Forensic Science International 236 (2014) 164–169

Contents lists available at ScienceDirect

Forensic Science International

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

Detection of object-based manipulation by the statistical features of object contour

Chen Richao, Yang Gaobo *, Zhu Ningbo

School of Information Science and Engineering, Hunan University, Changsha 410082, China

ARTICLE INFO

Article history:
Received 29 January 2012
Received in revised form 18 December 2013
Accepted 20 December 2013
Available online 7 January 2014

Keywords:
Video forensics
Passive forensics
Object-based forgery
Video in-painting
Object detection

ABSTRACT

Object-based manipulations, such as adding or removing objects for digital video, are usually malicious forgery operations. Compared with the conventional double MPEG compression or frame-based tampering, it makes more sense to detect these object-based manipulations because they might directly affect our understanding towards the video content. In this paper, a passive video forensics scheme is proposed for object-based forgery operations. After extracting the adjustable width areas around object boundary, several statistical features such as the moment features of detailed wavelet coefficients and the average gradient of each colour channel are obtained and input into support vector machine (SVM) as feature vectors for the classification of natural objects and forged ones. Experimental results on several videos sequence with static background show that the proposed approach can achieve an accuracy of correct detection from 70% to 95%.

1. Introduction

In the era of digital media, the proliferation of image and video editing tools makes the tampering or forgery of digital media much easier. Even ordinary users can produce forged digital media and spread them over Internet for malicious purposes. This leads to an increasing concern about the trustworthiness of public digital media [1]. To verify the authenticity, originality and integrity of digital media, digital media forensics arises to analyse, collect and preserve evidences from digital media. The existing techniques for digital media forensics can be divided into two categories: active and passive forensics [2]. Compared with active forensics, passive forensics does not need any data such as digital watermark or signatures. Thus, passive forensics is becoming a hot research topic in the field of information security.

Compared with digital image, the tampering of digital video is often more sophisticated and time-consuming. However, it is becoming easier with the popularity of video editing tools, such as Video Edit Magic. In the literature, there are many works about digital image forensics [3,4]. However, the research on digital video forensics is still in its infancy. The most representative works are summarized as follows: (1) forensics by the inconsistent trails during the imaging process such as PRNU [5], noise level functions [6]; and (2) forensics by the traces of video tampering, such as ghosting shadow [7], block artefacts [8], GOP periodicity [9] and motion compensated edge artefacts (MCEA) [10]. These methods are effective to detect traditional forgery operations, including copy–paste, double MPEG compression and frame-based tampering.

Object-based manipulations are usually malicious for digital video. For example, if an object is added into, or deleted from digital video, it might have direct influence on the content of digital video that it conveys [11]. Digital video is often believed to provide stronger forensic evidence than still images. As a consequence, the forensics of digital video is extremely important, especially when it used for legal evidence or news report. However, there is still few work reported in literatures about the passive forensics of object-based forgery in digital video to the best of our knowledge. In fact, object-based manipulations will inevitably leave some splicing traces [12], which are resulted from the limited accuracy of video object detection and extraction. Therefore, the statistical features within the boundary areas near video object will be inconsistent. This provides valuable clues for passive video forensic. In this paper, we are motivated to propose a passive video forensic method for object-based tampering. The statistical properties of video object and its variable-width boundary areas are fully utilized to determine the classification of natural objects and forged ones.

The rest paper is organized as follows. In Section 2, motion object is detected from static background-by-background subtraction technique, and then the object boundary is located. In Section 3, the statistical features of variable-width object area are

* Corresponding author at: School of Information Science and Engineering, Hunan University, Lushan South Road, Changsha 410082, Hunan, China.
Tel.: +86 0731 88821341.
E-mail addresses: yanggaobo@hnu.edu.cn, gbyang_hunu@hotmail.cm (Y. Gaobo).

0379-0738/$ – see front matter © 2014 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.forsciint.2013.12.022

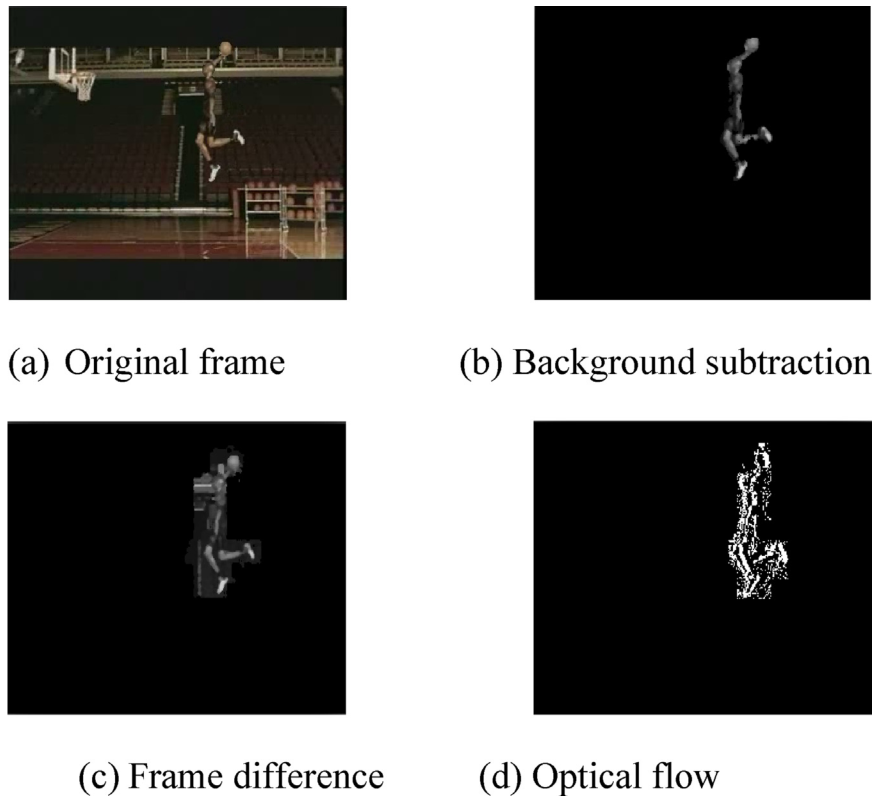

Fig. 1. Comparison of object detection methods: (a) original frame, (b) background subtraction, (c) frame difference, and (d) optical flow.

extracted and input into support vector machine (SVM) for pattern classification. Experimental results are reported and discussed in Section 4. We conclude the paper in Section 5.

2. Video object detection

Object-based video tampering refers to the generation of faked videos by adding, deleting or altering new video object. It usually consists of object detection/tracking, object manipulation, video in-painting and video layer fusion [13]. Therefore, object detection is the first step for digital video forensics to locate the object contour and its bounding areas. Then, the statistical features are extracted from the bounding areas around the object contour. With the help of pattern classifier, the originality and integrity of digital video is verified.

For motion object detection, the most conventional methods are optical flow, frame difference and background subtraction [14]. Optical flow method can obtain accurate diction results when tracking fast-moving object, but with intensive computation. Frame difference method is computationally efficient but very sensitive to scene change such as illumination. Therefore, it is relatively reasonable to choose background subtraction for motion object detection, especially for those video with static background. By establishing appropriate background model, the cumulative average of background frame can be obtained. Thus, motion object can be detected by making the difference between current frame and background frame. Apparently, the key issue for background subtraction technique is the background modelling and updating to adapt the external environment change. Among these background models, Gaussian Mixture Model (GMM) is most widely used [15]. It is a probabilistic approach that uses a mixture of normal distributions to model a multimodal background. For each pixel, each normal distribution in its background mixture corresponds to the probability of observing a particular intensity or colour in the pixel.

Fig. 1 shows the experimental results of *Jordan* sequence by the above-mentioned three object detection methods. Apparently, background subtraction method achieves the best object detection because the obtained object contour is more smooth and accurate. This will be beneficial to the successive statistical feature extraction from object contour and its bounding areas, and then the final classification result for passive forensics will be greatly improved.

After object-based tampering such as object removal, the structure in-painting, texture in-painting or combined structural and textural in-painting are usually performed to remove the motion artefacts. However, there are still some left traces for object-based video forgery, which always exist near the object boundary and its boundary areas. In our earlier work of object extraction, a new concept of adjustable width object boundary (AWOB) is introduced by mathematical morphology [16]. Let l be the extracted binary object, $\oplus$ be the dilation operation, and δ_s be the symmetric structure element, AWOB is defined as follows:

$$AW\,OB = \underbrace{\delta_s \oplus (\ldots \delta_s}_{n\ \text{times}} \oplus (l)) \qquad (1)$$

From Eq. (1), it is obvious that the object area gets larger with the increase of times n for dilation. In Fig. 2, an example is given for the AWOB generation of the detected object by background subtraction (in Fig. 1), where n equals 2.

3. The statistical features extraction

Because the gap between semantic object and low features used in object detection and extraction, there are always some irregularities near object boundary in the process of object-based tampering. Especially, when no dedicated video in-painting is performed after the object-based manipulation, there will be some subtle tampering artefacts near the object boundary and its