Detection of Regional Copy/Move Forgery in MPEG Videos using Optical Flow

Amir Bidokhti
Department of Electrical Engineering
Sharif University of Technology
Tehran, Iran
bidokhti@ee.sharif.edu

Shahrokh Ghaemmaghami
Electronics Research Institute and
Department of Electrical Engineering
Sharif University of Technology
Tehran, Iran
ghaemmag@sharif.edu

Abstract—With rapid proliferation of affordable video capturing devices and state-of-the-art video editing software tools, it is now easier than ever to manipulate video contents. In this paper a passive method for copy/move video forgery detection in MPEG videos is proposed. The method first divides each video frame into suspicious and apparently innocent parts. Subsequently, an optical flow coefficient is computed from each part. Forgeries are located when an unusual trend in the optical flow coefficient of the suspicious object is detected. Experiments on a set of forged and original sequences validate the justifications made by the proposed method.

Keywords—Digital Forensics; Video Forgery Detection; Optical Flow; Multimedia Security; Copy/Move Forgery

I. INTRODUCTION

Following the invention of video cameras, videos were used as sources of evidence. However, by the emergence of digital videos and video doctoring applications, they were not as reliable as before for critical applications such as proofs of crime in courts, etc. Nevertheless, for security purposes, video surveillance systems have become popular and continue to grow. Besides, affordable camera recorders and mobile phones with high quality cameras are proliferating. Due to digital nature of current videos and availability of modern video/image editing tools, even to ordinary people, these digital videos are subject to malicious manipulations.

Video forgery refers to manipulating a video in such a way that changes its content perceptually. Video forgery methods are categorized into inter-frame forgery and intra-frame forgery [1]. In intra-frame forgery methods, some parts of a frame of video are removed from or altered inside the same frame, while in inter-frame forgery methods a frame’s content is changed using the contents of other frames of the same or another video.

A small amount of research has been devoted to video forgery detection. There are two main approaches to this problem: active and passive. In active methods, usually a watermark is embedded into the video at the recording time. If the video undergoes any change, the corresponding watermark is subject to change as well. The verifier can investigate the watermark in order to reveal possible forgeries made to the video. In passive methods nothing is embedded during recording and the verifier must check the suspicious video for possible clues of inconsistency. Active methods are not applicable to general situations because most hardware and software video recorders available in the market do not include a watermark embedding module. This makes the passive method the only choice in most cases.

We propose a passive method for video forgery detection which relies on inconsistencies in optical flow coefficients, which are defined later in this paper. This method first divides frames into two parts and then searches for primary and secondary periodic trends in optical flow coefficients. A secondary periodic trend indicates forgery.

This paper is organized as follows. In section II an overview of passive digital video forgery detection methods is given. Section III covers some basic concepts of the MPEG standard and the optical flow. The proposed method is described in sections IV and our experimental results are presented and discussed in section V. Section VI concludes the paper.

II. RELATED WORKS

Image forgery detection has been widely investigated and several working methods for detecting different types of image forgeries have been proposed in the literature. However a few methods for video forgery detection are available.

One approach to video forgery detection is source identification, i.e., determining the capturing device. If the video has undergone inter-frame forgery, then different parts of the sequence might have been captured using different devices (see [11] and [12]).

Some researchers try to reveal forgery through detecting double compression. Double compression, usually occurs while resaving a video after editing. In contrast to images which are usually edited (resizing, adding special effects, etc.) after capturing, videos are rarely edited. Therefore, double compression could be an indicator of forgery (see [13] and [14]).

Authors in [8] used a temporal artifact emerging while re-compressing an MPEG video after removing some frames. In MPEG videos, P- and B-frames in a GOP (group
of pictures) are predicted based on I-frame of that GOP. During removing process, some of these predicting frames may go to another GOP, causing high prediction errors. These high errors will repeat periodically. That is the frames retained in the same GOP make smaller errors, while the frames moved to other GOPs often show higher error values. We will further discuss key concepts of MPEG standard in section III.

Recently, optical flow has been used for forgery detection [6,7]. We discuss these methods in section IV, particularly because our work is closely related to that given in [6].

III. PRELIMINARY CONCEPTS

In this section some basic concepts used in the proposed method are explained. We first describe fundamentals of the MPEG video compression standard [9] and optical flow concept between two adjacent video frames.

A. MPEG standard

The MPEG standards are widely used in digital multimedia industry. In the MPEG encoding, video frames are grouped into groups of around 15 frames. Such a group is called “GOP” which means Group of Pictures. A GOP can contain the following frame types: I-frame, P-frames and B-frames. Each GOP starts with an I-frame that is encoded independently of other frames. However, to encode P-frames, they are first divided into small blocks and then the motion is estimated between these blocks and a previously encoded frame. Only the motion vector and the difference between matched blocks in two frames are encoded. The procedure is the same for B-frames, in spite of the fact that for these frames, the two neighboring frames (one preceding and one succeeding), are used for motion compensation. Fig. 1 shows a sample GOP structure in the MPEG standard. Note that only I- and P-frames are used for motion compensation.

![Figure 1. Illustration of a sample GOP structure in MPEG standard. Arrows show prediction procedure through motion compensation. Red and black arrows show backward and forward prediction, respectively.](http://example.com/image1)

B. Optical Flow

Optical flow is an important concept in video processing and computer vision. Optical flow is apparent velocity of each pixel of a video frame in transition to the next frame. A velocity vector of two elements is calculated for each pixel in the current frame. We denote the brightness of pixel at location \((x, y)\) at time \(t\) by \(I(x, y, t)\). Assuming that the intensity of the pixel is preserved during transition, the optical flow equation could be derived, as [6]:

\[
I_xu + I_yv + I_t = 0 \tag{1}
\]

where \(I_x\) and \(I_y\) are gradients of image intensity in \(x\) and \(y\) directions, respectively; \(I_t\) is time derivative of frame intensities; \(u\) and \(v\) are horizontal and vertical components of estimated optical flow, respectively. This is a single equation for each pixel with two unknown variables. Therefore the problem is ill-posed. There are different methods used for solving this problem, one of which is Lucas-Kanade [3] method. This method assumes that the optical flow in a small vicinity of a pixel is identical. This assumption leads to the following optimization problem [6]:

\[
\min_{\Omega} \sum_{p \in \Omega} W(p) (I_xu + I_yv + I_t)^2 \tag{2}
\]

in which \(I_x, I_y, I_t, u\) and \(v\) are same as in (1) and \(W(p)\) is a weighting function with more weights at the center and \(\Omega\) is a close small vicinity of the given pixel. Usually Gaussian pyramids of the two frames are constructed and a coarse-fine strategy is utilized in order to estimate optical flow [3].

C. Copy/Move Video Forgery

One type of inter-frame video forgery methods is copy/move forgery. In these methods some frames or parts of frames are moved or copied to another parts of the video. For example a car moving on a road is relocated copied from one point in time to another, while backgrounds of the corresponding frames are left unchanged.

Since the frames in a GOP are compressed through motion compensation and comparison of two adjacent frames, we expect that some periodicity in the amount of prediction error; i.e. lower errors at the beginning of a GOP and higher errors at the end of it. This phenomenon could be investigated for forgery detection. Reordering or moving video content in the time domain and then compressing them with MPEG will cause large errors at unexpected places. This introduces secondary periodicity trend.

IV. OUR METHOD

In [6] authors proposed a method to find inconsistencies in optical flow of frames. First the optical flow is computed for all pixels. Then the sum of magnitude of these optical flow vectors is calculated for all pixels in the frame to obtain \(OF(k)\), as:

\[
OF(k) = \sum_{i=1}^{M} \sum_{j=1}^{N} \sqrt{u(i,j)^2 + v(i,j)^2} \tag{3}
\]

where \(M\) and \(N\) are frame width and height, respectively. Another ratio for each frame, except for the first and the last frames, is introduced as:

\[
\alpha(k) = \frac{2 \cdot OF(k)}{OF(k-1)+OF(k+1)} \tag{4}
\]

where \(k\) is the frame number. Due to smoothness of motion in genuine videos, \(OF(k)\) changes smoothly between two neighboring frames and \(\alpha(k)\) has values close to 1. Conversely, forged videos obtained through copy/move
Forgeries tend to have larger or smaller values of $\alpha(k)$ in some frames. By investigating $\alpha(k)$ for all frames, the authors were able to detect different types of forgeries like frame deletion, frame insertion, and frame duplication.

Our contribution is threefold: a) focusing on partly copy/move attacks, b) dividing frames into suspicious and innocent parts as a preprocessing step, and c) detecting and describing a new forgery indicator.

### A. Focusing on partly copy/move attacks

In some cases, only a portion of video frames are copied, deleted, or inserted in another place. For example, in the sequence 07_forged.avi of REWIND dataset for video forgery [10], detection of only a rectangular part of frames, containing road and motion, has undergone forgery. In this paper we focus on this type of video forgery.

### B. Concept of “Region of Interest”

We introduce the concept of region of interest or ROI in this work. ROI is defined as a region with which attackers are more willing to tamper. It is also assumed that copy/move forgeries impose some inconsistencies on the video contents, particularly in tampered areas. Based on this assumption, we divide each video frame into two parts: suspicious areas (Region of Interest or ROI) and the remaining areas. Then, we try to compare forgery indicators in these two parts. Splitting video frames into suspicious and innocent parts is a general idea which is applicable to other forgery detection methods. It is incorporated in the proposed method as well.

It is not easy to define a general measure for automatically determining ROI. However, some intuitive reasoning could be made. For example, regions with large motion are more likely to be ROI because they convey a great deal of perceptual information in a video sequence. This could be justified to some extent in case that a moving object is removed from a scene because even in this case the forgery is done in a place which most probably contains motion in adjacent frames. A good example is recordings of surveillance cameras of urban roads.

For static scene videos, we can evaluate motion in some consecutive frames and consider parts with large motions as ROI. By investigating forged sequences in REWIND database, we observe that forgeries are often done in places with known geometric shapes. Moreover even if forged areas are not known shapes, they can be approximated with known shapes and results will be better than considering the frames as a whole. Therefore, it is reasonable to search for ROI’s with shapes such as square, rectangle, circle, polygons, etc.

### C. Describing secondary peaks in $\alpha(k)$

In section III-A, we talked about periodic trends found in prediction error of an MPEG video. During our experiments same trends were found in $\alpha(k)$ and a large value of $\alpha(k)$ at the beginning of each GOP was found. This phenomenon is due to compression noise rather than to a natural reason. If some parts of a frame are displaced in time and this displacement is not a multiplier of GOP length, then secondary periodic peaks will be introduced to $\alpha(k)$ coefficients. If a part of video frame is kept unchanged and another part of frame is replaced with fake materials, then their $\alpha(k)$ will be different. If we consider a time frame of length equal to the length of GOP, the unchanged part will have only one peak in this time frame and the forged parts will have secondary peaks as well. This phenomenon is clearly demonstrated in Fig. 2. Note that we have separated the two plots by adding 1 to one of them, for ease of comparison. Also, we have smoothed $\alpha$ sequence with a moving average filter of length 3, for the peaks to be more observable. In the top plot (\(\alpha(k)\) for ROI) green and red arrows show primary and secondary peaks, respectively.

![Figure 2. Plots of $\alpha$ sequence (smoothed with a moving average filter) for ROI and other parts of video. For demonstration purposes, 1 is added to $\alpha$ for ROI.](image)

**Algorithm 1 – Forgery Detection Algorithm**

<table>
<thead>
<tr>
<th>Input: Video Frames, ROI mask</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output: Video is forged or genuine?</td>
</tr>
<tr>
<td>Step 1. Extract $N$ video frames.</td>
</tr>
<tr>
<td>Step 2. Compute optical flow between each two adjacent frames using Lucas-Kanade method.</td>
</tr>
<tr>
<td>Step 3. Compute $OF_{ROI}(k)$ for $k = 1, ..., N - 1$ using (3)</td>
</tr>
<tr>
<td>Step 4. Compute $OF_{other}(k)$ for $k = 1, ..., N - 1$ using (3)</td>
</tr>
<tr>
<td>Step 5. Compute $\alpha_{ROI}(k)$ for $k = 2, ..., N - 2$ using (4)</td>
</tr>
<tr>
<td>Step 6. Compute $\alpha_{other}(k)$ for $k = 2, ..., N - 2$ using (4)</td>
</tr>
<tr>
<td>Step 7. Apply a moving average filter with length 3 to $\alpha_{ROI}$ and $\alpha_{other}$ coefficients.</td>
</tr>
<tr>
<td>Step 8. Check for secondary peaks in $\alpha_{ROI}$ and $\alpha_{other}$:</td>
</tr>
<tr>
<td>- If secondary peaks are found then the video is “forged”; otherwise it is “genuine”</td>
</tr>
</tbody>
</table>

As shown in Fig. 2, in the intervals between frames 90 to 125 and from 140 to 180, there are no clear periodic trends in $\alpha$ sequence of ROI. This is due to a fundamental change in the video content: these intervals have large motions in the ROI.
In our experiments, we first determine the suspicious parts of frames which are more likely to be forged. Then \( \alpha(k) \) is calculated for suspicious parts and remaining parts. After that we search for primary and secondary periodic peaks in \( \alpha \)'s. This procedure is described concisely in Table 1. The ROI mask, which is given to Algorithm 1, is a binary matrix of size equal to each video frame, with ones indicating pixels in ROI.

An important step in Algorithm 1 is to determine whether the secondary peaks exist in the sequence of \( \alpha \) coefficients. Autocorrelation is used for finding periodic trends in time series. Autocorrelation of a sequence takes its maximum value at 0. For a quasi-periodic sequence, there will a single peak in the autocorrelation and other values will be small. On the other hand, when the secondary peaks are present in the sequence, they will weaken the effect of the primary peaks of the autocorrelation, so they will have smaller values. Having this in mind, we define a ratio between first and second largest values of \( \alpha \) sequences’ autocorrelation, as:

\[
\Lambda = \frac{\text{Largest value of autocorrelation function of } \alpha}{\text{2nd largest value of autocorrelation function of } \alpha}
\]  

\( 5 \) We expect that original sequences will have larger \( \Lambda \)'s. So, we are able to separate forged and original sequences by a simple threshold over \( \Lambda \) value of each sequence. Detailed procedure is explained in Table 2. In our experiments, we observed that a good value for \( thr \) would be 1.1. Larger values for \( thr \) lead to more “original” indicator and therefore higher false negatives are encountered. Similarly, smaller values for \( thr \) lead to higher false positives.

It is notable that \( \Lambda \) is calculated for periods of 30 to 45 frames (approximately 3 GOPs in MPEG encoding). It is to suppress the effects of temporal events such as a moving car, etc.

### Table 2. Secondary Peak Detection Algorithm

<table>
<thead>
<tr>
<th>Algorithm 2 – Secondary Peak Detection</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong> Two time series, ( \alpha_{ROI} ) and ( \alpha_{other} )</td>
</tr>
<tr>
<td><strong>Output:</strong> Do secondary peaks exist?</td>
</tr>
</tbody>
</table>

**Step 1.** Calculate autocorrelation function of time series: 
\[ R_{\alpha,ROI}[n] \text{ and } R_{\alpha,other}[n] \text{ for } n = 1, ... , 30. \]

**Step 2.** Compute \( \Lambda_{ROI} \) for \( R_{\alpha,ROI} \) using (5)

**Step 3.** Compute \( \Lambda_{other} \) for \( R_{\alpha,other} \) using (5)

**Step 4.** If \( \frac{\Lambda_{other}}{\Lambda_{ROI}} > thr \) then secondary peaks exist, otherwise they do not.

### V. Experiments and Results

There are no large-scale video datasets available for video forgery detection. Nevertheless there are two small datasets available online. One of them is for REWIND project [4] which consists of 10 pairs of original and forged videos of a length about 10 seconds. The other one can be found in Surrey University Library for Forensic Analysis [5] which consists of 5 forged sequences. Fig. 3 shows some sample frames of these databases.

We applied our method and the method in [6] to 10 pairs of forged and original videos from REWIND dataset. We changed and optimized threshold values used in the method in [6] to get the best results. Also we manually designed an ROI mask for each video sequence. For example ROI for “07_forged.avi” sequence is a rectangle in the bottom half of frames. Fig. 4 shows some examples of this ROI mask. In practice, this mask could be designed manually.

To obtain more confident results, we need much more data samples. To this goal, we first label each frame of forged sequences as forged or genuine (note that even in forged sequences, some frames are not manipulated). Then, each sequence is split into groups of 50 overlapping frames. Adjacent frames have 40 common frames. Groups with even forged frame are considered as forged. Next, we apply our algorithm to each of these groups. Final results are summarized in Table 3.

![Figure 3](image316x392to567x558)

**Figure 3.** Sample frames of the sequence 07_original.avi (top row) and 07_forged.avi (middle row). Bottom row is the difference of two frames multiplied by 4. It is clear that a rectangular area of video has been forged. The number below each picture is the frame number.

![Figure 4](image320x145to555x329)

**Figure 4.** Examples of ROI masks used in our experiments. Gray and black areas are suspicious and innocent areas respectively.
The method given in [6] has a poor performance when applied to this video dataset because it assumes that in the forgery process frames are displaced or deleted as a whole, while in this case copy/move forgery is done only in a part of each frame.

TABLE 3: COMPARISON OF OUR METHOD WITH METHOD IN [6].

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Method in [6]</th>
<th>Our Method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Original</td>
<td>Forged</td>
<td>Original</td>
</tr>
<tr>
<td>Original</td>
<td>461</td>
<td>367</td>
<td>94</td>
</tr>
<tr>
<td>Forged</td>
<td>147</td>
<td>35</td>
<td>112</td>
</tr>
</tbody>
</table>

A. When this method does not work

An anti-forensics technique could be used in order to weaken the method proposed in this paper. As mentioned earlier, this method is effective only if the displacement of forged parts is not a multiplier of GOP’s length. So, attackers may constrain the displacement to multipliers of the GOP’s length as a counter-forensics technique.

The proposed method is highly sensitive to ROI selection. For example, if ROI in the forged sequence is selected as a rectangle in the right half of each frame, then the $\Lambda_{ROI}$ and $\Lambda_{other}$ in Algorithm 2 will be nearly the same and no secondary peak will be detected.

Another weakness of this method is that in GOPs with high amounts of motion, the periodic trends are disturbed and the proposed method loses its performance.

VI. CONCLUSION AND FUTURE WORKS

In this paper, we have proposed a passive method for digital video forgery detection. A forgery indicator has been introduced using unusual trends in optical flow coefficients of video frames. We have found that secondary periodic trends are signs of copy/move forgery. The arguments have been justified when the method is applied to ten pairs of forged and genuine video sequences.

REFERENCES


