Digital Signal Processing 24 (2014) 197–208

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

A novel speech content authentication algorithm based on Bessel–Fourier moments

Zhenghui Liu, Hongxia Wang *

School of Information Science and Technology, Southwest Jiaotong University, Chengdu, Sichuan 610031, China

ARTICLE INFO

Article history:
Available online 19 September 2013

Keywords:
Content authentication
Bessel–Fourier moments
Synchronization codes
Tamper localization
Digital watermarking

ABSTRACT

For audio watermark schemes, the method robust against desynchronization attacks based on synchronization codes faces security challenges. In this paper, a content-based method robust against insertion and deletion attacks is given, which is aimed to solve the insecurity problem of synchronization codes embedding, and a speech content authentication algorithm based on Bessel–Fourier moments is proposed. The definition and fast computation of Bessel–Fourier moments of discrete signal are given, and the attack on synchronization codes embedding method is described. For the scheme proposed, the non-synchronized signals caused by desynchronization attack can be re-synchronized by finding the frame that the watermark generated and extracted are equal. Comparing with the synchronization codes embedding method, the scheme not only is robust against insertion and deletion attacks, but also improves the security of watermark system. Theoretical analysis and experimental evaluation results show that the scheme is effective.

1. Introduction

Currently, for audio signals, there are a lot of research results in protecting audio copyright and authenticating the veracity and integrity of audio content [1–5]. Comparing with audio signals, speech signals are more likely to cause attacker's interest and maliciously attacked. If the attacked speech signals are not detected, the authentication client will consider that the attacked speech signal is veracity, which may cause serious consequences.

For speech signals, there are a lot of research results in the speaker recognition and identification [6–9], while the speech content authentication schemes are rare. In [10], the scheme used to detect speech forgery is proposed. The scheme can detect three kinds of alterations or forgeries, in which the cyclic pattern embedding method is used to overcome synchronizing problems. To some extent, the method increases the load of the watermarked signal. In [11], a watermark scheme for compressed speech based on compression technique and codebook-excited linear prediction is proposed, in which the features used to generate watermark bits are extracted during compression. For speech signals uncompressed or compressed based on other speech codecs (not based on codebook-excited linear prediction), the scheme is powerless. And watermark bits are embedded in the lest significant bits (LSBs), which is very fragile to common signal processing operations. For common signal processing operation, the scheme will regard it as hostile attack. In [12], Yuan and Huss introduced an integrity and authenticity mechanic for real-time multimedia communication and produced a method for real-time speech integrity and authentication incorporating with GSM 610 full-rate coder, which is mainly used in the real-time communication process. In practical application, for the convenience of storage, the format of special requirements and many other reasons, speech signal will inevitably be subjected to a certain degree of common signal processing operations. In this situation, the schemes proposed in [11,12] are unsuitable.

Moments and orthogonal moments have been widely utilized as features for image watermark and image pattern recognition. Kim and Lee [13] introduced a robust image watermark based on the normalized Zernike moments (ZMs) of an image. The watermark is generated by modifying some selected ZMs. Revaud et al. [14] proposed an improved ZMs for 2D/3D object recognition. The improved algorithm is more robust against noise and geometric deformation. Xiao et al. [15] introduced a new set of orthogonalized moments based on the Bessel function of the first kind, and named Bessel–Fourier moments (BFMs). Li et al. [16] proposed a watermarking scheme based on BFMs, and analyzed the robustness against various geometric attacks experimentally.

More recently, moments or orthogonal moments used for audio and speech watermark are rare. Xiang et al. [17] proposed a robust audio watermark scheme based on ZMs. In this paper, authors analyzed the linear relationship between the audio amplitude and its ZMs, and watermark the audio ZMs in lower orders by scaling the sample values. Wang et al. [2] proposed a digital audio watermarking algorithm based on pseudo-Zernike moments (P-ZMs) and synchronization code. With the spatial watermarking technique, synchronization codes are embedded based on signal's energy, and then the watermark bits are embedded into the average value of modulus of the low-order P-ZMs.

* Corresponding author.

1051-2004/$ – see front matter
http://dx.doi.org/10.1016/j.dsp.2013.09.007

For audio watermarking systems, it's known that desynchronization attacks, such as deletion and insertion attack, are the most difficult attacks to resist, because that they desynchronize the location of the watermark and cause incorrect watermark detection. The schemes robust against desynchronization attacks are based on synchronization codes commonly, which are embedded by quantifying signal's energy or the average value [1,2,18]. This introduces two issues. The first is synchronization codes embedded increase the load of the watermarked signal. The second is, for the calculation of energy and the average value is public to attackers, it introduces security challenges. Such as, for the synchronous codes embedding method based on signal energy is public, it is easy for attackers to get the energy of signals, and then extract the synchronous codes embedded. After that, the attackers select other signal randomly, and embed the synchronous codes in it. Then substitute the original signal using the signal that the synchronous codes embedded by the attackers, which will not be detected by the certified authority. That is the traditionally method robust against desynchronization attacks is insecurity, which inspires the content-based method robust against deletion and insertion attacks proposed in this paper.

The number, position and distribution of zeros of the radial polynomials of orthogonal moments correspond to the ability of the polynomials to catch the signals information [15,19,20]. Radial polynomials of the BFMs have more zeros than Zernike polynomials of the same degree, and the BFMs of 2D speech signals (obtained by the 1D to 2D map) are more robust than ZMs and p-ZMs. So, comparing with ZMs and p-ZMs, the authors believe that the BFMs are much more suitable for speech watermark schemes.

Considering the shortcomings of synchronous codes embedding method based on signal's energy used commonly, the application of speech content authentication in real life and the advantages of BFMs used for speech watermark schemes, a novel speech content authentication algorithm based on BFMs is proposed in this paper, and a content-based method robust against insertion and deletion attacks is given. The BFMs of discrete 2D signals and the fast computation method are shown, then the properties of BFMs of discrete 2D signals are analyzed, containing the robustness of BFMs and liner relationship between speech signals and its BFMs. Based on the liner relationship between the magnitudes of BFMs and the speech samples value, watermark bits are embedded by quantizing the amplitudes of the selected BFMs, which is completed by multiplying the speech samples value. The non-synchronized signals caused by desynchronization attacks can be re-synchronized by finding the frame, from which the watermark generated and extracted are equal. Theoretical analysis and experimental evaluation results show that the scheme proposed is inaudible and has excellent ability of tamper detection, and robust against insertion and deletion attacks.

The organization of this paper is as follows. Section 2 introduces the definition and fast computation of Bessel–Fourier moments of discrete signals. In Section 3, the properties of Bessel–Fourier moments are analyzed theoretically and experimentally. Section 4 describes the scheme proposed. Section 5 analyzes the performance of the algorithm theoretically. In Section 6, experiment results and comparison with other reported algorithms are presented, which illustrate the effectiveness of the proposed scheme. Finally, we summarize the conclusion in Section 7.

2. Bessel–Fourier moments

2.1. Bessel–Fourier moments of discrete signals

The definition of the Bessel function of the first kind is in Eq. (1).

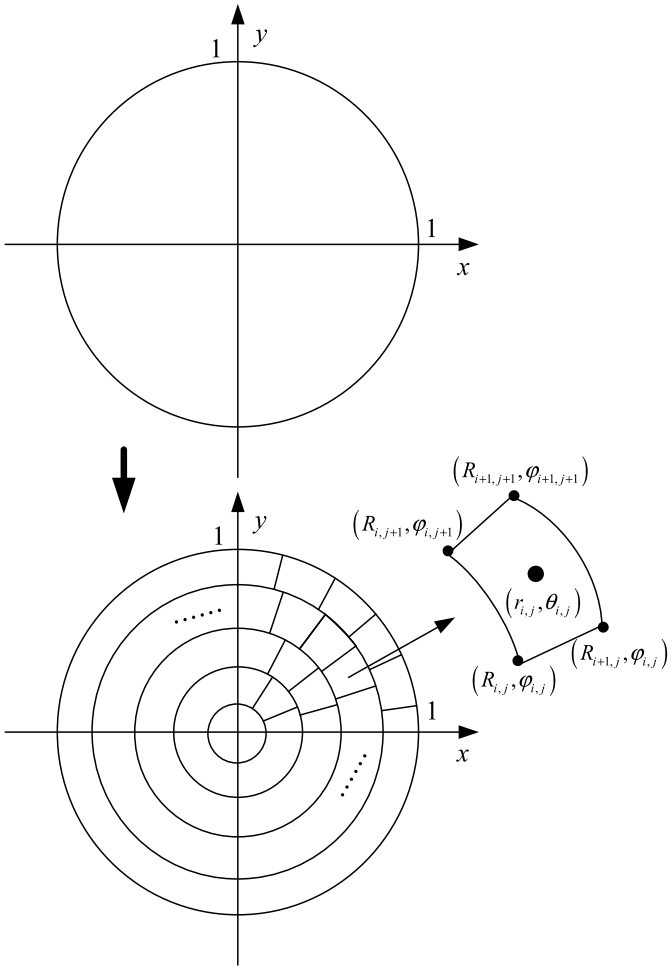

Fig. 1. The segmentation of unit disk.

$$J_v(x)=\sum_{k=0}^{\infty}\frac{(-1)^k}{k!\Gamma(v+k+1)}\left(\frac{x}{2}\right)^{v+2k} \tag{1}$$

where v is a real constant, and $\Gamma(\cdot)$ is the gamma function [15,16].

The Bessel–Fourier moments (BFMs) of continuous function $f(r,\theta)$ in polar coordinates, using Bessel function of the first kind, is defined in Eq. (2).

$$B_{nm}=\frac{1}{2\pi a_n}\int_0^{2\pi}\int_0^{1} f(r,\theta)J_v(\lambda_n r)\exp(-\tilde{i}m\theta)r\,\mathrm{d}r\,\mathrm{d}\theta \tag{2}$$

where $n=0,1,2,\ldots$, $m=0,\pm1,\pm2,\ldots$, $\tilde{i}=\sqrt{-1}$, and $a_n=[J_{v+1}(\lambda_n r)]^2/2$ is the normalization constant, λ_n is the n-th zero of $J_v(r)$. B_{nm} is the BFMs of order n with repetition m.

It should be pointed out that, the BFMs of discrete 2D signals can't be calculated by using Eq. (2) directly. The computation of BFMs of discrete 2D signals in polar coordinates is based on the segmentation idea. The unit disk is split into a set of non-overlapped circular sectors, and each sector is represented by one point in its center, as shown in Fig. 1. Based on the segmentation method, Eq. (2) can be rewritten as

$$B_{nm}=\frac{1}{2\pi a_n}\lim_{\substack{I\to\infty\\ J\to\infty}}\sum_{i=1}^{I}\sum_{j=1}^{J} f(r_{i,j},\theta_{i,j})$$