Signal Processing: Image Communication 47 (2016) 289–302

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

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

Inter-frame video forgery detection and localization using intrinsic effects of double compression on quantization errors of video coding

Javad Abbasi Aghamaleki, Alireza Behrad *

Department of Electrical Engineering, Faculty of Engineering, Shahed University, Tehran, Iran

ARTICLE INFO

Article history:
Received 21 September 2015
Received in revised form
3 June 2016
Accepted 5 July 2016
Available online 9 July 2016

Keywords:
Double compression
Frame deletion or insertion
Passive forensics
Quantization error
Residual error
Video forensics

ABSTRACT

In this paper, a new passive approach is proposed for tampering detection and localization in MPEGx coded videos. The proposed algorithm can detect frame insertion or deletion and double compression with different GOP structures and lengths. To devise the proposed algorithm, the traces of quantization error on residual errors of P frames are mathematically studied. Then, based on the obtained guidelines, a new algorithm is proposed to detect quantization-error-rich areas in the P frames and reduce the effect of motion on residual errors of P frames. Subsequently, a wavelet-based algorithm is proposed to enrich the traces of quantization error in the frequency domain. Finally, the processed and spatially constrained residual errors of P frames are employed to detect and localize video forgery in the temporal domain. Experimental results and a comparison of the proposed method with an existing approach show the efficiency of the proposed algorithm especially for videos with high compression rates.

1. Introduction

Progress in video capturing devices and multimedia editing software has largely facilitated the video tampering capability. Nowadays, most cell phones have been equipped with video and image capturing devices in companion with modern video editing software, which can be easily used for video editing and tampering. This has degraded the general trust to video and image contents and consequently, the validation of video and image contents has become more indispensable. To handle this problem, considerable research efforts have been devoted to determine the originality of video or media contents. The aim of a video forensic algorithm is to determine the validity of video content or discover forgery signs. Various algorithms have been proposed for video forensics including active and passive approaches. In active approaches like watermarking [1,2], a special digital signature is embedded in the video content during video formation process. Then, the forensic analyzers look for a pre-embedded signature in the video content for video validation. Active approaches require extra and generally expensive hardware for video formation, which restrict their use for general purposely used video cameras. On the contrary, passive approaches utilize inherent fingerprints of video formation process for forensic analysis. Recently, significant research efforts have been devoted to passive algorithms due to their efficiency and general practicability.

Milani et al. [3] presented an overview of various passive approaches for video forensics. Depending on the fingerprints of video formation and editing process, they categorized existing techniques into algorithms based on acquisition, compression and editing clues. In another categorization, existing approaches for video forensic are divided into four categories including pixel based, camera based, geometric based and format based methods [4].

Pixel based methods find pixel-level anomalies in video or image contents for tampering detection. Most existing pixel based methods have focused on copy-move and cloning (copy-paste) forgeries. Copy-paste or copy-move forgeries may be used in an intra-frame based manner to conceal an object in a frame. They may also be utilized in an inter-frame based approach to hide a shot in a video. The principal approach to detect copy-paste or copy-move forgeries is to measure the similarity between frames or different regions in frames, which is computationally expensive. In [5], structural similarity between two frames of videos is employed to detect inter-frame or temporal copy-move forgery. Liao and Huang [6] used Tamura texture features for inter-frame copy-move forgery detection. To reduce the computational overhead of the algorithm, eigenvector matrix of the video frames is constructed with Tamura texture features. Then, the eigenvector matrix of video is sorted and by computing the differences between the adjacent feature vectors, the copy-move forgery is detected. In [7] scale invariant feature transform (SIFT) features and correlation

* Corresponding author.
E-mail address: behrad@shahed.ac.ir (A. Behrad).

http://dx.doi.org/10.1016/j.image.2016.07.001
0923-5965/

between blocks of noise residue frames are employed to detect intra-frame and inter-frame copy-move forgery. Pixel based approaches generally suffer from high computational burden. They are also unable to detect frame insertion or deletion forgeries.

Cameras generally use some common structures and processing pipeline for image formation, where each stage of the pipeline may introduce some special anomalies or artifacts on the resultant image. In camera based approaches, a forensic expert leverages camera introduced artifacts like photo response non uniformity (PRNU) noise [8] or photon shot noise [9] for forensic analysis. Although the camera introduced artifacts like PRNU are successfully used for forensic analysis in uncompressed or low compressed images, their application for video forgery analysis is restricted because of higher compression effects in video coding.

Geometric based approaches employ geometric or physic inconsistencies in the content of an image or video for tampering detection. In [10] a technique was proposed to detect physically implausible trajectories of objects in video sequences. Moreover, various approaches have been proposed to detect geometric inconsistencies using image or frame contents [11–13]. Geometric based approaches are powerful tools for video integrity determination; however, they are restricted to very special image or video contents. Their automatic implementations are very difficult and generally need human supervision.

In [14,15] the influence of inter-frame forgery on motion or velocity of consecutive frames is employed for automatic forgery detection. Wu et al. [15] proposed inter-frame forgery detection by taking into account the velocity field consistencies of video frames. In this algorithm, velocity field intensity (VFI) is calculated by estimating velocity fields using particle image velocimetry (PIV) approach. Subsequently, inconsistencies in VFI are used for inter-frame video forgery detection comprising frame deletion and duplication. Methods based on motion or velocity inconsistencies lose their efficiency for videos with camera motion.

Most existing approaches for video forensics have been focused on format based fingerprints. Generally, video streams are available in compressed format because of their large amount of data for storage and transmission. Format based approaches mostly utilize coding or compression fingerprints to reveal video forgery. To forge a video content, one should decode, apply changes and recompress it. Therefore, recompression and its fingerprints are the main characteristics of a forged video that are used in several algorithms for video forensic analysis.

Intra-coded frames in moving picture experts group (MPEG) videos use a similar principle to joint photographic experts group (JPEG) compression. Inspired by this similarity, some authors have extended JPEG double compression detection algorithms to video forgery analysis. Jiang et al. [16] employed a 162-D feature vector based on the second order statistics of Markov model to detect double compression in MPEG4 videos. Their approach only uses the information of intra-coded frames for feature extraction. A comparison of this method with features based on first digit distribution [17,18] and periodicity of discrete cosine transform (DCT) histogram [19] showed the efficiency of features based on Markov model. Ravi et al. [20] analyzed characteristic of the compression noise to detect double compression. They used modified Huber Markov random field (HMRF) to extract and model the compression noise in the spatial domain. Then the noise of each frame is divided into 16×16 blocks and transition probability matrix (TPM) is obtained in 8 directions. An 18-D feature is subsequently extracted by combining 8 TPMs linearly, and support vector machine (SVM) is finally used for double compression detection. The method cannot expose inter-frame forgery such as frame insertion or deletion.

Tampered videos generally comprise inter-frame forgery like frame insertion and deletion. In representative work by Wang and Farid [21], it was shown that the frame deletion or insertion generates a periodic pattern in residual errors of interceding P frames. Then the Fourier transform of the residual errors of P frames is manually inspected for forgery detection. Their algorithm is not automatic and need a human user inspection for forgery detection. Moreover, the approach is restricted to MPEG videos with a constant number of frames in a group of pictures (GOP). It was also reported that the peaks of forgery are apparent when the number of deleted or inserted frames are the multiple of the sub-GOP length. The algorithm does not give any information about the forged location in video streams.

Stamm et al. [22] employed an automatic approach to detect peaks in the residual errors of P frames in the frequency domain for the detection of frame insertion or deletion. In this work, which is capable of detecting frame insertion and deletion in variable and constant GOP lengths, the residual errors of P frames are first filtered using a median filtering. Then the difference between original and filtered residual errors is employed for forgery detection. As shown later, this approach generates weak results in videos with high compression rates.

Bestagini et al. [23] addressed local tampering detection in video sequences. Their method deals with two restricted forgery scenarios comprising replacing a spatio-temporal region of the video sequence by either a series of fixed image repeating in time or a portion of same video in a different time interval. The method utilizes residual errors of adjacent frames and correlation analysis for tampering detection. The method loses its efficiency for videos with high compression rate or silent scene.

Shanableh [24] proposed an approach to detect frame deletion using a number of discriminative features, which are extracted from video bit stream or the extracted video frames. The effective performances of the features were shown for both variable bit rate (VBR) and constant bit rate (CBR) coding formats. However, the algorithm suffers from its inability to determine the locations of deleted frames. Xu et al. [25] used the periodicity in the histogram of DCT coefficients in double compressed videos to detect video transcoding. They demonstrated that frequency contents of reconstructed DCT coefficients in transcoded video signals have extra peaks because of periodicity. In [26] a method was also proposed for frame deletion or insertion in MPEG video sequences using the power of high frequency components of DCT coefficients. Zhang et al. [27] used the quotients of correlation coefficients between local binary pattern (LBP) coded frames for detecting frame insertion and deletion. It was shown that quotients of correlation coefficients between LBP coded frames show inconsistency for frame insertion and deletion. Vazquez et al. [28] proposed variation of prediction footprint (VPF) for double compression detection and GOP size estimation. The method has affordable computational cost; however, it cannot calculate correct GOP size in the case of inter-frame tampering. To handle the problem, authors in [29] extended VPF to expose frame insertion and deletion by detecting phase discontinuity of the VPF and time domain analysis. The proposed approach is able to detect video forgery even when different codecs are used for the first and second compressions as well as CBR and VBR coding formats. However, the method is restricted to videos with constant GOP size.

In [30] motion-compensated edge artifact (MCEA) was used to detect inter-frame manipulations. It was shown that frame insertion or deletion or changes in the structure of GOP increase MECA difference between adjacent P frames and generate a periodic pattern. Therefore, the spike in frequency domain is used as an evidence of tampering detection. The method cannot determine forgery location.

In this paper, a novel approach is proposed for video forgery detection and localization based on the traces of quantization