Int. J. Electron. Commun. (AEÜ) 69 (2015) 188–197

Contents lists available at ScienceDirect

International Journal of Electronics and Communications (AEÜ)

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

Perception-based audio watermarking scheme in the compressed bitstream

Baiying Lei [a,*], Ing Yann Soon [b]

[a] Department of Biomedical Engineering, School of Medicine, Shenzhen University, National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, Shenzhen 518060, China
[b] School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore

ARTICLE INFO

Article history:
Received 25 January 2012
Accepted 23 August 2014

Keywords:
Audio watermarking
MPEG-4 SLS audio
Perception
Copyright protection

ABSTRACT

In this paper, a new perception-based watermarking scheme for MPEG-4 scalable to lossless (SLS) audio is proposed. Under the control of the psychoacoustic model, the significant part of integer modified discrete cosine transform (IntMDCT) coefficients are adaptively modified during MPEG-4 SLS audio compression taking robustness, imperceptibility and security into consideration. The chaotic watermark generated by chaos is simply embedded to ensure security. Moreover, the adaptive spread spectrum method is exploited to further tradeoff robustness and transparency of this scheme. Extensive experimental results confirm that the proposed scheme is robust against common signal processing attacks while the inaudibility of the scheme is preserved.

© 2014 Elsevier GmbH. All rights reserved.

1. Introduction

In the last few years, with the development of Internet and availability of the state-of-the-art digital multimedia, it is possible to share a large amount of audio or video files, but it also runs the risks of widespread multimedia production forgeries and copyright violation. Consequently, the problem of digital property and copyright protection of multimedia products has attracted a lot of interest and is gaining more attention. As a solution to copyright protection issues, digital audio watermarking technology has become a focus in information security and achieved unprecedented development.

Meanwhile, with the advancement in the audio standards, audio techniques and features can be exploited to protect copyright with watermarking method. Therefore, multitudes of watermarking schemes are available for the protection of the audio formats like MP3, MPEG-2 and MPEG-4 advanced audio coding (AAC) to prevent copyright pirating and malicious modifications. However, as the newest standard released in 2006 by ISO/IEC, MPEG-4 scalable lossless audio [1] coding integrates the functions of lossless, perceptual and fine granular scalable audio coding in a single framework and allows the scaling up of a perceptually coded representation such as MPEG-4 AAC, there is no report on watermarking schemes for MPEG-4 SLS audio watermarking schemes in the literature to our best of knowledge. The widespread use of the compressed MPEG-4 SLS audio format on the Internet, the commercial importance of copyright protection, and the potential profit loss caused by illegally copying necessitate a proposal for the MPEG-4 SLS audio copyright protection. Actually, embedding watermarks for MPEG-4 SLS audio in the compressed domain based on IntMDCT is very feasible.

Recently, some work has shown that the chaotic maps can be adopted in digital watermarking to increase the security [2–7]. In [8], due to the unique features of chaos, chaotic sequences have shown superior robustness when compared to the widely used pseudo noise (PN) sequences in watermarking applications. It is evident that watermark embedding process can easily integrate the existing coding schemes with chaotic watermark generation. Thus chaotic sequences are adopted to enhance the security, anti-counterfeit and noninvertibility of the proposed watermarking scheme.

The goal of this paper is to design and analyze a robust, feasible and applicable watermarking scheme based on chaos and human auditory system (HAS) model to protect MPEG-4 SLS audio clip in the IntMDCT domain. Finding a balance point between robustness and transparency will be an important problem adequately solved and addressed in this paper. The watermark should be embedded into the MPEG-4 SLS audio compressed bitstream and extracted directly. The chaotic watermark data as copyright information is embedded into the IntMDCT coefficients and detected from the coded bitstream successfully. The modifications should be as small as possible to ensure the watermark inaudibility. Moreover, the watermark should also be robust to various attacks like

* Corresponding author. Tel.: +86 755 26534314; fax: +86 755 26534940.
E-mail addresses: leiby@szu.edu.cn, baiying@email.unc.edu (B. Lei).

http://dx.doi.org/10.1016/j.aeue.2014.08.012
1434-8411/© 2014 Elsevier GmbH. All rights reserved.

signal processing, and decoding/recoding attacks. Lastly, the proposed method should introduce some possibility of commercial realization and be tailored for a wide range of applications with the exploitation of MEG-4 SLS audio features and encoding and decoding process. All in all, the main contributions of the proposed watermarking framework are threefold. Firstly, it is compatible with the SLS bitstream without much audio quality distortion and robust to MPEG-4 SLS audio compression and coding process. Secondly, it is based on psychoacoustic model and significant state to ensure the imperceptibility. There is no additional psychoacoustic model as we use the SLS perceptual model directly. Chaotic sequence rather than PN sequence is adopted to improve the security. Our method is simple and convenient to implement. Thirdly, it is applied in the IntMDCT domain and adaptive spread spectrum method is used. The watermark can be adaptively modified in a way that tradeoffs between robustness and transparency.

The rest of this paper is organized as follows. A brief review of the recent watermarking schemes is introduced in Section 2. Section 3 describes MPEG-4 SLS structure and related characteristics. Psychoacoustics and the significant state applied in the proposed watermarking scheme are presented in Section 4. Section 5 introduces chaotic sequences in the watermarking field and the proposed watermarking method. The watermark extraction is given in Section 6. Preliminary performance analysis is conducted in Section 7. Experimental results are discussed in Section 8 followed by the paper summary in Section 9.

2. Related work

Currently, digital watermarking techniques mainly focus on the design of watermark generation, embedding and extraction in the frequency, spatial, cepstrum or mixed domain with symmetric and public key. Indeed, irrelevant and perceptual audio characteristics can be exploited to hide the extra signal. The essential principles and techniques are of great significance to design a novel and hybrid scheme. A large number of audio watermarking techniques can be found in the literature. The most widely used watermarking techniques are echo hiding, spread spectrum, patchwork and content based methods. Echo hiding method [9,10] embeds a watermark by adding an echo signal to increase robustness. Echo defined as delay or offset can determine imperceptibility of the modifications. However, it is signal dependent or offset dependent and not suitable for speech signal with frequent silent intervals. Spread spectrum method is the mainstream audio watermarking technique weighted in the time (or frequency) domain by adopting PN or chaotic sequence to generate and spread watermark [11,12]. In this scheme, computational complexity and synchronization are the main challenging problems. Patchwork method is implemented by adding and subtracting a constant value from two corresponding sample sets [13]. It is a typical watermarking method that takes advantage of host signal to tradeoff between robustness and imperceptibility. Increasing the watermark strength within the constraints of masking threshold is a good way to find the balance point in this method. Content-based method explores dynamic and unique audio features for watermarking approaches [14,15]. This scheme is signal dependent or adaptive by modifying audio frames or scale factor. Besides, HAS model is usually adopted in this technique.

At the same time, there are some recent watermarking techniques based on MPEG-2 AAC audio [16–19]. For example, watermarking prototype for MPEG-2 AAC advocated in [16] was a solution to the Pulse Coding Modulation (PCM) watermarking problem. The weak point of the method is that the overall performance of the system is not optimal and the bit error rate is still very high. Besides, the watermark is not a blind detection system. Tachibana et al. [17] proposed a bitstream watermarking scheme for MPEG-2 AAC audio based on a two-dimensional pseudo random array and detect the watermark with the correlation method. It has lower robustness as there is no HAS model. Meanwhile, the algorithm still has much room for improvement because no listening test was performed to verify the imperceptibility. Another novel watermarking scheme for MPEG-compressed audio designed by Quan et al. [18] employed wet paper codes and inserted data directly by modifying the MPEG audio quantization process. The novel enhanced spread spectrum AAC watermarking scheme proposed by Cheng et al. was very fast for real-time application as it used the quantization indices directly [19]. The general scheme applied in the DCT domain is also suitable for other domain and other audio formats. This robust watermarking algorithm with low complexity adopts spectral filtering to reduce noise and improve the detection bitrate.

Apart from MP3 and MPEG-2 AAC data protection framework, the latest high data rate audio watermarking techniques are designed in the latest publications [20,21] as well. In [20], Li et al. introduced three techniques to embed watermark adaptively based on advanced audio zip (AAZ). The masking threshold is used in order to hide a watermark signal. Under the constraints and control of the masking threshold, the watermark is adaptively and transparently embedded. The system is highly dependent on the threshold that makes its computation intensive for real-time application. A similar data hiding scheme for audio property protection based on IntMDCT was suggested in [21] for high data rate audio such as MPEG-4 AAC. The data hiding technique using the IntMDCT time–frequency transforms is still possible and applicable for other compressed audio data and spatial sound information. The disadvantage of the hiding technique is that the perceptual model is too simple to achieve higher data rate with less audio distortion. In other words, higher data rate has higher audible degradation that is not suitable for real application.

Furthermore, there are some watermarking schemes based on chaos and HAS in the compressed domain in the literature too. For instance, a novel watermarking method introduced in [2] used the chaotic sequence and MPEG-1 psychoacoustic model to improve robustness in the frequency domain. Besides, this scheme has relatively high robustness, flexibility, and supports multiple watermarking. However, listening test was not conducted to test the imperceptibility. There is also no performance comparison with other methods. Bassia et al. [22] developed an algorithm that modified each audio sample by adding a signal-dependent, low-pass shaped watermark signal. This amplitude modification watermarking scheme generates watermark by thresholding a chaotic map to form a bipolar sequence. It is a non-blind scheme and can resist most of attacks. However, it is not robust to time scale modification and synchronization attacks.

Finally, a description of perceptual audio watermarking schemes can be found in many other publications too [15,23,24]. In 1996, Boney et al. [23] put forward an algorithm to utilize the MPEG psychoacoustic model in order to obtain necessary frequency-masking values to achieve audio transparency. The watermark is generated by filtering a PN sequence with a filter that approximates HAS frequency masking characteristics. Thus these watermarks are signal dependent and different for other audio signals. This technique was further explored by Swanson et al. [15] which took advantage of perceptual coding techniques in order to embed the watermark efficiently and appropriately.

3. MPEG-4 SLS structure

MPEG-4 SLS codec adopts the IntMDCT based lossless coding approach [25,26]. The input integer PCM format is losslessly

Download English Version:

https://daneshyari.com/en/article/444944

Download Persian Version:

https://daneshyari.com/article/444944

Daneshyari.com