IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 5, 2018, accepted March 16, 2018, date of publication March 26, 2018, date of current version May 24, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2819624

Coarse-to-Fine Copy-Move Forgery Detection for Video Forensics

SHAN JIA[1,2], ZHENGQUAN XU[1,2], HAO WANG[1,2], CHUNHUI FENG[3], AND TAO WANG[1,2]

[1]State Key Laboratory of Information Engineering in Surveying Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China
[2]Collaborative Innovation Center of Geospatial Technology, Wuhan 430079, China
[3]College of Computer and Information Sciences, Fujian Agriculture and Forestry University, Fuzhou 350002, China

Corresponding author: Zhengquan Xu (xuzq@whu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 41301443, and in part by the Natural Science Foundation of Fujian Province of China under Grant 2016J01278.

ABSTRACT Video copy-move forgery detection is one of the hot topics in multimedia forensics to protect digital videos from malicious use. Several approaches have been presented through analyzing the side effect caused by copy–move operation. However, based on multiple similarity calculations or unstable image features, a few can well balance the detection efficiency, robustness, and applicability. In this paper, we propose a novel approach to detect frame copy–move forgeries in consideration of the three requirements. A coarse-to-fine detection strategy based on optical flow (OF) and stable parameters is designed. Specifically, coarse detection analyzes OF sum consistency to find suspected tampered points. Fine detection is then conducted for precise location of forgery, including duplicated frame pairs matching based on OF correlation and validation checks to further reduce the false detections. Experimental evaluation on three public video data sets shows that the proposed approach is effective and efficient in detecting both unsmooth manipulation and common smooth forgery and also with high robustness to regular attacks, including additive noise, filtering, and compression.

INDEX TERMS Copy-move forgery, optical flow, coarse–to-fine detection, video passive forensics.

I. INTRODUCTION

The high speed development and spread of image and video processing software, such as Photoshop, Adobe Premiere and Final Cut Pro, makes it easier to tamper with digital visual media without leaving obvious traces. However, malicious tampering may cause serious legal and social problems. For example, tampered images or videos may be used to provide false evidence in court, or mislead the public about the truth in news reports. Meanwhile, the vast and growing quantity of multimedia information makes it difficult to detect tampering using only human intuition. As a result, from 2001, automated methods for digital visual media forensics have become as routine as the application of physical forensic analysis [1].

Generally, digital forensic techniques can be classified into active approaches and passive or blind approaches. Two typical active forensic technologies are watermarking [2] and digital signatures [3], [4], both of which embed specific validation data in visual media during their production; but these methods require specialized hardware, limiting their application. The passive or blind forensic approaches verify the genuineness by exploring intrinsic features in the media left by acquisition devices or manipulation acts, without using any pre-embedded signals. It is a new research field emerging in the last decade, and has been a promising tool in the authentication field of digital visual media [5].

However, most of the passive forensic approaches were devoted to the analysis of still images [6]. In recent years, researchers have increasingly focused on video forensics, not only because the amount of video data is increasing at an explosive speed, but also because video tampering is becoming more and more easy, to which a wide range of possible alterations can be applied, such as frame deletion [5], [7], frame insertion [8]–[10], and video compression [11], [12]. Among them, copy-move forgery to extend or hide specific objects in the same video is one of the common methods. It is fairly easy to operate, but difficult to distinguish since the moved objects or frames are from the same videos. Based on different operational domains, video copy-move forgeries can be classified into regional forgery and frame cloning. Regional copy-move tampering changes only parts of the frame images, which is similar to image copy-move, and can be detected by relatively mature image

forensic techniques [1], [13]–[15]. The second type, frame copy-move, occurs in the time domain. It is performed by copying successive video frames and pasting them to another non-overlapping position, aiming to conceal objects, clone regions, or extend the time of some specific activities to forge the event records.

In frame copy-move forgery, cloning and pasting successive frames in the same video improves the imperceptibility and calculation difficulty, making it ineffective to detect the changes of color, shooting parameters, illumination condition, etc. [1], but it leads to abnormal points in the parameter distribution, and creates a high level of correlation between the original and duplicated frames. Based on this idea, different approaches have been developed to detect frame copy-move forgeries, which can be divided into two categories: image feature based and video feature based. Algorithms of the first category extract and explore image features of each frame to detect correlation, including gray values [16], [17], image texture [18], [19], color modes [20], and noise features [21], [22]. In the second category, they exploit the unique features in videos, such as motion features [23]–[25], video compression and coding features (including size, bitrate, and frame type) [26], [27] to analyze the side effect caused by copy-move operation. Although various detection solutions towards video copy-move forgery have been proposed, current schemes are faced with the following challenges:

- **High computation complexity.** Pixel based or directly correlation based approaches generally suffer from high computational burden. It will be quite time consuming to analyze a large number of frames in videos with a bulk of data far greater than that of still images.
- **Unstable detection performance.** Methods based on image features, including texture, color modes, noise, and pixel gray values, are vulnerable to regular attacks or post-processing on videos, like secondary compression and additive noise. Few of the existing detection approaches take the detection robustness into account, and generally set fixed sensitive parameters for detection.
- **Limited applicability.** Some methods have restrictions to the detected videos in terms of video formats, number of tampered frames, tampering ways (only for unsmooth manipulation) or shooting ways (only with static camera), which limit the practical applicability in video forensics.

These challenges imply that a practical frame copy-move forgery detection scheme is in high demand, which should satisfy three basic requirements: low computation complexity, high accuracy with good robustness, and strong applicability. In this paper, we try to take the above three requirements into consideration, and propose a new method to detect copy-move forgery. Our main contributions are as follows.

- To address the complexity of processing videos, we propose a coarse-to-fine approach based on the unique video motion features, Optical Flow (OF), to detect frame copy-move forgery. Coarse detection based on OF sum consistency aims to quickly find candidate tampered points, but it will lead to some false alarms. The detail features, OF, of candidate frames are then compared by OF correlation to match duplicated frame pairs, and false alarms can be further reduced by validation checks. The calculation efficiency is superior to most existing methods while guaranteeing the detection accuracy.
- The algorithm maintains high detection accuracy even under common attacks. Based on robust motion features, we also design adaptive or stable parameters to improve the robustness, and verify it by detecting tampered videos with regular attacks or secondary tampering, including additive noise, filtering, and compression.
- The proposed method achieves strong feasibility and applicability. It can deal with both unsmooth manipulation and common smooth forgery, and has no restriction to video formats and shooting ways (whether with static or moving cameras). Experimental results on three public video databases with different kinds of videos validate its strong applicability.

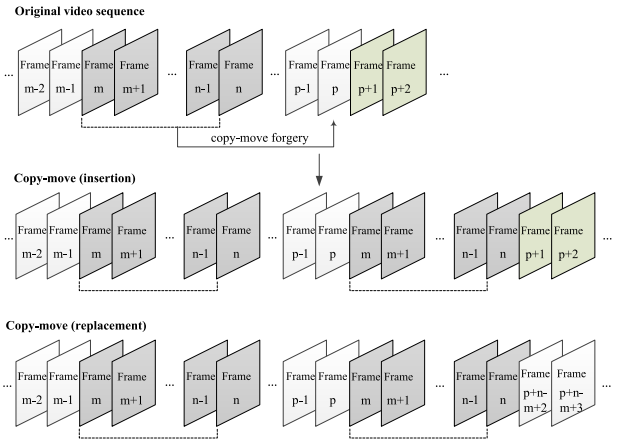


FIGURE 1. Illustration of frame copy-move forgery.

The rest of the paper is organized as follows. In Section II, we briefly introduce the related work for frame copy-move detection. Section III describes the preliminaries in this work. Our approach and experiments are presented in Sections IV and V, respectively. Finally, the discussions are provided in Section VI, followed by conclusions in Section VI.

II. RELATED WORK

Existing approaches for frame copy-move forgery detection have been presented through analyzing the side effect caused by copy-move operation, namely, the high feature correlation between the original and duplicated frames caused by either frame insertion or replacement, as shown in Fig. 1. Based on the extracted feature type, it can be divided into two categories: image feature based and video feature based.

A. IMAGE FEATURE BASED METHODS

This category of algorithms explores image features of each frame to detect frame correlation, such as pixel gray values, image texture, color modes, and noise features. Wang and Farid [16] earlier proposed a method based on temporal and spatial correlation matrices of pixels in gray images to detect duplication, finding that a high correlation indicates an instance of frame duplication forgery. In [17], the consistency of correlation coefficients of gray values after normalization and quantization was calculated to identify inter-frame forgeries. Liu and Huang [18] presented a dual positioning algorithm of video inter-frame forgery detection by analysing Zernike opponent chromaticity moments (ZOCMs) and coarseness (one attribute of Tamura texture features). Because the correlation calculation is based on low-order ZOCMs, it has high calculation efficiency. Liao and Huang [19] also used Tamura texture features for tamper detection. Three components: directionality, contrast and roughness were extracted and compared to detect video copy-move forgery. In [20], the authors designed a coarse-to-fine approach based on histogram difference of two adjacent frames in the RGB color space to detect video duplication forgery in the temporal domain. Inspired by the reliability of sensor pattern noise in identifying camera sources, Hsu and Hung *et al.* [21] explored the correlation of noise residue to detect video forgeries and achieved promising detection accuracy for fine-quality videos. Kobayashi *et al.* [22] also proposed an approach to detect suspected regions in videos by using the noise characteristics of the acquisition device.

However, these methods are usually vulnerable to regular attacks or post-processing on videos, like secondary compression and additive noise. For example, the performance of methods based on singular value decomposition (SVD) [10], gray value [17], and histogram difference [20] will suffer from noise or filtering, whereas noise-based approaches [21], [22] drops dramatically when the video is compressed by conventional codecs, such as MPEG-2 or H.264.

B. VIDEO FEATURE BASED METHODS

This kind of algorithms generally achieves higher robustness by exploiting the unique features in videos, such as motion features, video compression and coding features (including size, bitrate, and frame type). Chao *et al.* [9] and Chao [23] calculated Optical Flow (OF) consistency to detect video inter-frame forgery. They used a rough detection method and binary searching scheme to achieve good performance. Kingra *et al.* [24] analyzed gradients of prediction residual and OF for the detection of frame-based tampering in MPEG-2 and H.264 encoded videos. It can deal with frame insertion, removal or duplication, but the performance is not satisfactory, especially for videos with high illumination. Another forensic technique proposed in [25] used Motion Vector Pyramid (MVP) consistency to detect inter-frame forgery for static-background videos. In terms of video compression and coding features, Subramanyam and Emmanuel [26] made use of compression properties of MPEG-2 video codec to select the frames in a GOP, and combined histograms of oriented gradients (HOG) features to detect video forgery. Intrinsic effects of double compression on quantization errors of video coding were explored and traced in [27] to detect frame insertion or deletion and double compression with different GOP structures and lengths. However, these methods only analyzed the situation of MPEG-x videos specifically, and mainly focused on static-background videos or videos with no significant motion.

C. SUMMARY

To sum up, pixel based or directly correlation based approaches generally suffer from high computational burden, such as [16], [20], [22], [23], and [28] with multiple calculations or comparisons of correlation matrices, whereas methods based on video compression and coding features are limited in applicability. Besides, few of the two categories of approaches take the detection robustness into account, and generally set fixed sensitive parameters for detection. Therefore, in this paper, we try to make a tradeoff among the detection efficiency, robustness, and applicability, and propose a coarse-to-fine approach based on video OF features and stable parameters to address frame copy-move forgery in video forensics.

III. PRELIMINARIES

This section briefly describes the OF in video sequences, and explores the influence of frame copy-move forgery on OF correlation and OF sum consistency.

A. OF IN VIDEOS

OF is the distribution of apparent movement velocities of brightness patterns in videos [29], which can give important information about the image spatial arrangement and change rate of objects. Because of its highly descriptive motion information, it has been widely employed in multimedia processing and computer vision field including image segmentation, target tracking, face coding, mosaic construction, etc.

Differential methods are the most widely used techniques for OF computation in image sequences. Among them, the Lucas-Kanade Optical Flow, proposed by Lucas and Kanade [30], is a local least square calculation to compute OF sparsely for each blob. Because of the rapid computation, simple application, and robustness under noise [31], OF vectors extracted by Lucas-Kanade algorithm have been widely studied and used. Fig. 2 gives an example of a video sequence and shows the motion change vectors of the corresponding pixel between adjacent frames in Lucas-Kanade OF fields. The OF describes the details of movement changes in each frame and reflects the difference or similarity of frames in video sequences.

Since video copy-move forgeries across the temporal domain always aim to conceal the motion records or change

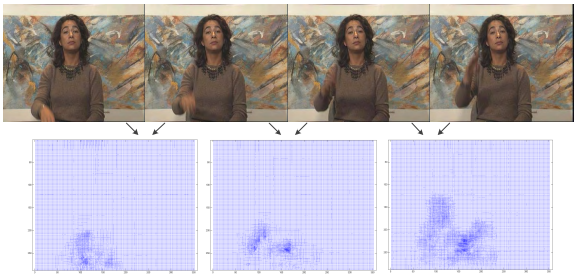


FIGURE 2. Illustration of motion changes in the OF field.

the time of some specific activities, videos recording moving objects are much easier to tamper. Therefore, in videos with copy-move forgery, OFs of adjacent frames can be extracted to record the detailed difference of frame images; high similarity of OFs between original and duplicated frames created by copy-move operation permits detection.

B. OF CORRELATION

To describe the OF similarity between frame images, the correlation coefficient is taken as a measure. For two adjacent frames i and $i+1$, the Lucas Kanade OF vector OF_i is decomposed into two figures: OX_i in X direction and OY_i in Y direction. In a video with N frames, N-1 OF vectors will be extracted and the correlation coefficients between every two OFs can be calculated with the following equation.

$$cor(i,j) = \frac{\sum_{m=1}^{wid}\sum_{n=1}^{hei}(OX_i(m,n)-\overline{OX_i})(OX_j(m,n)-\overline{OX_j})}{\sqrt{\sum_{m=1}^{wid}\sum_{n=1}^{hei}(OX_i(m,n)-\overline{OX_i})^2 \cdot \sum_{m=1}^{wid}\sum_{n=1}^{hei}(OX_j(m,n)-\overline{OX_j})^2}} \quad (1)$$

where X can be replaced with Y; $\overline{OX_i}$ and $\overline{OX_j}$ are the respective means of OX_i and OY_i. wid and hei are the numbers of pixels in each row and each column of the OF figures, which are the same with the video frames. $cor(i,j)$ is the element in the correlation coefficient matrix ranging from -1 to 1. A higher value indicates a higher similarity between OX_i and OY_i, therefore, meaning the adjacent frames i and $i+1$ have higher similarity with j and $j+1$, respectively.

To demonstrate how copy-move forgery affects the OF correlation, an example originated from a raw YUV sequence is given in Fig. 3 and Fig. 4. The correlation coefficient matrices of both *OX* and *OY* in video sequences are shown after removing the diagonal elements. In Fig. 3(a) and (b), the correlation coefficients between every two OFs are small in an original video because the OF records the motion change details of each corresponding pixel between two adjacent frames; different OFs have a relatively low correlation. But copy-move

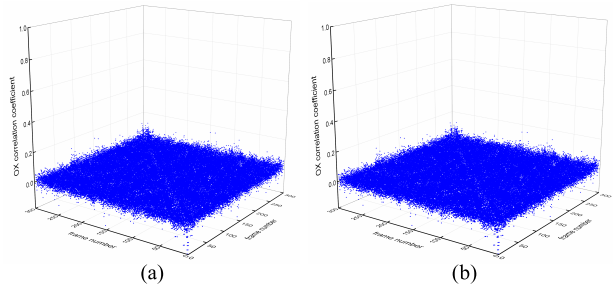


FIGURE 3. OF correlation coefficient matrices of an original video. (a). *OX* correlation coefficient matrix. (b). *OY* correlation coefficient matrix.

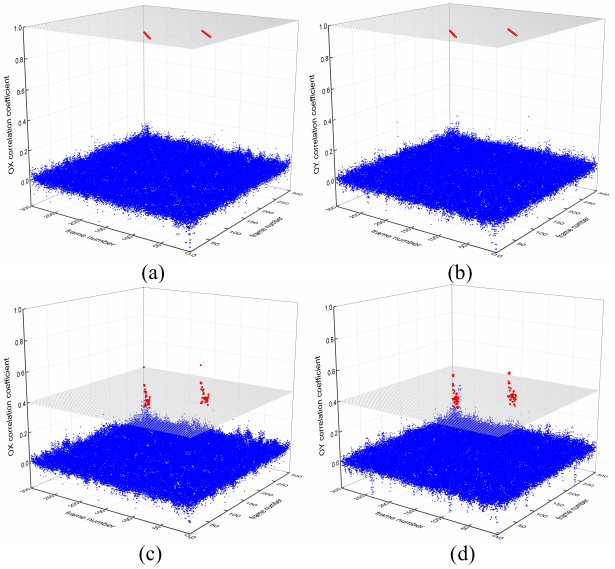


FIGURE 4. OF correlation coefficient matrices of a copy-move tampered video. (a). *OX* correlation coefficient matrix. (b). *OY* correlation coefficient matrix. (c). *OX* correlation coefficient matrix after video compression. (d). *OY* correlation coefficient matrix after video compression.

forgery will result in high correlation between original frame OFs and duplicated frame OFs in videos. From Fig. 4(a), (b), we can see that in a copy-move tampered video sequence, the correlation coefficients between the original frame OFs and the duplicated frame OFs are up to 1, significantly higher than the normal values. Moreover, post-processing on videos may be performed along with copy-move forgery, making the tampering difficult to detect, such as adding noise, filtering, and secondary loss compression. As shown in Fig. 4 (c), (d), although the video sequence is subjected to H.264 compression, high correlation between the original frame OFs and the duplicated frame OFs is still apparent, which can be identified by setting a smaller threshold.

Because of the robustness of OFs, even though additional operation introduces differences between initially identical frame sequences, the motion features will change little. Accordingly, the high OF correlation between original and duplicated frames still exists and can provide evidence for copy-move forgery detection.

C. OF SUM CONSISTENCY

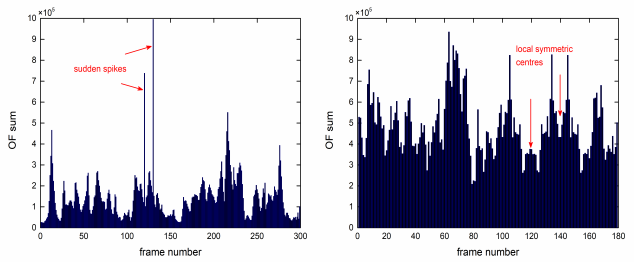
Since the calculation of OF correlation is point by point in frames, the computational cost is high and will increase rapidly as the image size and video length increase. Therefore, the consistency of OFs, as a global feature, is helpful to locate suspected tampered positions, and reduce multiple calculations or comparisons of correlation matrices in forgery detection.

We analyze the OF sum consistency to identify candidate tampered points. In a video with N frames, for the i^{th} frame, the absolute values of OX_i and OY_i in each pixel (m,n) are added with (2) as the OF sum, then the sum sequence composed of $N-1$ values is obtained.

$$sum\_OF_i = \sum_{m=1}^{wid}\sum_{n=1}^{hei}(|OX_i(m,n)| + |OY_i(m,n)|), \quad i = 1, 2, \ldots, N-1 \qquad (2)$$

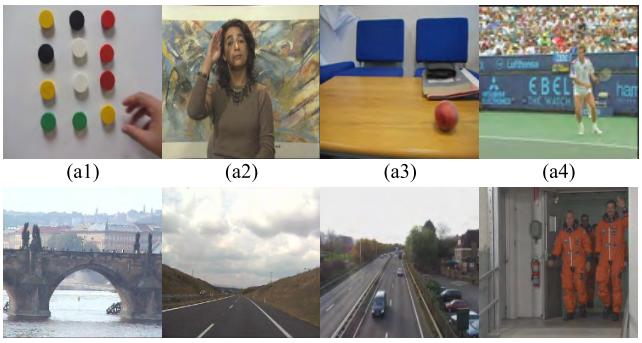
Based on how frame copy-move forgery affects the OF sum consistency, we classify it into two major types. The first type is to directly clone some frames to a different position in videos. The manipulation will generally result in sudden motion spikes in the OF sum sequence, such as inevitably unsmooth insertion because of motion in videos, or manipulation on non-key frames that may aim to change or extend the time of some key frames to obfuscate the event records. Because of the continuity and regularity of the motion in videos, the OF sum sequence will be relatively consistent, meaning no obvious spikes in the sequence. But this type of copy-move forgery will destroy the consistency due to frame replacement or insertion, and bring larger difference between adjacent frames, therefore, leading to anomalies in the OF sum sequence. Fig. 5 (a) shows an example of the OF sum sequence of a copy-move tampered video. There are some small fluctuations in the sequence caused by movement of objects, but these OF sums fluctuate slightly or gradually and have minor differences from the neighbouring OF sum values. However, spikes are manifest at the start and end points of the duplicated frames because a copy-move forgery destroys the consistency of the OF sums. These abnormal spikes can be detected to locate the tampered positions.

Another type is the careful manipulation which smoothly integrates the duplicated frames into videos to avoid the abnormal motion spikes. The easiest way is to insert frames in reverse order behind the tampered position, which is difficult to be detected by human eyes. For example, as shown in Fig. 5(b), frames 100 to 119 are inserted in reverse order behind frame 120, and then, frames 101 to 120 are inserted in proper order behind the inserted part. Therefore, the OF sum sequence is smooth at both the start point (the 120^{th} frame) and end point (the 160^{th} frame) of the duplicated frames, making it difficult to detect copy-move forgery based on abnormal spikes. However, there are obvious local symmetries because of the reverse insertion, where the symmetric centres (actually the start or end points of the duplicated frames) can be detected to locate the tampered



(a) (b)

FIGURE 5. OF sum sequence of tampered videos. (a). Tampered video sequence with spikes. (b). Tampered video sequence with symmetries.



(a1) (a2) (a3) (a4)

(a5) (a6) (b1) (b2)

FIGURE 6. Examples of video frames. (a1). Hand movement. (a2). Sitting person. (a3). Moving ball. (a4). Side-to-side motion. (a5). Remote monitoring. (a6). Moving camera. (b1). Moving car. (b2) Walking person.

positions. Note that inserting frames in reverse order is generally effective in videos with less or no direction attributes (see Fig. 6(a1-a6) as examples), but not feasible for videos with moving cars or persons (as shown in Fig. 6(b1-b2)), which can be easily detected by human observation.

D. SUMMARY

In summary, due to the highly detailed description of motion information, OFs of adjacent frames can reflect the difference or similarity of two pairs of frames. The high OF correlation between original and duplicated frames created by copy-move operation forms the basis for detection, but direct correlation calculation will lead to high computation cost. The OF sum consistency in video sequences can help to locate suspected tampered positions first and reduce multiple calculations of correlation matrices. Based on the OF and its features in video sequences, a coarse-to-fine detection scheme will be proposed to detect and locate frame copy-move forgery for video forensics.

IV. PROPOSED DETECTION SCHEME

In this section, we demonstrate the details of the proposed coarse-to-fine detection scheme for frame copy-move forgery. We first introduce the overall framework of the method. Then we describe the coarse detection based on OF sum consistency to locate suspicious tampered points. Finally, fine detection composed of duplicated frame pair

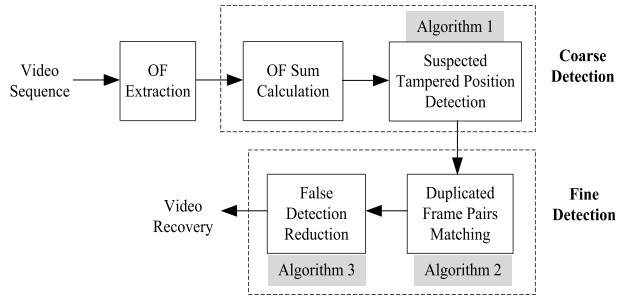


FIGURE 7. The detection process of the proposed scheme.

matching and false detections reduction is presented in detail.

A. PROPOSED DETECTION SCHEME

OFs and their high and stable correlation in copy-move tampered videos offer basis for effective detection. For calculation cost reasons, the consistency of OFs is analyzed first to locate suspected tampered positions. This process will help to reduce multiple calculations and comparisons of correlation matrices, but may lead to more false detections. Fine detection based on OF correlation is then proposed to match the duplicated frame pairs, and reduction of false detections based on validation checks will be conducted further for precision. If necessary, the duplicated sequence can be differentiated and deleted to realize video recovery. The whole detection process is shown in Fig. 7.

B. COARSE DETECTION

Temporal consistency is ubiquitous in original videos, where temporally adjacent video shots usually share similar visual and semantic content [32], leading to the similarity of features extracted from adjacent video shots. Therefore, we define the OF sum consistency as a high similarity in OF sums of adjacent video frames. Copy-move forgery will affect the consistency due to frame replacement or insertion, because larger difference of OFs at the start and end points of the duplicated frame sequence will lead to anomalies (i.e., sudden spikes and local symmetries) in OF sum sequences, providing a quantitative measurement for video analysis and forgery detection. Therefore, coarse detection based on the OF sum consistency helps to extract abnormal points as suspected tampered positions, avoiding multiple calculations in correlation analysis. The algorithm of abnormities detection is proposed as follows.

For a video sequence with N frames, we firstly extract all individual frames and compute the Lucas Kanade OF of every two adjacent frames i and $i+1$ ($i = 1, 2, \ldots, N-1$), obtaining OX_i matrices in X direction and OY_i in Y direction. Then we compute the OF sums with (2) and get the OF sum sequences composed of N-1 values. For the i^{th} frame, we next determine whether it is a suspected tampered position that leads to a sudden spike or local symmetry. The mean value of OF sums of its adjacent $2T$ frames is calculated with (3) to detect whether it is a sudden motion spike.

$$\overline{sum\_OF_i} = \frac{1}{2T}\sum_{k=1}^{T}(sum\_OF_{i-k} + sum\_OF_{i+k}) \quad (3)$$

where T is the window size for determining the number of adjacent frames. The rate of change β_i is defined to describe the fluctuation extent of the i^{th} frame and is measured by (4).

$$\beta_i = sum\_OF_i / \overline{sum\_OF_i} \quad (4)$$

If β_i is larger than a threshold THR_F, meaning an abnormal spike in sum_OF_i is manifest, then the i^{th} frame and its adjacent frames, $(i-1)^{th}$, $(i+1)^{th}$ frames, are identified as suspected tampered positions.

Meanwhile, we determine whether the i^{th} frame is a local symmetric centre in the OF sum sequence to detect the copy-move forgery of continuous reverse and forward insertion. If the OF sum around the i^{th} frame satisfies

$$sum\_OF_{i+k} \approx sum\_OF_{i-k-1}, \quad k = 0, 1, \ldots, T \quad (5)$$

It indicates that the frames before and after the local symmetric centre have the approximate equivalent OF sums, and maybe the duplicated frame pairs. Therefore, the i^{th} frame is identified as suspected tampered positions. The suspected tampered position detection process is summarized in Algorithm 1.

Algorithm 1 Suspected Tampered Position Detection

Input: sum_OF_i ($1 \le i \le N-1$), window size T, spike threshold THR_F
Output: Suspected tampered positions: sudden spike set S, local symmetric centre L

```
S = Ø, L = Ø
for i = 1; i < N; i++ do
    calculate sum_OF_i (overline) according to Eq. (3)
    calculate β_i according to Eq. (4)
    if β_i > THR_F then
        add i, i + 1, i − 1 into S
    end if
    if 7/10 < sum_OF_{i+k} / sum_OF_{i−k−1} < 10/7 (k = 0, 1, ..., T) then
        add i into L
    end if
end for
```

In copy-move forgeries, the suspected tampered positions may be the start or end points of the duplicated frame sequences. After the coarse detection, fine detection can find duplicated frame pairs only around the tampered positions, improving detection efficiency with little computation in the OF correlation computing.

C. FINE DETECTION

Coarse detection based on rough OF sum features helps to locate suspected tampered positions, but whether the anomalies are caused by copy-move forgery needs fine detection based on more detailed features to identify. In this section,

two steps of fine detection are proposed, including duplicated frame pairs matching based on OF correlation and reduction of false alarms based on video inherent features.

1) DUPLICATED FRAME PAIR MATCHING

OF correlation calculation is used to match the duplicated frame pairs after coarse detection. We extract OFs around each suspected tampered point, and calculate their correlation coefficients either with all the other OFs (for sudden spikes) or with the OFs of the adjacent frames (for local symmetric centres). Note that OX and OY have the same size with the video frame image, meaning the calculation of the OF correlation coefficients will be heavy. It is necessary to sub-sample the input OFs to reduce the number of pixels involved in the computation. Meanwhile, as shown in Fig. 3 and Fig. 4, the correlation coefficient matrices for OX and OY in the video sequences are nearly the same. Therefore, we only calculate the OF correlation coefficients of OX to reduce the computation load. The computing efficiency will be improved, but it will have little influence on the OF correlation coefficients distribution.

We first sub-sample the input OFs to reduce the number of pixels involved in the computation. A factor d^2 (scale every axes by d) is introduced to sub-sample the full-size OF matrix OX, obtaining OX'. Then, the process of duplicated frame pair matching is shown in Algorithm 2.

Algorithm 2 Duplicated Frame Pair Matching

Input: OF sequence $OX_i^{'}(1 \leq i \leq N)$, sudden spike set S, local symmetric centre L, threshold THR_C_1, THR_C_2
Output: Duplicated frame pair set D

```
D = Ø
for each frame number i ∈ S do
    for j = 1; j < N; j++ do
        calculate cor(i, j) according to Eq. (1)
    end for
    obtain the maximum correlation coefficient
    cor(i, j)_max, i ≠ j
    if cor(i, j)_max ≥ THR_C1 then
        add (i, j), (i + 1, j+1) into D
    end if
end for
for each frame number i ∈ L do
    k = 0
    while cor(i + k, i − k − 1) ≥ THR_C2 &
    cor(i + k + 1, i − k − 2) ≥ THR_C2
    do
        k = k + 2
    end while
    add (i − k −1, i + k + 1) into D
end for
```

Algorithm 2 runs as follows. First, for each suspected tampered frame number i that is detected as a sudden spike, it calculates the OF correlation coefficients $cor(i, j)$ $(j = 1, 2, \ldots, N-1)$ to find the maximum correlation coefficient $cor(i, j)_{max}$. The threshold THR_C_1, which is always significantly larger than the average value of all the OF correlation coefficients $cor(i, j)$, is used to determine whether the related frame pairs (i, j), $(i+1, j+1)$ of $cor(i, j)_{max}$ have high correlation coefficients. Then, for each suspected tampered frame number i that is detected as a local symmetric centre, it calculates the OF correlation coefficients of the frame pairs before and after i frame. Note that the while loop will be repeated for at most n_t times, where n_t is the number of copy-moved frames. The threshold THR_C_2 is used to get the successive frames with high correlation coefficients. The final outputs (with either two points from abnormal spikes or three points from symmetric centres) of the algorithm are the candidate start or end points of tampered frame sequences.

2) REDUCTION OF FALSE DETECTIONS

It is worth noting that fine detection for copy-move forgeries depends on the coarse detection results with abnormal points in OF sum sequences. However, tampering is not the only factor accounting for the outliers in coarse detection phrase. Other factors may also produce spikes or local symmetric centres, leading to false detections. For example, some spikes may come from the weaker OF sum consistency in videos with quickly moving content, while local symmetry may be derived from continuous static scenes or smooth movement in videos. In fine detection based on correlation analysis, adjacent frames with high similarity will also lead to false alarms. Besides, additional operations may be performed after copy-move forgery to cause interference and cover up the abnormities.

Therefore, validation checks based on the inherent features of videos will be introduced to reduce the interference frames as further fine detection. We define three inherent features of videos in copy-move forgery detection as follows.

Similarity. Adjacent frames or frames in a short time interval have high similarity because of video consistency.

Similarity leads to high correlation between original adjacent frames and may cause false detections in matching duplicated frame pairs. However, it can be distinguished by detecting the frame number differences between the suspected duplicated frame pairs. A small difference means the two frames are close to each other, and the high correlation is caused by video similarity instead of copy-move operation.

Continuity. Videos with continuous multi-frames carry more information, and will be more likely to be tampered than discontinuous or short-length frames with scarce actual meaning.

Continuity ensures that the tampered frames are a successive sequence. That is, for a suspected frame pair (i, j), if both $(i-1, j-1)$ and $(i+1, j+1)$ frame pairs have low correlation, (i, j) should be removed as a false detection; if both $(i-1, j-1)$ and $(i+1, j+1)$ frame pairs have high correlation, (i, j) should also be removed because it is not the end or start point of the tampered frame sequence.

*Regularity.*The detected duplicated video sequence has the same length with its original sequence, meaning both the intervals of the start points and the end points are equal.

Regularity means that the two detected sequences with high correlation (i.e., the duplicated frame sequence and its original sequence) should have the same length, and ensures the integrity of the detection results.

Making use of these features in videos for fine detection after OF correlation calculation, false alarms will be effectively reduced. The main process is present in Algorithm 3. After that, the tampered video can be recovered by removing the duplicated frames.

Algorithm 3 False Detections Reduction

Input: Candidate duplicated frame pair set D, threshold THR_C_2, minimum number of tampered frames $W = 10$
Output: Duplicated and original frame sequences

```
for each frame pair (i, j) ∈ D do
  if (|j − i| < W)||(cor(i − 1, j − 1), cor(i+1, j + 1) <
     THR_C2)||(cor(i − 1, j − 1), cor(i + 1, j + 1) >
     THR_C2) do
    delete (i, j) from D
  end if
end for
choose (i_p, j_p), (i_q, j_q) ∈ D, i_p < i_q, and |i_p − j_p =
   |i_q − j_q|, output {i_p, i_p+1, ..., i_q, i_q+1} as duplicated
   frames, {j_p, j_p+1, ..., j_q, j_q+1} as original frames
for each frame pair (k, i, j) ∈ D do
  if |k − j| < 2W do
    delete (k, i, j) from D
  else output {i + 1, ..., j − 1, j} as duplicated
       frames, {k, k + 1, ..., i − 1} as original frames
  end if
end for
```

V. EVALUATION

We conduct a series of experiments to evaluate the performance of the proposed detection scheme in this section. The experimental data and evaluation standards are introduced first. Then the involved parameters are determined with a subset of tampered video sequences. Finally, we present the experimental results and comparison analysis with four existing classical algorithms in terms of detection accuracy, robustness, efficiency, and applicability.

A. EXPERIMENTAL DATA AND EVALUATION STANDARDS

As there are no large-scale video datasets available for copy-move forgery detection [33], most of which are for regional copy-move forgery, we simulated the frame copy-move forgery by randomly selecting a sequence of frames in the experiment for general detection as [16]–[20], [23]–[28], and [34] did. The original test data composed of 115 videos come from three public video databases: 1) 55 standard YUV sequences downloaded from the video trace library (http://media.xiph.org/video/derf/), denoted as VTL; 2) 36 videos in the AVI or MOV format downloaded from SULFA (Surrey University Library for Forensic Analysis, http://sulfa.cs.surrey.ac.uk/videos.php), denoted as SULFA; 3) 24 videos in the y4m format from Derf's Test Video Collection (http://media.xiph.org/video/derf/), denoted as DERF. Among them, 70 sequences were taken from stationary cameras and 45 were taken from moving cameras, both including videos with slow or fast movements.

TABLE 1. Details of the video datasets.

Database	Format	Resolution	Number	Length	Sequence Number		
					Original	Type I	Type II
VTL	YUV	352×288	50	90-300	5	25	25
SULFA	AVI/MOV	320×240	30	199-468	6	15	15
DERF	Y4M	176×144	20	180-288	4	10	10

Video copy-move tampering was realized in two ways. We simulated the first type, Type I, in 50 sequences by selecting a random location in each original video sequence, and duplicating a number of successive frames to another non-overlapping position. For the second type of copy-move forgery, Type II, we selected a random location and made continuous reverse and forward insertion of the frame sequence after the position in each of another 50 videos. The remaining 15 video sequences are original without copy-move forgery. Table 1 shows the details of the video datasets. Then, the tampering process was repeated three times, each time selecting different number of frames, namely, 10, 20, and 40 respectively, to be duplicated to a new location. Finally, the video dataset with 3 groups, totally 315 tampered videos, was formed; it will be available from http://202.114.114.212/whu/yuv_download.html.

To analyze the performance of the proposed scheme, the recall rate R and precision rate P were used to evaluate the results, as defined in (6), (7).

$$R = \frac{N_c}{N_c + N_m} \tag{6}$$

$$P = \frac{N_c}{N_c + N_f} \tag{7}$$

where N_c is the number of correct detection, N_m represents the number of missed detections, and N_f denotes the number of false positives. A high R means a low missing detection rate, and a high P indicates a low false detection rate.

The computing environment was a windows 7 system with an Intel i3 processor, and the programming language was C/C++ in Microsoft Visual Studio 2005.

B. PARAMETER DETERMINATION

After generating the tampered video set, we used a subset (as shown in Table 2) to determine optimum parameters to get the proper R and P, including two variable factors (the window size T and sub-sampling factor d), and three thresholds

TABLE 2. Sequence numbers of the subset for parameter determination.

Duplicated frame number	Type I			Type II		
	VTL	SULFA	DERF	VTL	SULFA	DERF
10	10	6	4	10	6	4
20	10	6	4	10	6	4
40	10	6	4	10	6	4

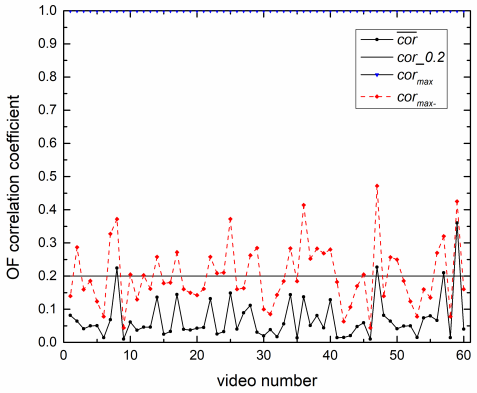


FIGURE 8. Examples of videos to determine *THR_C*.

(the fluctuation threshold THR_F, the correlation thresholds THR_C_1 and THR_C_2).

We first describe the process to select the threshold THR_C_1, which is used to determine whether the maximum correlation coefficient $cor_{\max}$ of each suspected tampered spike (see Algorithm 2) is large enough to be detected as duplicated frame pairs. Considering that the value of $cor_{\max}$ is related to the video content, we studied the relationship between $cor_{\max}$ and $\overline{cor}$ of each Type I tampered video to determine THR_C_1. We set both the window size T and sub-sampling factor d to 2, the threshold THR_F to 1.5, and THR_C_2 to 0.2. The result on 60 Type I tampered video sequences is shown in Fig. 8. It can be seen that the $\overline{cor}$ values are generally smaller than 0.2. Taking into account the $cor_{\max\_}$ values, which represent the maximum values removing the tampered frame value, we set THR_C_1 as an adaptive threshold, as follows.

$$THR\_C_1 = \begin{cases} 0.3, & \overline{cor} < 0.2 \\ 2 \times \overline{cor}, & \overline{cor} \geq 0.2 \end{cases} \tag{8}$$

A smaller $\overline{cor}$ means the suspected frame has low OF correlation with other frames (because the slow or local motion of video content leads to small values of OF, especially in most surveillance videos, see Fig. 9(a2) and (b2) as an example). Then a smaller threshold will be set to locate its duplicated frame with the maximum correlation coefficient. Inversely, from Fig. 9(c2), we can see that the video captured by a moving camera trends to have a higher $\overline{cor}$, and therefore needs a larger threshold to find duplicated frame pairs. The robustness of the relationship between $cor_{\max}$ and $\overline{cor}$ was also tested on videos with additive Gaussian white noise, filtering, or secondary H.264 compression as secondary forgery.

TABLE 3. Detection time (s/frame) of different window size *T*.

Database	$T=1$	$T=2$	$T=4$
VTL	0.124	0.164	0.196
SULFA	0.110	0.122	0.145
DERF	0.032	0.042	0.051

TABLE 4. Detection time (s/frame) of different sub-sampling factor *d*.

Database	$d=1$	$d=2$	$d=4$
VTL	0.184	0.164	0.128
SULFA	0.143	0.122	0.102
DERF	0.047	0.042	0.035

As shown in Fig. 9(a3-a5), (b3-b5), (c3-c5), the results still satisfy (8) because of the robustness of OF features.

Then, the window size T was determined. The detection process and results under different T values are illustrated in Fig. 10(a1), (a2) and Table 3. We can see that a larger window size T gets a relatively higher P but a lower R and longer computation time, while a smaller window size takes shorter computing time but is more likely to result in false detection. Because a smaller T helps to locate more abnormal values as spikes or local symmetric centres, it therefore results in a higher recall rate R but a lower precision rate P. For a rational R, P and computation time, 2 was selected as the value of window size T.

To determine THR_F, we varied it from 1.1 to 3.1 with a 0.4 step and got the detection results as shown in Fig. 10(b1), (b2). With the increase of THR_F, the average precision rate P of three datasets rises, but the recall rate R declines because some sudden spikes caused by copy-move forgery are not obvious enough to be detected. For both higher R and P, THR_F was set at 1.5.

The detection result for selecting the sub-sampling factor d in OF correlation calculation was illustrated in Fig. 10(c1), (c2) and Table 4. A larger d contributes to a shorter computing time but sub-sampling the OF will reduce the difference and lead to more false detection. Therefore, for a relatively high R, P and short computation time, d was selected as 2.

The last step was to determine THR_C_2 for finding duplicated frame pairs around suspicious symmetric centres. As shown in Fig. 10(d1), (d2), when THR_C_2 increases, both the average precision rate P and recall rate R can reach 1 for the Type II copy-move forgery. For parameter stability reasons, THR_C_2 was set to be a smaller value 0.2.

To summarize, the experiment results show the effects of different parameters on the detection performance. These parameters are then tested to evaluate the robustness and efficiency of the proposed method in the remaining video subset.

C. ACCURACY AND ROBUSTNESS

In order to evaluate the accuracy and robustness of the algorithm, some common attacks were simulated as secondary forgery after copy-move forgery, including additive Gaussian

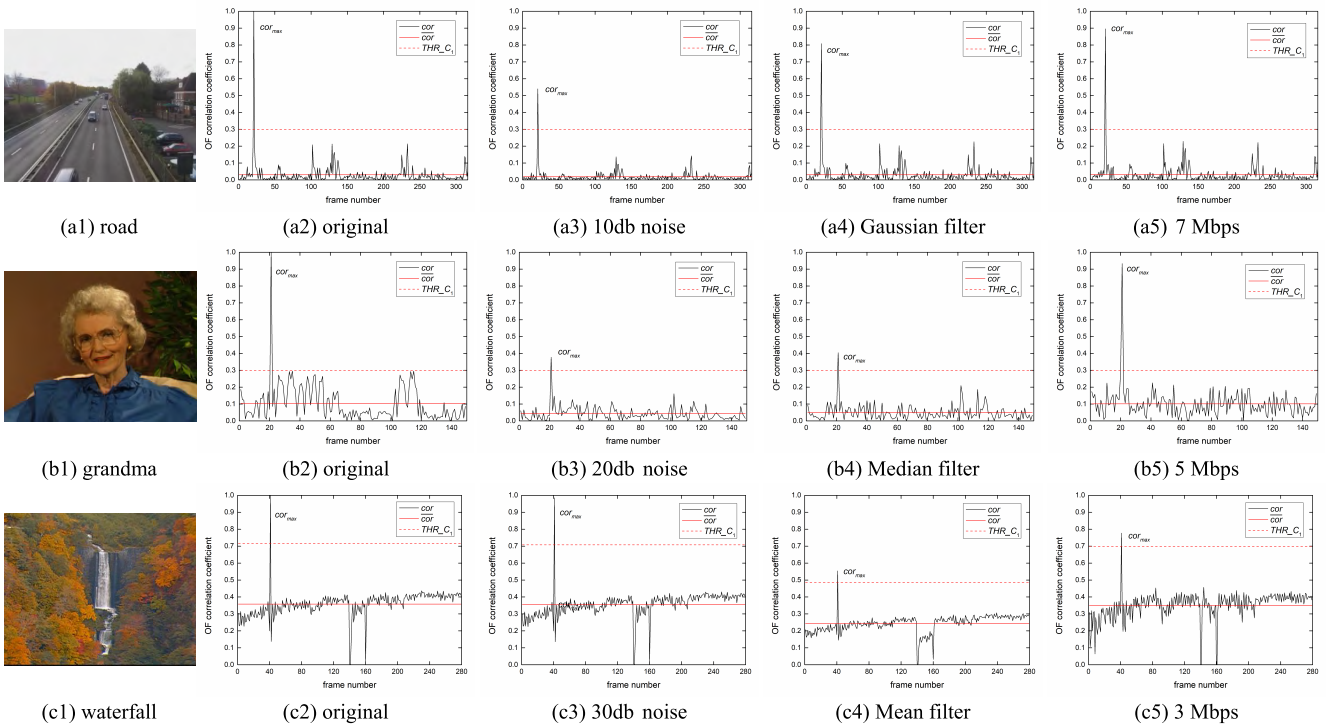


FIGURE 9. Examples of videos to determine THR_C_1.

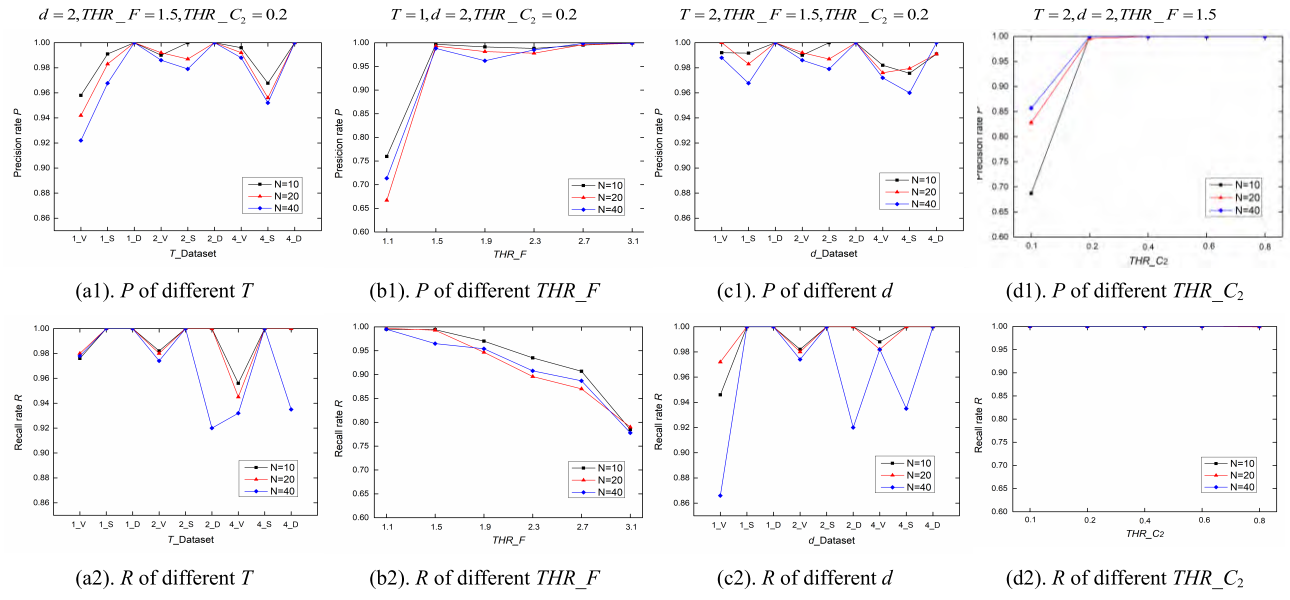


FIGURE 10. Detection results of different parameters.

white noise, filtering, and secondary H.264 compression. The detection performance of the proposed approach has been tested on the remaining tampered videos with different types and different tampered numbers. Table 5 shows the average results of the algorithm on three databases with no attack, with additive Gaussian white noise using different PSNRs, filtering with different 3×3 filters, and secondary H.264 compression using different bit rates, respectively.

The experimental results show that under different intensity of Gaussian noise or filtering, the precision rate P remains above 0.950 owing to the fine detection, while the recall rate R was above 0.920 except the cases that 10 dB or 20 dB noise was added, which reduced R to 0.904 and 0.917

TABLE 5. Detection results of the proposed method.

Attacks		Precision rate P	Recall rate R
Under the condition of no attack		0.985	0.985
Gaussian noise	10 dB	0.953	0.904
	20 dB	0.967	0.917
	30 dB	0.982	0.936
3×3 Filter	Mean Filter	0.963	0.935
	Median Filter	0.952	0.930
	Gaussian Filter	0.960	0.932
H.264 Compression	0.5 Mbps	0.931	0.747
	1 Mbps	0.912	0.921
	3 Mbps	0.958	0.920
	5 Mbps	0.967	0.950
	7 Mbps	0.978	0.980

TABLE 6. Comparison results in efficiency and applicability.

Criteria	Chao *et al.*[9]	Wang and Farid [16]	Yang *et al.*[10]	Liu *et al.* [18]	Proposed
Calculation Time(μs /pixel)	3.196	3. 275	2.580	1.095	1.623
Location of duplication	Yes	No	Yes	Yes	Yes
Common smooth forgery	No	Yes	No	No	Yes
Accuracy	Medium	High	Medium	Low	High
Robustness	Strong	Medium	Weak	Medium	Strong

respectively. This is because more noise in poor-quality images will introduce more differences between initially identical frame sequences, leading to a higher omission ratio. Similarly, when the bit rate of H.264 compression declines, especially to 1Mbps and 0.5 Mbps, the poorer quality of videos has a great influence on the OF correlation, and results in a lower R. But in most cases, the proposed algorithm achieved both high precision rate and recall rate for frame duplication detection.

Next, we compare the performance of the proposed method with four existing algorithms in terms of detection accuracy and robustness. We re-implemented the competing algorithms according to the algorithm description in their papers and used the same way to simulate frame copy-move forgery. To show the influence of datasets on detection performance, we tested on different datasets, i.e., VTL, SULFR, and DERF, to compare the detection results, as shown in Fig. 11.

By vertical comparison from the trend of each curve in Fig. 11, we can see that the detection performance of OF-based methods (the proposed and [9]) is stable because of the robustness of OF. The Zernike moment based method in [18] converted each frame from the three-dimensional RGB space into two-dimensional opponent chromaticity space, eliminated some noise of images, and also achieved relatively good robustness. The robustness of the methods in [10] and [16] is mainly influenced by filtering process because filtering on poor-quality video frames will have a larger influence on the correlation of pixel-based methods, therefore leading to more false or missing detection.

The horizontal comparison indicates that the proposed method achieved a higher R and P owing to the characteristics of OF and the coarse-to-fine strategy. Note that although the fixed and sensitive threshold parameters (to detect outliers or high correlation) in [9], [10], and [18] have been adjusted to detect robustness in the experiments, these methods cannot deal with the type of copy-move forgery with smooth insertion, which was not included in the detection results. The method in [18] performs poorly in detection due to its limitation in dealing with videos with fast movements of the tampered area, while [9] was affected by its rough calculation on spikes detection. The method proposed by Yang *et al.* [10] performed better based on SVD (singular value decomposition) feature and double-checking, and [16] benefited from its similarity analysis of both temporal and spatial correlation matrices. Moreover, through comparison of Fig. 11(a1-a2) and (b1-b2), we can see that the performance on dataset SULFR is poorer. The reason is that it includes more surveillance videos with large range of static scene, leading to less obvious characteristics of tampering.

D. EFFICIENCY AND APPLICABILITY

Table 6 summarizes the performance of the four comparison methods and the proposed algorithm in terms of detection efficiency and applicability. The method [18] has the shortest computation time (1.095 μs/pixel) due to its coarse-to-fine scheme and lower-dimensional ZOCMs features for similarity calculation. The proposed method also benefited from the coarse-to-fine strategy and required about 1.623 microseconds for each pixel, much less than that of [9], [10] and [16]. At the same time, it achieved better applicability, higher detection accuracy and stronger robustness than other approaches.

VI. DISCUSSIONS

Based on the experimental results, we discuss the parameter robustness and limitation of the proposed method in this section.

We attribute the robustness of the method not only to the robustness of OF features, but also to the stability of parameters. Two variable factors, namely, the window size T and sub-sampling factor d, mainly for improving calculation efficiency, are related to the resolution or length of video sequences. Both fixed at 2, they could generally apply to most video sequences. The thresholds, THR_F, THR_C_1 and THR_C_2, making more contribution to the robustness, are also stable for different videos. The main reasons are as follows. 1) The fluctuation threshold THR_F is a ratio, used to determine whether the fluctuation extent can be detected as spikes caused by tampering. In this paper, we describe the fluctuation extent by (4): $\beta_i = sum\_OF_i / \overline{sum\_OF_i}$. It can be seen that for videos with faster motion, both the sum_OF_i and $\overline{sum\_OF_i}$ will be larger, while for videos with slower motion,

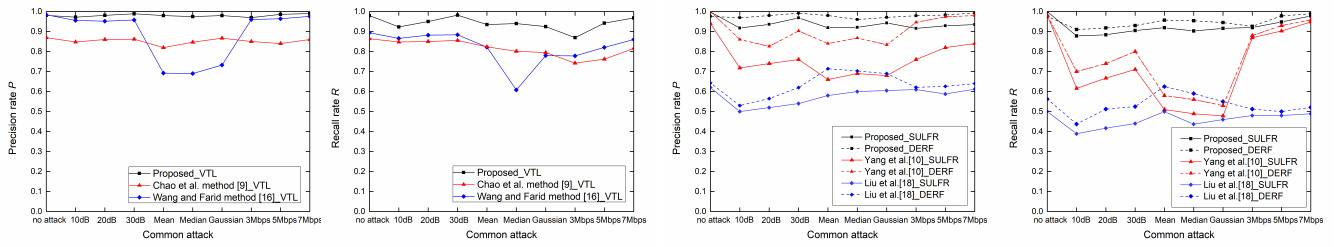


FIGURE 11. Comparison results in accuracy and robustness. (a1). *P* on dataset VTL. (a2). *R* on dataset VTL. (b1). *P* on dataset SULFR and DERF.

they both will be smaller. Therefore, the ratio is relatively stable. 2) For THR_C_1, we set it as an adaptive threshold, which is dependent on the video content to determine whether the maximum correlation coefficient of each suspected tampered frame is large enough to be detected as duplicated frame pairs. 3) THR_C_2 is used to find duplicated frame pairs around suspicious symmetric centres. It is an empirical value. As the correlation coefficient of two frames in the original video tends to be very small due to OF's highly detailed description of motion; therefore, it can be set as a smaller value for parameter stability reasons (see Fig. 10(d1), (d2)).

However, one limitation of the proposed method is that the recall rate of detection is sensitive to the coarse detection. It may miss some copy-move forgeries which do not have an influence on OF sum consistency, such as tampered videos with a largely static scene and other types of carefully prepared manipulation. Reducing the threshold THR_F helps to detect these cases (according to Fig. 10(b2)), but the advantage of high calculation efficiency of coarse detection will be reduced.

According to the idea of taking different detection algorithms as a forensics tool set (FTS) to provide reliable and sufficient evidence [35], the proposed method will serve as a promising tool with high efficiency and robustness for comprehensive detection of frame copy-move forgery in combination with other forensic tools.

VII. CONCLUSION

A practical frame copy-move forgery detection scheme should achieve low computation complexity, high accuracy with good robustness, and strong applicability. In this paper, we present a coarse-to-fine approach based on video OF features and stable parameters to make a tradeoff among the three requirements in frame copy-move forgery detection. We validated the method on different kinds of videos with two common types of copy-move tampering, i.e. one with unsmooth forgery and one with smooth manipulation. Experimental results show that the proposed detection approach achieves high accuracy under different common attacks with low computation complexity and strong applicability.

Our future research will focus on improving its ability to deal with tampered videos with largely static scene and more careful manipulation. For tampered videos with a largely static scene, we will research on how to describe the extent of video movement to adaptively adjust the parameters and detection process. For carefully prepared copy-move forgery or regional copy-move forgery in videos, we will try to combine different video features and techniques to enrich the method. To create a more convincing and large-scale video forgery dataset is also a goal of future work.

ACKNOWLEDGMENT

The authors are grateful for the editors and anonymous reviewers who made constructive comments and improvements on this paper.

REFERENCES

[1] R. Sekhar and R. S. Shaji, "A methodological review on copy-move forgery detection for image forensics," *Int. J. Digit. Crime Forensics*, vol. 6, no. 4, pp. 34–49, 2014.
[2] A. B. Pawar *et al.*, "Data encryption and security using video watermarking," *Int. J. Eng. Sci.*, vol. 4, no. 4, p. 3238, 2016.
[3] C.-S. Lu and H.-Y. M. Liao, "Structural digital signature for image authentication: An incidental distortion resistant scheme," *IEEE Trans. Multimedia*, vol. 5, no. 2, pp. 161–173, Jun. 2003.
[4] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia, "Feature-based dynamic signature verification under forensic scenarios," in *Proc. Int. Workshop IEEE Biometrics Forensics (IWBF)*, Mar. 2015, pp. 1–6.
[5] C. Feng, Z. Xu, W. Zhang, and X. Yanyan, "Automatic location of frame deletion point for digital video forensics," in *Proc. 2nd ACM Workshop Inf. Hiding Multimedia Secur.*, 2014, pp. 171–179.
[6] P. Bestagini, S. Milani, M. Tagliasacchi, and S. Tubaro, "Local tampering detection in video sequences," in *Proc. IEEE 15th Int. Workshop Multimedia Signal Process. (MMSP)*, Sep./Oct. 2013, pp. 488–493.
[7] S. Milani *et al.*, "An overview on video forensics," *APSIPA Trans. Signal Inf. Process.*, vol. 1, no. 1, pp. 1229–1233, 2012.
[8] Z. Zhang, J. Hou, Q. Ma, and Z. Li, "Efficient video frame insertion and deletion detection based on inconsistency of correlations between local binary pattern coded frames," *Secur. Commun. Netw.*, vol. 8, no. 2, pp. 311–320, Jan. 2015.
[9] J. Chao, X. Jiang, and T. Sun, "A novel video inter-frame forgery model detection scheme based on optical flow consistency," in *Proc. Int. Workshop Digit. Forensics Watermarking*, 2013, pp. 267–281.
[10] J. Yang, T. Huang, and L. Su, "Using similarity analysis to detect frame duplication forgery in videos," *Multimedia Tools Appl.*, vol. 75, no. 4, pp. 1793–1811, Feb. 2016.
[11] X. Jiang, W. Wang, T. Sun, Y. Q. Shi, and S. Wang, "Detection of double compression in MPEG-4 videos based on Markov statistics," *IEEE Signal Process. Lett.*, vol. 20, no. 5, pp. 447–450, May 2013.
[12] W. Wang and H. Farid, "Exposing digital forgeries in video by detecting double MPEG compression," in *Proc. 8th Workshop Multimedia Secur.*, 2006, pp. 37–47.
[13] J. Zhao and J. Guo, "Passive forensics for copy-move image forgery using a method based on DCT and SVD," *Forensic Sci. Int.*, vol. 233, nos. 1–3, pp. 158–166, 2013.

[14] C.-S. Lin, C.-C. Chen, and Y.-C. Chang, "An efficiency enhanced cluster expanding block algorithm for copy-move forgery detection," in *Proc. Int. Conf. Intell. Netw. Collaborative Syst. (INCOS)*, Sep. 2015, pp. 228–231.
[15] D. Chauhan, D. Kasat, S. Jain, and V. Thakare, "Survey on keypoint based copy-move forgery detection methods on image," *Procedia Comput. Sci.*, vol. 85, pp. 206–212, Jan. 2016.
[16] W. Wang and H. Farid, "Exposing digital forgeries in video by detecting duplication," in *Proc. 9th Workshop Multimedia Secur.*, 2007, pp. 35–42.
[17] Q. Wang, Z. Li, Z. Zhang, and Q. Ma, "Video inter-frame forgery identification based on consistency of correlation coefficients of gray values," *J. Comput. Commun.*, vol. 2, no. 4, p. 51, 2014.
[18] Y. Liu and T. Huang, "Exposing video inter-frame forgery by Zernike opponent chromaticity moments and coarseness analysis," *Multimedia Syst.*, vol. 23, no. 2, pp. 223–238, Mar. 2017.
[19] S.-Y. Liao and T.-Q. Huang, "Video copy-move forgery detection and localization based on Tamura texture features," in *Proc. 6th Int. Congr. IEEE Image Signal Process. (CISP)*, vol. 2. Dec. 2013, pp. 864–868.
[20] G. Lin and J. Chang, "Detection of frame duplication forgery in videos based on spatial and temporal analysis," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 26, no. 7, pp. 1250017-1–1250017-18, 2013.
[21] C.-C. Hsu, T.-Y. Hung, C.-W. Lin, and C.-T. Hsu, "Video forgery detection using correlation of noise residue," in *Proc. IEEE 10th Workshop IEEE Multimedia Signal Process.*, Oct. 2008, pp. 170–174.
[22] M. Kobayashi, T. Okabe, and Y. Sato, "Detecting forgery from static-scene video based on inconsistency in noise level functions," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 4, pp. 883–892, Dec. 2010.
[23] J. Chao, "Research on digital video inter-frame forgery passive detection algorithm based on visual content," M.S. thesis, Shanghai Jiao Tong Univ., Shanghai, China, 2013.
[24] S. Kingra, N. Aggarwal, and R. D. Singh, "Inter-frame forgery detection in H.264 videos using motion and brightness gradients," *Multimedia Tools Appl.*, vol. 76, no. 24, pp. 25767–25786, Dec. 2017.
[25] Z. Zhang, J. Hou, Z. Li, and D. Li, "Inter-frame forgery detection for static-background video based on MVP consistency," in *Proc. Int. Workshop Digit. Watermarking*, 2015, pp. 94–106.
[26] A. V. Subramanyam and S. Emmanuel, "Video forgery detection using HOG features and compression properties," in *Proc. IEEE 14th Int. Workshop IEEE Multimedia Signal Process. (MMSP)*, Sep. 2012, pp. 89–94.
[27] J. A. Aghamaleki and A. Behrad, "Inter-frame video forgery detection and localization using intrinsic effects of double compression on quantization errors of video coding," *Signal Process., Image Commun.*, vol. 47, pp. 289–302, Sep. 2016.
[28] X. J. Yuan *et al.*, "Digital video forgeries detection based on textural features," *Comput. Syst. Appl.*, vol. 21, no. 6, pp. 91–95, 2012.
[29] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artif. Intell.*, vol. 17, nos. 1–3, pp. 185–203, Aug. 1981.
[30] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. Int. Joint Conf. Artif. Intell.*, 1981, pp. 674–679.
[31] A. Bruhn, J. Weickert, and C. Schnörr, "Lucas/Kanade meets horn/schunck: Combining local and global optic flow methods," *Int. J. Comput. Vis.*, vol. 61, no. 3, pp. 211–231, 2005.
[32] J. Yang and A. G. Hauptmann, "Exploring temporal consistency for video analysis and retrieval," in *Proc. ACM Int. Workshop Multimedia Inf. Retr.*, 2006, pp. 33–42.
[33] A. Bidokhti and S. Ghaemmaghami, "Detection of regional copy/move forgery in MPEG videos using optical flow," in *Proc. Int. Symp. Artif. Intell. Signal Process.*, Mar. 2015, pp. 13–17.
[34] C. Feng, Z. Xu, S. Jia, W. Zhang, and Y. Xu, "Motion-adaptive frame deletion detection for digital video forensics," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 271, no. 12, pp. 2543–2554, Dec. 2017.
[35] A. J. Fridrich, B. D. Soukal, and A. J. Lukáš, "Detection of copy-move forgery in digital images," in *Proc. Digit. Forensic Res. Workshop*, 2003, pp. 1–10.

SHAN JIA received the B.S. degree in electronic and information engineering from Wuhan University, Wuhan, China, in 2014, where she is currently pursuing the Ph.D. degree in communication and information system with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing. She is currently a visiting Ph.D. student in computer science and electrical engineering with West Virginia University. Her research interests include digital image and video processing, biometrics, and information security.

ZHENGQUAN XU received the B.S. degree in radio technology and information system and the M.S. degree in communication and electronic system from Tsinghua University, Beijing, China, in 1985 and 1988, respectively, and the Ph.D. degree in biomedicine engineering from The Hong Kong Polytechnic University, Hong Kong, in 1998. He is currently a Professor with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China. His current research interests include multimedia processing, information security, cloud computing security, and biometrics.

HAO WANG received the M.S. degree from South-Central University for Nationalities, Wuhan, China, in 2014. He is currently pursuing the Ph.D. degree in communication and information system with Wuhan University, Wuhan, China. His research interests include data mining, privacy preserving, and information security.

CHUNHUI FENG received the M.S. degree in graphic communication engineering and the Ph.D. degree in communication and information system from Wuhan University, Wuhan, China, in 2010 and 2015, respectively. She is currently with the College of Computer and Information Sciences, Fujian Agriculture and Forestry University, Fuzhou, China. Her research interests include image and video processing and multimedia forensics.

TAO WANG received the Ph.D. degree in communication and information system from Wuhan University, Wuhan, China, in 2015. He is currently with the Collaborative Innovation Center of Geospatial Technology, Wuhan, China. His research interests include information security, big data, and privacy preserving.