Neurocomputing 205 (2016) 84–91

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Exposing frame deletion by detecting abrupt changes in video streams

Liyang Yu [a,b], Huanran Wang [b], Qi Han [b,*], Xiamu Niu [b,**], S.M. Yiu [c], Junbin Fang [d], Zhifang Wang [e]

[a] School of Software, Harbin University of Science and Technology, Harbin, China
[b] School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China
[c] Department of Computer Science, The University of Hong Kong, Hong Kong
[d] Department of Optoelectronic Engineering, Jinan University, Guangzhou, China
[e] School of Electronic Engineering, Heilongjiang University, Harbin, China

ARTICLE INFO

Article history:
Received 26 August 2015
Received in revised form
14 March 2016
Accepted 22 March 2016
Communicated by Lu Xiaoqiang
Available online 12 May 2016

Keywords:
Video forensics
Anomaly detection
Frame deletion detection
Video stream analysis

ABSTRACT

Many existing methods for frame deletion detection attempt to detect abnormal periodical artifacts in video stream, however, due to a number of reasons, the periodical artifacts can not always be reliably detected. In this paper, we propose a new method for frame deletion detection. Rather than detecting abnormal periodical artifacts, we devise two features to measure the magnitude of variation in prediction residual and the number of intra macro blocks. Based on the devised features, we propose a fused index to capture abnormal abrupt changes in video streams. We create a dataset which consists of 6 subsets, and test the detection capability of our method in both video level and GOP (Group of Pictures) level. The experimental results show that the proposed method performs stably under various configurations.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

With different kinds of video editing software, it becomes more and more easier for people to distort the content of digital videos. Particularly, an attacker can remove events by simply deleting a group of video frames. For example, the attacker could delete a sequence of consecutive frames where a person is walking through a scene, to destroy the evidence that the man was present for some accident. Such operations are rather simple for even non-expert users. In this sense, it is necessary to check the integrity of a given video clip when it is used as an evidence in some serious scenarios, e.g. in the court.

The authenticity of a given video clip can be examined by active tools such as digital watermarking [1] or media fingerprints [2]. However, in most cases, the watermarking and fingerprints are not available. As an alternative to the previous active approaches, a number of passive techniques have been proposed during the past decade. To detect frame deletion, in [3], Wang and Farid observe that in MPEG-1- or MPEG-2-coded videos, frame deletion can result in periodic increase in prediction residual of P frames, and peaks in the middle frequency region in the Fourier domain can expose forgeries. Since [3] only applies to videos with fixed GOP (Group of Pictures) size and the peaks are detected by manual inspection, Stamm et al. extend [3] in [4], the improved method applies to videos with both fixed and adaptive GOP sizes, and is able to automatically detect the forgery trace. The methods in [3,4] can be used to judge whether a given video has been tampered, but the location of frame deletion can not be detected. Instead of prediction residual increase, [5] uses periodical increase in the number of I-MBs (intra macro blocks) in P frames to expose forgery. The previous methods suffer from the same limitations: the periodical spikes can be easily concealed by the inherent fluctuations in the residual signal, and on the other hand, when the frame deletion point is close to the end of the video, or the resulted video is of static scene (e.g., as mentioned above, deleting the frames that a man passing a scene in a surveillance video, which is quite common and simple to perform), the periodical inconsistency can hardly be detected. Besides the above mentioned limitations, the method in [5] relies on excessively strict assumptions: the GOP sizes used in two compressions must be different, and the input video should contain only I and P frames.

* Corresponding author.
** Principal Corresponding author.
E-mail addresses: qi.han@hit.edu.cn (Q. Han), xm.niu@hit.edu.cn (X. Niu).

http://dx.doi.org/10.1016/j.neucom.2016.03.051
0925-2312/© 2016 Elsevier B.V. All rights reserved.

To overcome the previous problems, in this paper, we devise 2 features which are able to capture anomaly in video stream. Based on these features, we construct a fused index to detect frame deletion. Our method does not rely on any assumptions with respect to the encoding process such as those in [5], therefore it is more applicable to realistic forensics. Moreover, our method does not need to find inconsistent periodic artifacts, which makes it more robust against inherent fluctuations in video streams.

The rest of this paper is organized as follows: In Section 2, we introduce necessary background knowledge. In Section 3, we briefly review the techniques for frame deletion detection. The proposed method is detailed in Section 4. We present the experimental results in Section 5 and conclude this paper in Section 6.

2. Basic concepts

For the mainstream compression standards such as MPEG 2, 4, and H.264, a video sequence to be compressed is firstly segmented into groups of pictures (GOP). One GOP typically contains one I frame and a number of P and B frames. One GOP must start with an I frame, which is independently encoded, i.e., I frames can be decoded without referencing to any other frames. P frames are predicted by calculating the motion according to the previous I or P frame. The difference between the predicted frame and the original frame is referred to as prediction residual (PR). Since PR is subject to lossy compression in the encoding pipeline, the reconstructed pictures are typically different from their uncompressed counterparts. B frames are predicted by the immediate previous and future I or P frames.

The frames are divided into macro blocks (MB) before compression, and the MBs can be roughly classified into three types: P-MBs, which are predicted by the MBs in the reference frames, I-MBs, which are generated with the information of the frame to which they belong, and S-MBs, which can be directly copied from the MBs in the reference frame.

3. Related work

Besides the methods discussed in Section 1, in [6,7] the authors exploit abnormal variation of MCEA (motion-compensated edge artifact) to detect forgeries, and these methods are not suitable for the videos whose frames are quite similar to each other. In [8], the author proposes a machine learning-based method, however, being similar to [3,4], this method only judges whether an input video has been tampered or not, while the frame deletion location can not be located. Since intra- and inter-frames use different quantization matrices during the quantization stage, the authors of [9] argue that, in a tampered video, the B and P frames which are previously coded as I frames will miss more high frequency energy during the first compression, therefore inconsistencies in the energy of the high frequency components in neighboring B or P frames indicate deletion operations. This method applies only to the MPEG-2 videos.

Recently, two content-based methods have been proposed. In [10,11], the authors use optical flow to trace the variation across frames, and abrupt changes in optical flow indicate frame deletion. These two methods are rather slow and become quite weak when dealing with compressed videos.

4. Proposed method

When a video clip is re-compressed after being deleted a sequence of frames, if the deleted range does not happen to be one or more entire GOPs (deleting exactly one or more entire GOPs should be of rather low probability, considering that the deleted range depends on the content of the video), the remaining frames will be re-organized into new GOPs during the second compression. Please see an example in Fig. 1. When the attacker delete the last 5 frames in the second GOP and re-compress the resulted video, the first 5 frames in the third GOP of the first compression are moved forward into the second GOP, and the sixth frame of the third GOP becomes the I frame. It is obvious that the same displacement happens in all the succeeding GOPs. In theory, when a frame is reallocated into a new GOP and re-encoded as P frame during the second compression, the correlation between this frame and its reference frame become weaker [3]. As a consequence, the displacement of frames leads to the periodical increase in both PR and NIMBs in P frames. However, the expected abnormal periodical increase is difficult to detect for several reasons. Firstly, the PR is content-related and therefore inherent fluctuations in the PR images unavoidably conceal the periodic artifacts. Secondly, when the location of frame deletion is close to the end of the video clip, the period can not be effectively observed. Thirdly, if the content of the tampered video is static, except for a handful of (typically one or two) spikes within a rather small neighborhood of the deletion location, there will be hardly any spike. We show for each situation a typical example in Fig. 2. In Fig. 2(a), the periodical increase in PR can not be correctly estimated due to the inherent fluctuation of the PR values. In Fig. 2 (b), the forgery location is quite close to the end of the video, as a result, the number of P frames succeeding the forgery point is too small to exhibit the periodical increase. In Fig. 2(c), the content of the forged video is a static scene, which is obtained by deleting the frames with motion. In this kind of videos, except for the frames locating within a small neighborhood (the 17-, 18- and 19-th P frames) of the forgery location, the mean values of PR for most frames are about zero, therefore no periodical artifacts can be observed.

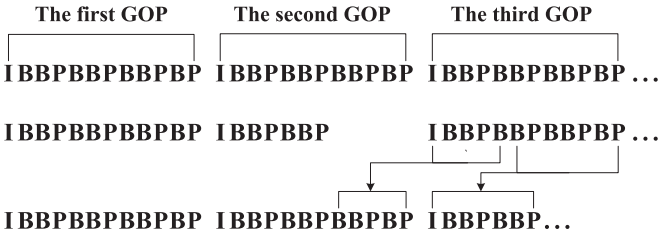

Fig. 1. An example of GOP re-organization after frame deletion. Shown in the figure is a standard H.264 GOP. Top row: the original GOPs. Middle row: the last 5 frames of the second GOP are deleted. Bottom row: the re-organized GOPs.

Although the abnormal periodical artifacts can not always be reliably detected, the frame deletion operations can be exposed by abrupt changes in the video stream. Let P_i denote the i-th P frame, and R_i and PR_i denote P_i's reference frame (i.e., the immediate previous P or I frame of P_i) and P_i's PR image, respectively, according to the encoding principal in popular video encoding standards such as MPEG-2, 4 and H.264, we have

$$P_i = \mathcal{M}(R_i) + PR_i, \tag{1}$$

where $\mathcal{M}(\cdot)$ denotes the motion compensation operation. Therefore,

$$PR_i = P_i - \mathcal{M}(R_i). \tag{2}$$

Suppose $U_{P,i}$ and $U_{R,i}$ are the uncompressed versions of P_i and R_i, respectively, since the I frames and the PR of P frames are subject to lossy compression, Eq. (2) can be re-written as:

$$PR_i = U_{P,i} + N_{P,i} - \mathcal{M}(U_{R,i} + N_{R,i}), \tag{3}$$

Download English Version:

https://daneshyari.com/en/article/405683

Download Persian Version:

https://daneshyari.com/article/405683

Daneshyari.com