Int. J. Electron. Commun. (AEÜ) 70 (2016) 416–426

Contents lists available at ScienceDirect

International Journal of Electronics and Communications (AEÜ)

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

A blind robust digital watermarking using invariant exponent moments☆

X.Y. Wang*, Q.L. Shi, S.M. Wang, H.Y. Yang*

School of Computer and Information Technology, Liaoning Normal University, Dalian 116029, PR China

ARTICLE INFO

Article history:
Received 9 June 2015
Accepted 5 January 2016

Keywords:
Audio watermarking
Robustness
Exponent moments (EMs)
Quantization index modulation (QIM)
Synchronization code

ABSTRACT

It is a challenging work to design a robust audio watermarking scheme against various attacks. Moments and moment invariants have become a powerful tool in robust audio watermarking owing to their description capability and invariance property. Exponent moments (EMs) is a new kind of orthogonal moment defined on the circular domain. EMs has a better reconstruction, lower noise sensitivity, lower computational complexity, and excellent invariant property. Based on invariant exponent moments and synchronization code technique, a new and robust audio watermarking method with good auditory quality and reasonable resistance against most attacks is proposed in this paper. Specifically, the watermark data is efficiently inserted in host audio by taking advantage of both invariant EMs and quantization index modulation (QIM). EMs magnitude is robust to common signal processing operations, and QIM can render watermarking scheme blind in nature. To resist against desynchronization attack, the synchronization code technique is also integrated into our invariant EMs based audio watermarking method. Compared with the typical and related audio watermarking schemes, the proposed method is more robust to most attacks and has higher imperceptibility. The effectiveness of our method is demonstrated by simulation results.

1. Introduction

FACING the ever-growing quantity of digital documents transmitted over the internet, it is more than ever necessary for efficient and practical data hiding techniques to be designed in order to protect intellectual property rights [1]. Digital watermarking techniques have historically been used to ensure security in terms of ownership protection and tamper proofing for a wide variety of data formats. This includes images, audio, video, natural language processing software, relational databases, and more, this paper focuses on audio watermarking [2]. Generally, digital audio watermarking is the technology of embedding a useful data (watermark data) within a host audio and the perceptual quality of the host audio should not be degraded substantially by the embedding. For different purposes, audio watermarking can be branched into two classifications: robust audio watermarking and fragile audio watermarking. Robust audio watermarking is used to protect ownership of the digital audio. In contrast, the purpose of fragile audio watermarking is digital audio authentication, that is, to ensure the integrity of the digital audio [3,4].

In the past decade, researchers have made the great efforts in developing robust audio watermarking algorithms, which were implemented in either the time domain [5,6] or transform domains such as the discrete Fourier transform (DFT) [7], discrete cosine transform (DCT) [8], discrete wavelet transform (DWT) [9], cepstrum [10], singular value decomposition (SVD) [11], etc. Transform-domain audio watermarking techniques can always provide higher audio quality and much more robustness than audio watermarking based on time domain, this is because that they can fully take advantage of signal characteristics and auditory properties. Aside from the employment of quantization index modulation (QIM), Wang et al. [12] explored the multiresolution analysis of DWT and the energy-compression characteristics of DCT to achieve efficient audio watermarking. Singh et al. [13] proposed a robust audio watermarking scheme for MPEG-1/Audio Layer II compressed domain. The scheme is implemented by modifying the subband coefficients using adaptive QIM. The watermarking procedure exploits perceptual frequency and temporal masking of the human auditory system (HAS) of MPEG coder to satisfy the requirements of robustness, security and transparency. In [14], the host audio is divided into two subsegments and the DCT coefficients

☆ This work was supported by the National Natural Science Foundation of China under Grant Nos. 61472171 & 61272416, and Liaoning Research Project for Institutions of Higher Education of China under Grant No. L2013407.

* Corresponding authors.
E-mail addresses: wxy37@126.com (X.Y. Wang), yhy_65@126.com (H.Y. Yang).

http://dx.doi.org/10.1016/j.aeue.2016.01.002
1434-8411/

of the subsegments are computed. The DCT coefficients related to a specified frequency region are then partitioned into a number of frame pairs. The DCT frame pairs suitable for watermark embedding are chosen by a selection criterion and watermarks are embedded into the selected DCT frame pairs by modifying their coefficients, controlled by a secret key. Wang et al. [15] presented a blind audio watermarking algorithm based on the vector norm and the logarithmic quantization index modulation (LQIM) in the DWT domain, integrating the robustness of the vector norm with the imperceptibility of the LQIM based on μ-Law (or mu-Law) companding. Hu et al. [16] presented a discrete wavelet packet transform (DWPT) and the DCT framework for blind audio watermarking. Via the exploitation of auditory properties, the quantization steps for QIM are not only perceptually determinable during watermark embedding but also retrievable during watermark extraction. The imperceptibility of the embedded watermarks is ensured because the disturbance caused by the QIM remains below the auditory masking threshold. Wang et al. [17] proposed an adaptive DWT-DCT audio watermarking scheme using the well-trained support vector regression (SVR). This adaptive audio watermarking approach uses the corresponding feature of the template in the training sample to achieve favorable tradeoffs between robustness and imperceptibility. Lei et al. [18] introduced a new audio watermarking based on self-adaptive particle swarm optimization (SAPSO) and quaternion wavelet transform (QWT). By obtaining optimal watermark strength using a uniquely designed objective function, SAPSO addresses the conflicting problem of robustness, imperceptibility, and capacity of audio watermarking using self-adjusted parameters. Khaldi et al. [19] proposed an adaptive audio watermarking scheme based on the empirical mode decomposition (EMD). Digital watermark is embedded in very low frequency mode (last IMF), thus achieving good performance against various attacks. Watermark is associated with synchronization codes and thus the synchronized watermark has the ability to resist shifting and cropping. Combining the robustness of vector norm with that of the approximation components after the DWT, Wang et al. [20] presented a blind and adaptive audio watermarking algorithm. In order to improve the robustness and imperceptibility, a binary image encrypted by Arnold transform as watermark is embedded in the vector norm of the segmented approximation components, the count of which depends on the size of the watermark, after DWT of the original audio signal through QIM with an adaptive quantization step selection scheme. Bhat et al. [21] presented a SVD-based blind watermarking scheme operated in the DWT domain. The watermark bits were embedded into the audio signals using QIM, of which the quantization steps were adaptively determined according to the statistical properties of the involved DWT coefficients. Lei et al. [22] integrated lifting wavelet transform (LWT), SVD, and QIM to achieve a very good tradeoff among the robustness, imperceptibility, and payload. Hu et al. [23] proposed a novel approach for blind audio watermarking. The proposed scheme utilizes the flexibility of DWPT to approximate the critical bands and adaptively determines suitable embedding strengths for carrying out QIM. The SVD is employed to analyze the matrix formed by the DWPT coefficients and embed watermark bits by manipulating singular values subject to perceptual criteria. Lei et al. [24] proposed a robust audio watermarking scheme based on LWT-DCT-SVD, DWT-DCT-SVD with exploration of differential evolution (DE) optimization and dither modulation (DM) quantization. The attractive properties of SVD, LWT/DWT-DCT, DE and quantization technique make the scheme very robust to various common signal processing attacks. Mohsenfar et al. [25] proposed an intelligent audio watermarking method in terms of collaborating QR decomposition and Genetic Algorithm (GA). At the outset, the host signal is segmented into several frames. Then, every frame is decomposed by using QR decomposition method, and, subsequently, the best place for embedding the watermark bit which has a high robustness to the possible attacks is searched by using GA.

In recent years, moment and moment invariants have found extensive applications in the field of image watermarking [26]. The reason for this is that, moments are efficient content descriptors, which have the advantage to fully reconstruct the initial host after the embedding of watermark information. Moreover, the invariant properties of the moments to remain unchanged under common geometric transformations significantly increase the robustness of the watermarks in such kind of attacks [27,28]. More recently, moments and moment invariants are introduced into robust audio watermarking owing to their description capability and geometric invariance property, and some moment-based robust audio watermarking schemes have been proposed. Xiang et al. [29] proposed a robust audio watermarking scheme based on Zernike moments (ZMs). By analyzing and deducting the linear relationship between the audio amplitude and ZMs, watermarking the low-order ZMs is achieved in time domain by scaling sample values directly. Thus, the degradation in audio reconstruction from a limited number of watermarked ZMs is avoided. Based on pseudo-Zernike moments (PZMs), Liu et al. [30] proposed an audio content authentication algorithm robust against feature-analyzed substitution attack, which is aimed at some insecure issues in the existing content-based audio content authentication schemes. Here, the features used to generate and extract watermark are unknown to attackers. So, the scheme proposed is robust against feature-analyzed substitution attack. Wang et al. [31] proposed a digital audio watermarking algorithm based on wavelet moment 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 wavelet moment. Liu et al. [32] presented a speech content authentication algorithm based on Bessel–Fourier moments. 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. One problem of the above moments based audio watermarking method is that, the imperceptibility and robustness of the embedded watermark is not well guaranteed, this is because that the kernel computation of above moments involves computation of a number of factorial terms, which inevitably cause the numerical instability of these moments. For example, the numerical stability of ZMs/PZMs will break down when the number of moments is increased up to a certain value, approximately 44 for ZMs while 23 for PZMs [33].

According to the relation between the exponential function and triangular function, Ping et al. [34] extended radial harmonic Fourier moments and introduced a new moment named exponent moments (EMs). Unlike the conventional moments such as ZMs/PZMs, the kernel computation of EMs is extremely simple and has no numerical stability issue whatsoever. This implies that EMs encompass the orthogonality and invariance advantages of ZMs/PZMs, but are free from their inherent limitations. As a result, we believe EMs magnitude are more suitable for robust watermarking. Based on invariant EMs and synchronization code technique, we proposed a new audio watermarking method with good auditory quality and reasonable resistance against most attacks. Digital watermark is inserted in host audio by taking advantage of both invariant EMs and QIM, and synchronization code is also integrated into the audio watermarking to resist against desynchronization attack.

The rest of this paper is organized as follows. Section 2 presents the basic principle of the proposed audio watermarking scheme. In