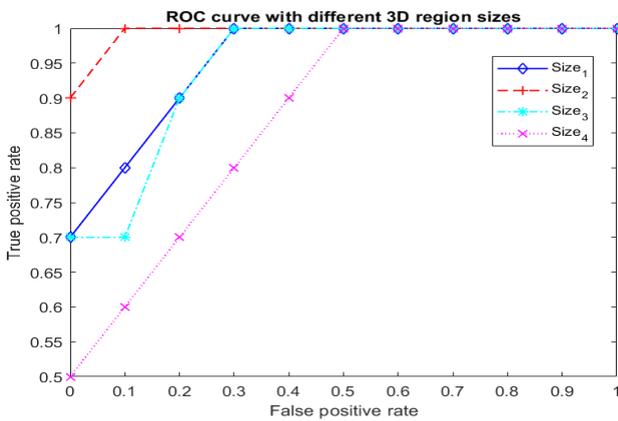


Fig. 11. ROC curves with different 3-D region size as follow: with Size1 is 24 x 32 x 30, Size2 is 48 x 64 x 30, Size3 is 80 x 100 x 30, and Size4 is frame size x 30

Similarly, we have also run the simulations from [6], [8] and [2] on this part dataset to detecting small 3-D regions duplication, and the results are shown in Fig. 12.

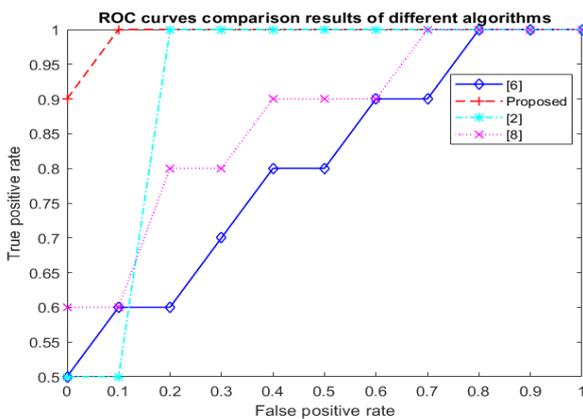


Fig. 12. Compare with [2], [6] and [8] scheme for 3-D regions duplication detection on the first part of REWIND_3D dataset

b. Detection of Frames Sequence Duplication

The second part of the REWIND_3D dataset also contains low-resolution videos and the forged videos have tampered with frames sequence duplication. So, it was used to evaluate the effectiveness of frames sequence duplication detection. It was performed with different sizes of 3-D region for assessing the effectiveness of the algorithm on different sizes of 3-D region and these experiments have given the highest detection accuracy with sizes of 3-D region as width of frame x height of frame x 30 shown in Fig. 13, in addition with the larger 3-D region size, algorithm detection robustness improved. This result was compared with previous researches on this dataset, as shown in Fig. 14.

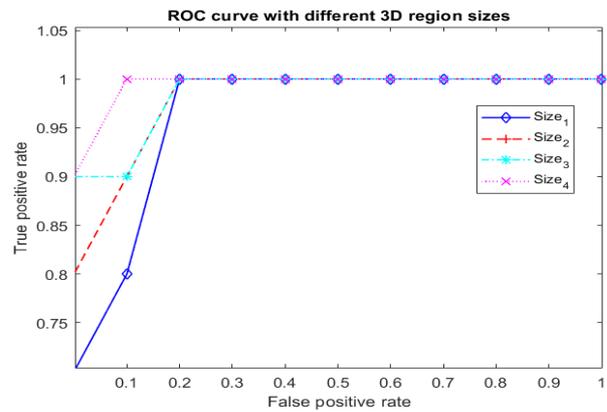


Fig. 13. ROC curves with different 3-D region size as follow: with Size 1 is 24 x 32 x 30, Size 2 is 48 x 64 x 30, Size 3 is 80 x 100 x 30, and Size 4 is width of frame x height of frame x 30

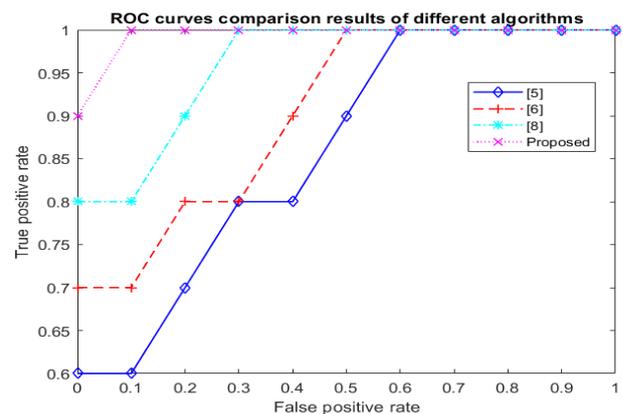


Fig.14. Comparison with [5], [6] and [8] schemes for frame sequences duplication on the second part of REWIND_3D dataset

c. Localization

Similarly, to evaluate forgery localizations which are found out by the algorithm in the REWIND_3D dataset. We depend upon the identified result as duplication or not by the algorithm on all pairs of 3D regions in REWIND_3D dataset to evaluate the effectiveness of locating 3-D regions duplication. The following result was performed with 3-D region size as 48 x 64 x 30 on the first part and frame size x 30 on the second part of REWIND_3D. The results of the algorithm are shown in Table 2.

Table 2. The result of localization on the REWIND_3D dataset.

Dataset	The first of REWIND_3D				The second part of REWIND_3D			
	Original Videos		Forged Videos		Original Videos		Forged Videos	
	<i>Localization of Detection</i>		<i>Localization of Detection</i>		<i>Localization of Detection</i>		<i>Localization of Detection</i>	
Video	Pair Number of True Negative	Pair Number of False Positive	Pair Number of True Negative and True Positive	Pair Number of False Positive and False Negative	Pair Number of True Negative	Pair Number of False Positive	Pair Number of True Negative and True Positive	Pair Number of False Positive and False Negative
01	175	0	155	20	8	0	7	1
02	250	0	250	0	10	0	10	0
03	325	0	325	0	13	0	13	0
04	250	0	250	0	10	0	10	0
05	475	0	475	0	19	0	19	0
06	200	0	200	0	8	0	8	0
07	325	0	325	0	13	0	13	0
08	225	0	225	0	9	0	9	0
09	450	0	450	0	18	0	18	0
10	175	0	175	0	7	0	7	0

The result for locating of 3-D regions duplication shows detection accuracy as 99.29% on the first part and 99.56% on the second part.

V. RESULT ANALYSIS

When detecting small 3-D region duplication in videos, selecting of 3-D region size in the algorithm for detecting is essential. If the size of 3-D region is small, it will increase the false-positive rate whereas the size is large will increase the false-negative rate. Therefore, selecting a suitable size of 3-D region increase detection accuracy. In the experiment on the first part of VFDD_3D dataset with high-resolution, the result has shown that 3-D region size with 80 x 120 x 30 pixels is the highest Area Under the ROC Curve (AUC) as 0.928 as shown in Fig. 7 and this result is better than previous researches as shown in Fig. 8 and Table 1.

Similarly, in the experiments on the first part of REWIND_3D with low-resolution, the results are given in Fig. 11, and Table 2 shows that selecting of 3-D region size with 48 x 64 x 30 has provided the highest UAC as 0.995. This result is better than previous researches, as shown in Fig. 12.

Detection of frames sequence duplication, the smaller the 3-D region size, the higher the false positive rate. So, selection of 3-D region size in spatial is as large as possible and the best result when selecting with frame size. In the experiments on the second part of VFDD_3D, REWIND_3D detecting frames sequence duplication with 3-D region size as frame size x 30 has the highest result. It gave the UAC results of 0.975 and 0.995 respectively. These results are shown in Fig. 9, Fig. 13, and Table 2. It is quite better than previously researches in Fig. 10, Fig. 14.

Localization of 3-D regions duplication in VFDD_3D and REWIND_3D datasets has been shown the results in Table 1 and Table 2 that is precisely on 99.9%. It is an impressive result for locating 3-D regions duplication positions of the proposed algorithm.

VI. CONCLUSIONS

In this research, we proposed the algorithm for 3-D regions duplication detection and forgery localization in videos. It is applied to detect two different kinds of forgery as duplication of small 3-D regions and duplication of frames sequence. In the experiments, the proposed algorithm checked on the VFDD_3D dataset with high-resolution and the REWIND_3D dataset with low-resolution. The results were validated on these datasets have proven that the proposed algorithm achieves high accuracy for detecting forged videos and perform better than previous algorithms. Especially, localization of forgery position in videos has a very impressive accuracy of over 99.9%.

We have also built two benchmark datasets for supporting algorithms evaluation when detecting video 3-D region duplication. The VFDD_3D dataset was completely built by us, it has features taken from modern cameras. This dataset has a variety of original and forged videos with 100 high-resolution videos. The REWIND_3D was expanded from REWIND dataset, videos in this dataset have features taken from three old cameras with 30 low-resolution videos.

Future works include researches to detect complex forgeries as inpainting, splicing, deleting forgery in videos. Besides, we will focus on researches that provide a method by which it can identify some types of forgeries in videos. Besides that, we should research the development of anti-forensic techniques that reduce the accuracy of detecting forgery videos, with the purpose to develop new methods for detecting videos forgery more strongly.

REFERENCES

- [1] Sitara, K. and B.M. Mehtre, Digital video tampering detection: An overview of passive techniques. *Digital Investigation*, 2016. 18: p. 8-22.
- [2] Bestagini, P., et al. Local tampering detection in video sequences. in *2013 IEEE 15th International Workshop on Multimedia Signal Processing (MMSP)*. 2013. IEEE.

- [3] Al Hamidi, S., VFDD (Video Forgery Detection Database) Version 1.0. <http://sites.scut.edu.cn/misip/main.psp>, 2017.
- [4] Qadir, G., S. Yahaya, and A.T. Ho, Surrey university library for forensic analysis (SULFA) of video content. 2012.
- [5] Wang, Q., et al., Video inter-frame forgery identification based on consistency of correlation coefficients of gray values. *Journal of Computer and Communications*, 2014. 2(04): p. 51.
- [6] Wang, W. and H. Farid. Exposing digital forgeries in video by detecting duplication. in *Proceedings of the 9th workshop on Multimedia & security*. 2007. ACM.
- [7] Yang, J., T. Huang, and L. Su, using similarity analysis to detect frame duplication forgery in videos. *Multimedia Tools and Applications*, 2016. 75(4): p. 1793-1811.
- [8] Singh, G. and K. Singh, Video frame and region duplication forgery detection based on correlation coefficient and coefficient of variation. *Multimedia Tools and Applications*, 2018: p. 1-36.
- [9] Chao, J., X. Jiang, and T. Sun. A novel video inter-frame forgery model detection scheme based on optical flow consistency. in *International Workshop on Digital Watermarking*. 2012. Springer.
- [10] Jia, S., et al., Coarse-to-fine copy-move forgery detection for video forensics. *IEEE Access*, 2018. 6: p. 25323-25335.
- [11] Liu, Y. and T. Huang, Exposing video inter-frame forgery by Zernike opponent chromaticity moments and coarseness analysis. *Multimedia Systems*, 2017. 23(2): p. 223-238.
- [12] Wang, W. and H. Farid. Exposing digital forgeries in video by detecting double quantization. in *Proceedings of the 11th ACM workshop on Multimedia and security*. 2009. ACM.
- [13] Subramanyam, A. and S. Emmanuel. Pixel estimation-based video forgery detection. in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. 2013. IEEE.
- [14] Ravi, H., et al. Compression noise-based video forgery detection. in *2014 IEEE International Conference on Image Processing (ICIP)*. 2014. IEEE.
- [15] Huang, Z., F. Huang, and J. Huang. Detection of double compression with the same bit rate in MPEG-2 videos. in *2014 IEEE China Summit & International Conference on Signal and Information Processing (ChinaSIP)*. 2014. IEEE.
- [16] Li, L., et al. Detecting removed object from video with stationary background. in *International Workshop on Digital Watermarking*. 2012. Springer.
- [17] Kobayashi, M., T. Okabe, and Y. Sato, detecting forgery from static-scene video based on inconsistency in noise level functions. *IEEE Transactions on Information Forensics and Security*, 2010. 5(4): p. 883-892.
- [18] Pandey, R.C., S.K. Singh, and K. Shukla. Passive copy-move forgery detection in videos. in *2014 International Conference on Computer and Communication Technology (ICCCCT)*. 2014. IEEE.
- [19] Subramanyam, A. and S. Emmanuel. Video forgery detection using HOG features and compression properties. in *2012 IEEE 14th International Workshop on Multimedia Signal Processing (MMSp)*. 2012. IEEE.
- [20] Bestagini, P., et al., REWIND Video: copy -move forgeries dataset. <https://sites.google.com/site/rewindpolimi/downloads/datasets/video-copy-move-forgeries-dataset>. 2012.
- [21] Al-Sanjary, O.I., A.A. Ahmed, and G. Sulong, Development of a video tampering dataset for forensic investigation. *Forensic science international*, 2016. 266: p. 565-572.

SUPPLEMENTARY DATA
The VFDD_3D dataset
The first part of the VFDD_3D

File Name	Width x Height	Length(s)	Camera	Frame Position of Forgery	Frames per second
01forgered	1920x1080	8	Canon EOS 1200D	region of 1->71,150->223	25
02forgered	1920x1080	7	Canon EOS 1200D	region of 145->175	25
03forgered	720x404	6	Canon EOS 70D(W)	region of 129->166	25
04forgered	720x404	9	Canon EOS 70D(W)	region of 176->225	25
05forgered	720x404	6	Canon EOS 70D(W)	region of 126->175	25
06forgered	720x404	10	Canon EOS 70D(W)	region of 100->200	25
07forgered	720x404	6	Canon IXUS 230 HS	region of 1->140	25
08forgered	720x404	6	Canon IXUS 230 HS	region of 100->140	25
09forgered	1920x1080	6	Canon EOS 1200D	region of 1->49	25
10forgered	720x404	8	Canon EOS 70D(W)	region of 51->101	25
11forgered	1920x1080	7	Canon EOS 1200D	region of 1->82	25
12forgered	720x404	13	Canon EOS 70D(W)	region of 52-> 305	25
13forgered	720x404	13	Canon EOS 70D(W)	region of 1->101	25
14forgered	720x404	12	Canon IXUS 95 1S	region of 289->360	30
15forgered	720x404	11	Canon EOS 70D(W)	region of 129->153	25
16forgered	1920x1080	6	Canon EOS 1200D	region of 66 ->172	25
17forgered	1920x1080	11	Canon EOS 1200D	region of 221->295	25
18forgered	720x404	22	Canon EOS 70D(W)	region of 142->550	25
19forgered	720x404	6	Canon EOS 70D(W)	region of 129->153	25
20forgered	720x404	9	Canon EOS 70D(W)	region of 169->225	25
21forgered	720x404	9	Canon EOS 70D(W)	region of 174->250	25
22forgered	640x480	9	Canon IXUS 95 1S	region of 135->223	30
23forgered	720x404	15	Canon EOS 70D(W)	region of 232->335	25
24forgered	720x404	10	Canon EOS 70D(W)	region of 179->250	25
25forgered	720x404	10	Canon EOS 70D(W)	region of 163->250	25

The second part of the VFDD_3D

File Name	Width x Height	Length(s)	Camera	Frame Position of Forgery	Frames per second
01forgered	720x404	7	Samsung Galaxy note 4	110-146 pasted to 147-183	25
02forgered	720x404	10	Canon IXUS 230 HS	203-254 pasted to 1-50	25
03forgered	1920x1080	8	Canon EOS 1200D	1-30 pasted to 31-61	30
04forgered	1920x1080	6	Canon EOS 1200D	45-95 pasted to 96-146	30
05forgered	720x404	8	Canon EOS 70D(W)	130-198 pasted to 199 -256	30
06forgered	720x404	13	Canon EOS 70D(W)	60-90 pasted to 91-121	30
07forgered	720x404	7	Canon EOS 70D(W)	1-55 pasted to 56-112	30
08forgered	720x404	8	Canon EOS 70D(W)	30-105 pasted to 106-281	30
09forgered	720x404	10	Canon EOS 70D(W)	55-118 pasted to 119-172	30
10forgered	1920x1080	6	Canon EOS 1200D	1-30 pasted to 31-61	30
11forgered	720x404	7	Canon IXUS 230 HS	100-140 pasted to 140-180	25
12forgered	720x404	7	Canon IXUS 230 HS	1-25 pasted to 157-182	25
13forgered	720x404	11	iPhone 6	125-156 pasted to 157-181	25
14forgered	720x404	7	iPhone 6	125-156 pasted to 157-181	25
15forgered	720x404	10	Canon EOS 70D(W)	150-175 pasted to 176 -191	30
16forgered	720x404	12	Canon IXUS 230 HS	200-250 pasted to 250-300	25
17forgered	720x404	9	Canon EOS 70D(W)	200-230 pasted to 231-261	30
18forgered	720x404	8	Canon EOS 70D(W)	20-60 - pasted to 101-161	30
19forgered	720x404	9	Canon EOS 70D(W)	50-100 pasted to 151-201	30
20forgered	720x404	7	Canon EOS 70D(W)	170-200 pasted to 201-231	30
21forgered	720x404	18	Canon IXUS 230 HS	1-50 pasted to 51-101	25
22forgered	720x404	14	Samsung Galaxy note 4	110-146 pasted to 147-183	25
23forgered	720x404	9	Canon EOS 70D(W)	140-170 pasted to 191-221	30
24forgered	720x404	6	Canon EOS 70D(W)	40-70 pasted to 101-131	30
25forgered	720x404	15	Canon EOS 70D(W)	1-65 pasted to 389-454	30

The REWIND dataset

The first part of the REWIND_3D

File Name	Width x Height	Length(s)	Camera	Frame Position of Forgery	Frames per second
01forgered	320x240	7	-	region of 1->210	30
02forgered	320x240	10	-	region of 110->170	30
03forgered	320x240	13	-	region of 200->280	30
04forgered	320x240	10	-	region of 20->150	30
05forgered	320x240	19	-	region of 10->160	30
06forgered	320x240	8	-	region of 190-260	30
07forgered	320x240	13	-	region of 65->140	30
08forgered	320x240	9	-	region of 110 ->170	30
09forgered	320x240	18	-	region of 200 ->310	30
10forgered	320x240	7	-	region of 100 ->150	30

The second part of the REWIND_3D

File Name	Width x Height	Length(s)	Camera	Frame Position of Forgery	Frames per second
01forgered	320x240	7	-	1->10 pasted to 31->40	30
02forgered	320x240	10	-	210->270 pasted to 150-210	30
03forgered	320x240	13	-	30->160 pasted to 180->310	30
04forgered	320x240	10	-	90->165 pasted to 165->240	30
05forgered	320x240	19	-	420->570 pasted to 1->150	30
06forgered	320x240	8	-	1->50 pasted to 205->265	30
07forgered	320x240	13	-	320->370 pasted to 1->50	30
08forgered	320x240	9	-	210->248 pasted to 240->278	30
09forgered	320x240	18	-	360->553 pasted to 167->359	30
10forgered	320x240	7	-	28->60 pasted to 61->93	30

Authors' Profiles



Xuan Hau Nguyen was born in Phu Yen, Viet Nam, in 1981, received the B.E. degree in Computer Science from the Ha Noi Technology University, Viet Nam, in 2005. Xuan Hau Nguyen graduated the Master degree in Information System from Ho Chi Minh City National University, University

of Natural Sciences, Viet Nam, in 2011 and is working toward Ph.D. degrees in electrical Information and Communication Engineering at the South China University of Technology, Guangzhou, P.R. China, from 2016.

In 2005, he worked at the Department of Computer Science, Mien Trung Industrial and Trade College, as a Lecturer, and in 2010 became the deputy of Computer Science Department at Mien Trung Industrial and Commercial College. He has published more than 6 peer reviewed papers. He currently

research interests include multimedia security, machine learning and database system.



Yongjian Hu was born in Wuhan, Hubei, graduated from Xi'an Jiaotong University in 1990 with a master's degree in information and control engineering. Yongjian Hu received the Ph.D. degree in communication and information systems from South China University of Technology in 2002.

Now he works as full Professor in School of Electronic and Information Engineering at South China University of Technology. From 2011 to 2013, he worked as Marie Curie Fellow in the Department of Computer Science, University of Warwick, UK. From 2006 to 2008, he worked as Research Professor in the Department of Computer Science, Korea Advanced Institute of Science and Technology (KAIST), South Korea. From 2005 to 2006, he worked as Research Professor in the School of Information and Communication Engineering, SungKyunKwan University, South Korea. Between 2000 and 2004, he visited the Department of Computer Science, City University of Hong Kong four times as a research assistant, senior research associate, and research fellow, respectively.

Dr. Hu is Senior Member of IEEE. He is also Senior Member of Chinese Institute of Electronics (CIE) and Senior Member of China Computer Federation (CCF). He has published more than 70 peer reviewed papers. His research interests include information hiding, multimedia security and machine learning.



Van Thinh Le was born in Phu Yen, Viet Nam, in 1976, received the B.E. degree in Computer Science from the Ha Noi Technology University, Viet Nam, in 2000, graduated the Master degree in Computer Science from Da Nang Technology University, Viet Nam, in 2011 and is

working toward Ph.D. degrees in the School of Computer Science and Engineering, Southeast University, PR China, from 2014.

In 2002 he worked at the Department of Computer Science, Mien Trung Industrial and Commercial College, as a Lecturer. He has published more than 12 peer reviewed papers. He currently research interests include multimedia security, machine learning and database system.



Gohar Hayat Khan was born in Mianwali, Pakistan, in 1991. received B.E. degree in Electrical and Electronics Engineering from University of Bradford United Kingdom, in 2014. Now is working Master degrees in electrical Information and Communication Engineering at the South China University of Technology, Guangzhou, P.R. China, from 2018.



Dinh Tu Truong was born in Phu Yen, Viet Nam, in 1979. Dinh Tu Truong received the Ph.D. degree in the School of Computer Science and Engineering from Southeast University, Nanjing, China in 2016.

He works at Ton Duc Thang University (TDTU), Ho Chi Minh City, Vietnam from 2016. His research area is network security, information hiding, network measurement, traffic sampling and machine learning.



Muhammad Ahmad Amin was born in Faisalabad, Pakistan, received his B.E. degree in Electrical and Electronics Engineering from The University of Faisalabad, Pakistan, in 2013. Muhamad Ahmad Amin received his Master's degree in Information and Communication Engineering from South China University of Technology, Guangzhou, P.R. China, in 2018 and is working towards his Ph.D. degree in Information and Communication Engineering at the South China University of Technology, Guangzhou, P.R. China, from 2018.

Currently, he working as Researcher in Research Centre of Multimedia Information Security Detection and Intelligent Processing, School of Electronics and Information Engineering at the South China University of Technology from 2016. From 2017 to 2019, he worked as Researcher and Algorithm Engineer in R&D institute at GRG Banking, Guangzhou, P.R. China. His research interests include machine learning, pattern recognition, artificial intelligence, multimedia security, and forensics.

How to cite this paper: Xuan Hau Nguyen, Yongjian Hu, Muhmmad Ahmad Amin, Khan Gohar Hayat, Van Thinh Le, Dinh Tu Truong, " Three-dimensional Region Forgery Detection and Localization in Videos", International Journal of Image, Graphics and Signal Processing(IJIGSP), Vol.11, No.12, pp. 1-13, 2019.DOI: 10.5815/ijigsp.2019.12.01