Chapter 11
Fast and Effective Copy-Move Detection of Digital Audio Based on Auto Segment

Xinchao Huang
School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China

Zihan Liu
School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China

Wei Lu
School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China

Hongmei Liu
School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China

Shijun Xiang
College of Information Science and Technology, Jinan University, Guangzhou, China

ABSTRACT

Detecting digital audio forgeries is a significant research focus in the field of audio forensics. In this article, the authors focus on a special form of digital audio forgery—copy-move—and propose a fast and effective method to detect doctored audios. First, the article segments the input audio data into syllables by voice activity detection and syllable detection. Second, the authors select the points in the frequency domain as feature by applying discrete Fourier transform (DFT) to each audio segment. Furthermore, this article sorts every segment according to the features and gets a sorted list of audio segments. In the end, the article merely compares one segment with some adjacent segments in the sorted list so that the time complexity is decreased. After comparisons with other state of the art methods, the results show that the proposed method can identify the authentication of the input audio and locate the forged position fast and effectively.

DOI: 10.4018/978-1-7998-3025-2.ch011

1. INTRODUCTION

With the continuous development of science, digital multimedia, especially digital audio, is widely used nowadays. Because digital audio is easy to be transmitted and stored, it makes our daily life more colorful. However, as is well-known that everything is a double-edged sword, digital audio can also cause harm to the society in that it is easy to be edited, or in other words, vulnerable. As a result, the authentication of digital audio is significant since it might play an important role like a piece of crucial evidence in forensics and court. What even worse is that some types of digital audio forgeries such as copy-move forgery are imperceptible, and it's difficult to be detected. Copy-move forgery of digital audio could be done as follows: copy some words from an original audio and paste the words to other positions of the same audio. It can be easily realized by using the audio editing application such as Adobe Audition CC and people can hardly detect the copy-move forgery through ears because of the copied segment derived from the same audio. In addition, some post-processing may be adopted to the copied segment for making the forgery harder to be detected. Therefore, copy-move forgery detection of digital audio has become an urgent issue in the area of audio forensics.

At present, some advanced technologies like digital watermarking and digital signature are used to protect the integrity of digital audio effectively. Such kind of technology is called active forensic technique. Many excellent audio watermarking algorithms (Bassia, Pitas & Nikolaidis, 2001; Wang & Zhao, 2006; Wu, Su & Kuo, 2000; Li, Xue & Lu, 2006; Xiang & Huang, 2007) have been proposed. However, the biggest limitation is that most of recording devices don't have the function to insert watermark or signature into digital audio data now. For this reason, another kind of technology, which is called passive forensic technique, is arousing attention in audio forensics nowadays. Passive forensic technique can just use the audio without adding any digital watermarking or signature for verifying the authenticity and integrity of audio, and our method for copy-move detection of digital audio is based on passive forensic technique.

There are many research achievements in the area of audio forensics. Farid (Farid, 1999) put forward to an assumption that in the frequency domain a natural signal has weak higher-order statistical correlations, and proposed a new scheme that use polyspectral analysis technique to detect the forgery. Cooper (Copper, 2010) analyzed the cross-correlation between the signal and second-order difference, and proposed a method that can detect the "butt-splicing'' in tempered audio. Alessandro (D'Alessandro & Shi, 2009) used frequency spectrum analysis to detect MP3 bit rate quality. Grigoras (Grigoras, 2005) proposed a new method that use ENF (electric network frequency) as feature for verifying the authenticity of audio. Maarten et al. (Huijbregtse & Geradts, 2009) improved Grigoras's method. They found that there are certain requirements about file length in Grigoras's method, only when file length reaches to a certain value could a precise result be gotten. So, Maarten et al. did some pre-processings for audio data at first, then calculated correlation coefficient and made improved algorithm effective for short-time audio files. Kraetzer et al. (Kraetzer, Oermann, Dittmann & Lang, 2007) detected the forgery by classifying the audio using statistical features of digital audio consisting of time domain-based features and mel-cepstral domain-based features, the method detects forgery by checking whether every audio frames were recorded under same circumstance or same equipment. Yang et al. (Yang, Qu & Huang, 2008; Yang, Qu & Huang, 2012) used the inconsistency of frame offset to detect the audio forgery in MP3 files. Chen et al. (Chen, Xiang, Liu & Huang, 2013) analyzed high-order singularity of wavelet coefficients and proposed an audio splicing detection model. Pan et al. (Pan, Zhang & Lyu, 2012) came