ENF Based Video Forgery Detection Algorithm

Yufei Wang, South China University of Technology, Guangzhou, China

Yongjian Hu, South China University of Technology, Guangzhou, China & Sino-Singapore International Joint Research Institute, Guangzhou, China

Alan Wee-Chung Liew, Griffith University, Gold Coast Campus, Australia

Chang-Tsun Li, Deakin University, Geelong Campus, Australia

ABSTRACT

The electric network frequency (ENF) is recorded in the videos taken under the lights powered by grid and can be used for digital forensics. However, due to the lack of data caused by the low frame rate of the video, the ENF-based forensics methods always need a reference signal extracted from the grid, which limits the practical application of these methods. In this article, a new ENF-based time domain video forgery detection algorithm is proposed to solve the problem of data lack. The cubic spline interpolation is used to generate suitable data points of the ENF signal, and the detection sequence generated based on the correlation coefficient between data points in adjacent periods is used to catch the phase continuity interruption of the ENF signal and detect the exact position of forgery. The proposed algorithm can be used independently without any reference signals. The experimental results show that the proposed algorithm has good performance in detecting forgery videos with varying degrees of deletion, duplication and insertion of frames.

KEYWORDS

Electric Network Frequency, Forensics, Time Domain, Videos

1. INTRODUCTION

The electric network frequency (ENF) is the frequency of the power distribution networks. The nominal value of the frequency is 50 Hz or 60 Hz, and it usually fluctuates around the nominal value because of the changing load in the grid. A research pointed out that the range of the fluctuation is ±0.6 Hz (Grigoras, 2005). The ENF signal would affect all devices connecting to the grid, and the effect has a high degree of uniformity within the same grid (Sanders, 2008).

The ENF signal has been used in audio forensics in the latest years. The sound recording equipment powered by the grid will record the ENF signal in the audio file, and this signal can be used as evidence for audio forensics. One approach of ENF based audio forensic algorithms is to compare the ENF signal extracted from the audio with the signal extracted from the grid to identify the generation time and location of the sound recording and to detect any tampering in the audio (Brixen, 2008; Cooper, 2008; Hajj-Ahmad et al., 2005; Hajj-Ahmad et al., 2013; Huijbregtse et al., 2009; Kajstura et al., 2005). In order to use these algorithms, the ENF signals reference databases need to be built (Elmesalawy et al., 2014; Liu et al., 2012). There are also some algorithms using the continuity of the ENF signal to detect forgery in the audio (Nicolalde et al., 2009; Rodríguez et al., 2010). These algorithms are independent of the reference signal and more flexible to use.

DOI: 10.4018/IJDCF.2020010107

The ENF signal has also been used in video forensics. In earlier work, the ENF signal for video forensics was extracted from the audio recorded during video shooting (Cooper, 2011; Grigoras, 2007; Grigoras, 2009), while these methods cannot be used when the video does not contain audio track. Later, some video forensic algorithms based on the ENF signal extracted from video had been proposed (Garg et al., 2011; Garg et al., 2013; Su et al., 2014a). These algorithms need to match the ENF signal from a video to a ENF reference database, and recent research effort has focused on finding new methods for rapid and accurate matching (Su et al., 2014b; Hajj-Ahmad et al., 2016). However, the reference ENF signal extracted from the power grid directly is hard to be satisfied in most of the time. As a result, the mentioned methods have serious limitation in practice.

It is natural to consider using the continuity of the ENF signal to detect video forgery. However, it is very difficult to use the continuity of ENF signal extracted from video since the sampling rate of data points is not high enough. The frame rate of the video is always not more than 30fps, which leads to the lack of the data points. In order to solve this problem, we analyze the ENF signal from video and employ a special method for ENF signal interpolation. As a result, we can use the reconstructed ENF signal to detect forgery without relying on a reference ENF database.

The proposed method focuses on detecting the inter-frame video forgery, which tampers the whole frame instead of the region in the frame. Frame deletion, duplication and insertion are three of the most common used inter-frame video forgery methods, so this paper mainly investigates the detection algorithm to these three forgery methods. The main contribution of this paper is an ENF based inter-frame video forgery detection method which does not need the reference ENF signal and can be used to detect the accurate forgery position in the surveillance video with static scene. The proposed method is much more practical than other existing ENF based video forgery detection methods in practice.

The rest of this paper is organized as follows. Section 2 describes the principle and implementation of our algorithm. Section 3 discusses the practical problems in the algorithm. Section 4 analyses the experimental results. Section 5 concludes the paper.

2. THE PROPOSED ALGORITHM

2.1 Reconstruction of ENF Signal

The main source of the ENF signal in video is the flicker of the lighting. In one period of the ENF signal, the voltage amplitude will reach its maximum value twice, which makes the frequency of the flicker twice the power grid frequency. As mentioned above, the nominal value of the ENF frequency is commonly 50 or 60 Hz, so the flicker frequency will be 100 Hz or 120 Hz. The flicker cannot be noticed by human because of its frequency, but it could be recorded in the videos taken under lighting.

Each frame in the video is a sample of the flicker signal. The sampling rate (the same as the video frame rate) is usually no more than 30Hz, which is lower than the flicker frequency. Fortunately, both the ENF signal and the flicker signal are narrowband. According to the sampling theorem, the sampling rate fs can be used to capture all the information from a narrowband signal if it satisfies the condition below:

$$f_s \geq 2B\left(1+\frac{\frac{f_H}{B}-\left\lfloor\frac{f_H}{B}\right\rfloor}{\left\lfloor\frac{f_H}{B}\right\rfloor}\right) \qquad (1)$$