JOURNAL OF INFORMATION SECURITY AND APPLICATIONS 26 (2016) 39–50

Available online at www.sciencedirect.com

ScienceDirect

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

Detecting video frame-rate up-conversion based on periodic properties of edge-intensity

Yuxuan Yao [a], *Gaobo Yang* [a,*], *Xingming Sun* [b], *Leida Li* [c]

[a] *School of Information Science and Engineering, Hunan University, Changsha 410082, China*
[b] *School of Computer and Software, Nanjing University of Information Science & Technology, Nanjing 410082, China*
[c] *School of Information and Electrical Engineering, China University of Mining and Technology, Xuzhou 221116, China*

ARTICLE INFO

Article history:
Available online 6 January 2016

Keywords:
Video forensics
Frame-rate up-conversion (FRUC)
Video edge-intensity
Kaufman adaptive moving average (KAMA)

ABSTRACT

Video frame-rate up-conversion (FRUC) is one of the common temporal-domain operations. From the earlier frame repetition and linear interpolation, FRUC has been developed to motion compensated frame interpolation (MCFI), which effectively overcomes the temporal jerkiness and ghosting shadows. In a broad sense, FRUC can be regarded as a video forgery operation. By experiments, it is observed that FRUC still leads to edge discontinuity or over-smoothing artifacts around object boundaries. In this paper, an edge-intensity based passive forensics approach is proposed to detect the possible FRUC operation in candidate video. After computing the edge intensities of every frame, Kaufman adaptive moving average (KAMA) is exploited to define an adaptive threshold to distinguish the interpolating frames by FRUC from the original frames. Moreover, the original frame-rate of up-converted video can be inferred. Experimental results show that the proposed approach is not only effective for simple frame repetition and linear interpolation, but also valid for advanced FRUC techniques such as MCFI. The detection accuracy is up to 94.5% on average. Its computation is simple as well.

1. Introduction

The rapid development of various video editing tools such as VideoEdit facilitates the improvement of visual quality for digital video. However, it is becoming much easier for video tampering and forgery as well. As a consequence, the forensics techniques are anticipated to verify the authenticity and integrity of digital video (Edward et al., 2009; Li et al., 2015). Active forensics techniques require pre-embed auxiliary information such as digital watermark into videos (Tian et al., 2015), or pre-designed side information such as forensics hash (Wei et al., 2015) in advance, and then the tampering is determined then by detecting the integrity of pre-embedded auxiliary information or pre-designed side information. Passive forensics is to detect the inconsistent regularities or specific artifacts of digital video for forgery detection (Milani et al., 2012). For example, Wang and Farid (2009) proposed a video forgery detection approach by exposing double quantization artifacts. In addition, Wang and Farid (2007) proposed a forensic approach to detect traces of tampering in interlaced and de-interlaced videos. Subramanyam and Emmanuel (2012) presented a blind detection approach for the spatial and temporal copy paste tampering, which is based on Histogram of

* *Corresponding author.* School of Information Science and Engineering, Hunan University, Changsha 410082, China. Tel.: +86 0731 8882 3141; fax: +86 0731 8882 1907.
E-mail address: yanggaobo@hnu.edu.cn (G. Yang).
http://dx.doi.org/10.1016/j.jisa.2015.12.001

Oriented Gradients (HOG) feature matching and video compression properties. Subramanyam and Emmanuel (2013) also proposed a double quantization detection approach by exploiting the principles of estimation theory. Each pixel of a given frame is estimated from the spatially colocated pixels of all the other frames in a Group of Picture (GOP). The error between the true and estimated value is subjected to a threshold to identify the double compressed frame or frames in a GOP. Apparently, passive video forensics does not need any priori information, which makes it more suitable for practical applications. Thus, passive video forensics is recently becoming one of the hottest topics in the field of video information security (Rocha et al., 2011).

Compared with still image, digital video brings extra temporal dimension. Frame-based video manipulation is specific to digital video. Until now, there are some representative works about the blind forensics of frame duplication, frame deletion and frame-adding. Yang et al. (2014) presented a detection approach for frame duplication forgery using frame-level similarity analysis. Shanableh (2013) proposed a machine learning based detection approach for frame deletion. The discriminative features are based on prediction residuals, percentage of intra-coded macroblocks, quantization scales and reconstruction quality. Wang and Li (2014) presented an inter-frame forgery identification approach based on the consistency of correlation coefficients of gray values. Moreover, velocity field consistency and optical flow consistency are exploited to expose the video inter-frame forgery, respectively (Chao et al., 2013; Wu et al., 2014). Zernike opponent chromaticity moments and coarseness analysis are also exploited to expose video inter-frame forgery (Liu and Huang, 2015). Stamm et al. (2012) presented a theoretical model of the forensically detectable fingerprints that frame deletion or addition leaves behind. This model is further exploited in temporal forensics and anti-forensics for motion compensated video and better detection performances are achieved for frame deletion or addition than the approach by Wang and Farid (2007).

Frame-rate up-conversion (FRUC) is the procedure of increasing the frame-rate of a video by temporal interpolation of frames (Choi et al., 2000; Kang et al., 2007). From simple frame repetition and linear interpolation, FRUC has been developed to advanced motion compensated frame interpolation (MCFI) techniques. By introducing all kinds of assumptions to refresh the motion vectors and optimize the texture, MCFI can achieve better visual quality of the resultant video (Xue et al., 2015; Yoo et al., 2013). Apparently, FRUC is a special frame-adding operation, which is originally proposed to improve the visual quality of low frame-rate video. However, FRUC can also be used for video forgery purpose. For example, when two videos with different frame rates are needed to be spliced together, the low frame-rate video is usually up-converted by FRUC to match the relatively high frame-rate video. In recent years, the detection of video FRUC has attracted the attention from the community of video information security. Bian et al. proposed a similarity-analysis-based detection approach for frame duplication (Bian et al., 2014). After dividing the video sequence into overlapping sub-sequences, the similarities between the sub-sequences are calculated, which are exploited to identify those video sequences with high similarity as candidate duplication frames. However, it only reports the detection results of simple FRUC approaches such as frame repetition and linear weighting average. Moreover, Huang and Chen (2011) propose to a video forgeries detection approach based on bidirectional motion vectors. It does not investigate advanced FRUC techniques as well.

Actually, there are lots of advanced MCFI approaches in recent years (Choi et al., 2000; Kang et al., 2007; Stamm et al., 2012; Yoo et al., 2013). They consider the motion between successive frames by overlapped block motion compensation (OBMC) and adaptive motion compensation. These advanced FRUC techniques obtain more natural and realistic videos, which consequently bring extra technical challenges for their passive forensics. In essence, FRUC is a special type of frame-adding operation. The interpolated frames are obtained by block-based average, no matter whether the inter-frame motion is compensated or not. Therefore, FRUC inevitably leads to blurring artifacts to some extent, especially for those pixels near the edge. That is, the edge intensity might be decreased for those interpolated frames. Moreover, since the interpolated frames are periodically inserted into the original frames, the frame-level edge intensity of up-converted video will exhibit some periodicity along the temporal axis. Motivated by this, a novel passive forensics approach is proposed to detect FRUC by exploiting the temporal periodicity of frame edge intensity. A local adaptive threshold is determined by using Kaufman adaptive moving average (KAMA) (Kaufman, 1995). Since the adaptive threshold considers the dynamic change of video content, it can expose the abnormal change of edge intensity caused by frame interpolation and differentiate interpolated frames from the original frames. Moreover, since FRUC inserts the interpolation frames into the original frame periodically, the frame rate of original video sequence can also be inferred.

The rest of this paper is organized as follows. Section 2 briefly describes the FRUC techniques, and comparisons are made among them in terms of the visual qualities of interpolated frames. Section 3 presents the proposed blind detection approach. Section 4 discusses the experimental results and analysis. We conclude this paper in Section 5.

2. Preliminaries of video FRUC techniques

The simplest FRUC techniques are frame repetition (FR) and frame averaging (FA). They do not consider the motion between successive frames, which easily lead to temporal jerkiness and Ghosting shadow for non-static regions. To improve the visual quality, advanced MCFI techniques have been proposed in recent years. The basic idea behind MCFI is to estimate the motions as close as possible to the true motions by introducing various assumptions. Thus, more complex searching pattern is designed or the estimated motion vectors are refined to make the resultant up-converted video more realistic.

2.1. *Simple FRUC techniques*

Simple video FRUC approaches include frame repetition and linear interpolation (Tekalp, 1995), which can be modeled with a weighted linear averaging of forward and backward reference frames. That is,