Content-Based Digital Watermarking for Compressed Audio

***Changsheng Xu , *Jiankang Wu & **David Dagan Feng**
*Kent Ridge Digital Labs
21 Heng Mui Keng Terrace
Singapore 119613
{xucs, jiangkang}@krdl.org.sg
**Department of Computer Science
The University of Sydney
NSW, 2006, Australia
feng@cs.usyd.edu.au

Abstract

This paper proposes a method to embed and extract the digital watermark into and from digital compressed audio. The watermark is embedded in partially uncompressed domain and the embedding scheme is high related to audio content. The watermark content contains owner and user identifications and the watermark embedding and detection can be done very fast to ensure on-line transactions and distributions. The experimental results illustrate that the embedded watermark not only does not affect the audio quality in audibility as well as change the bit rates in compressed domain, but also can survive common signal processing methods such as D/A and A/D conversions, adding noise, filtering, re-sampling, and especially the decoding and re-encoding process. The proposed method is very useful and effective for copyright protection, trace of illegal distributions and other applications.

1. Introduction

Today as the development of Internet technology, audio coding technique and digital signal processing techniques, digital compressed audio distribution through the Internet gets faster and more convenient. Compression algorithms for digital audio can preserve audio quality as well as reduce bit rate dramatically, increase network bandwidth, and save density storage of audio content. Among various kinds of compressed digital audio currently used, MP3 is the most popular one and gets more and more welcomed by music users. MP3 audio compression is based on psycho-acoustic models of human auditory system (HAS). It is an ideal format for distributing high-quality sound files online because it can offer near-CD quality at the compression ratio of 11 to 1 (128kb/s).

However, the open environment of Internet causes a problem of illegal distribution of privately owned digital audio and other multimedia products. To prevent digital media from illegal distribution, there is a demand for the copyright protection and trace of illegal distribution sources. Digital watermarking is one of the emerging technologies to solve these problems. It directly embeds the copyright information and user identification into the original audio and keeps the information present in the audio after all kinds of manipulations. Generally, a watermark inside the audio should be inaudible and robust to different kinds of attacks and collusion. Watermark detection must unambiguously identify the ownership and find the illegal distribution sources.

Currently digital audio watermarking techniques mainly focus on uncompressed audio. The methods can be classified into time domain based techniques (Pitas, 1996; Wolfgang & Delp 1996), frequency domain based techniques (Cox *et al*, 1995; Swanson *et al*, 1996), and time-frequency domain based techniques (Swanson *et al*, 1998). Some of these watermarking techniques can survive compression-decompression-recompression processing. Therefore, one possible method to protect compressed audio is to decompress it first, then embed watermark into decompressed audio, and finally recompress the watermarked decompressed audio. This can probably ensure the robustness of the watermark, but it is too time-consuming because the compression process will take a long time. For example, it will take more than 30 minutes to compress a five to six minute audio of WAV format to MP3 format with the bit rate of 128k/sec. So it is not suitable for on-line transaction and distribution. In order to improve

the embedding speed as well as maintain the robustness of watermark, fast and robust embedding schemes for compressed audio must be taken into consideration. But according to our searching, there are so fewer prior watermarking methods related to compressed audio. In (Sandford *et al*, 1997), the auxiliary information is embedded as a watermark into the host signal created by a lossy compression technique. Obviously, this method has low robustness since the watermark can be removed easily without affecting the quality of the host audio signal by decompress the compressed audio. In (Petitcolas, 1999), a watermarking method (MP3Stego) for MP3 files is proposed. MP3Stego hides information in MP3 files during the compression process. The watermark data is first compressed, encrypted and then hidden in the MP3 bit stream. The hiding process takes place at the heart of the Layer III encoding process namely in the inner_loop. The inner loop quantizes the input data and increases the quantizer step size until the quantized data can be coded with the available number of bits. Another loop checks that the distortions introduced by the quantization do not exceed the threshold defined by the psychoacoustic model. The part2_3_length variable contains the number of main_data bits used for scalefactors and Huffman code data in the MP3 bit stream. The bits were encoded by changing the end loop condition of the inner loop. Only randomly chosen part2_3_length values were modified and the selection was done by using a pseudo random bit generator based on SHA-1. This scheme is very weak in robustness. The author acknowledged that any attacker could remove the hidden watermark information by uncompressing the bit stream and recompressing it. On the other hand, MP3Stego does not directly embed watermark in compressed domain. The processed object is PCM audio and the watermark is embedded during the compress process, so it is time-consuming.

This paper provides an effective method to protect copyright and trace illegal distributions for digital compressed audio by embedding digital watermark in partially uncompressed domain. The watermarked audio is robust to various kinds of manipulations and attacks. In the meantime, the embedded information will not affect the audio quality in audibility. The watermark embedding and detection can be done very fast. The detected watermark information can provide proofs of copyright and distribution sources. For copyright protection, the watermark content must contain the owner identification information which is identical in each audio content. For tracing illegal distributions, the watermark content must contain the user identification which is different for each audio transaction. In order to balance the optimality between the audibility and robustness, a content-adaptive embedding method based on human auditory system is proposed. By use of this method, the watermark is high related to the audio content and it tightly follows the masking threshold of the human auditory system. Watermark embedding increases the data rate very little so that it will not cause perceptible distortion in audibility. Any attempt to remove or distort it, including re-encoding the audio content, will lead to perceptible distortion of the original audio content. Since the watermark is embedded in partially uncompressed domain, it will make the embedding speed very fast.

2. Watermarking Scheme

2.1 Generic Embedding Scheme

Usually, there are three generic watermark embedding scheme as shown in Figure1, Figure2 and Figure3. Figure 1 illustrates the generic procedure of watermark embedding in uncompressed domain. The ideal case is the plain audio (PCM format) is embedded with watermark before compression. In this scheme, the content of the watermark only includes the copyright information because we can not get any user information before distribution. This scheme can not trace illegal distribution of the audio content. Actually in most cases, a lot of compressed audio contents without watermarks are existing in the music server and other media for on-line distributions. Figure2 and Figure3 illustrate two watermark embedding schemes to embed the watermark into compressed audio. The scheme of Figure2 is to directly embed the watermark in compressed domain. This can make watermark embedding very fast, but the robustness of the watermark is weak. Any decompression-recompression process can easily remove the watermark. In order to improve the robustness of the watermark, Figure3 illustrate another embedding scheme. According to this scheme, the compressed audio is first decompressed, then the watermark is embedded in uncompressed domain, and finally the watermarked

content is recompressed to generate the watermarked compressed audio. This scheme can improve the robustness of watermark, but it is not suitable for on-line distribution because the compression process is time-consuming. To protect the copyright of these contents and trace illegal distributions as well as ensure on-line transactions, we proposed a novel content-based embedding scheme in this paper. Our scheme fully considers the audio coding algorithms and is high related to audio content so that it can get an optimal balance between audio quality and robustness of the embedded watermark and ensure the embedding speed suitable for on-line distribution.

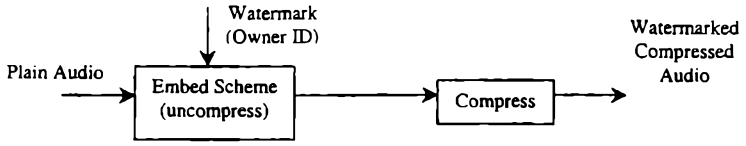

Figure1: Embedding scheme 1

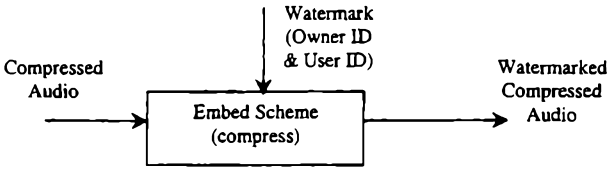

Figure2: Embedding scheme 2

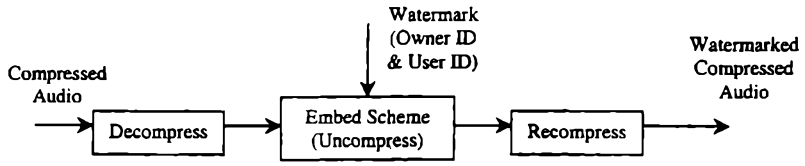

Figure3: Embedding scheme 3

2.2 Content-Based Embedding Scheme

In order to improve the robustness of the watermark embedded into the compressed audio as well as ensure the embedding speed, a content-based watermark embedding scheme is proposed in this section. According to this scheme the watermark will be embedded in partially uncompressed domain and the embedding scheme is high related to audio content. Figure4 illustrates the block diagram of the content-based watermark embedding scheme in partially uncompressed domain.

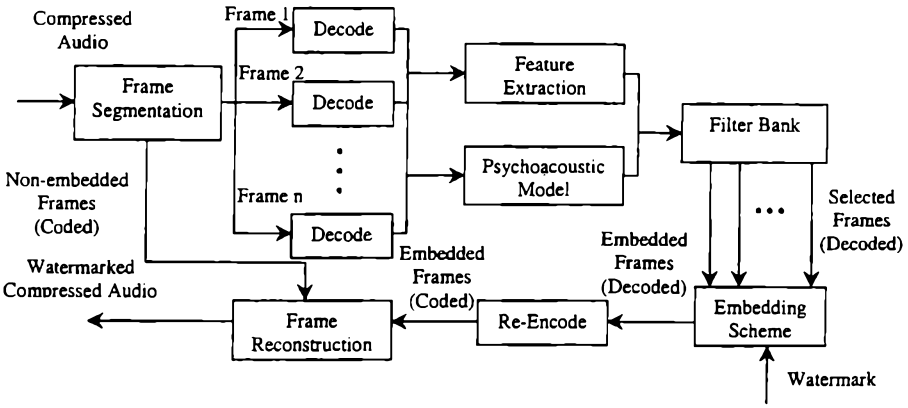

Figure4: Content-Based Watermark Embedding Scheme

The incoming compressed audio is first segmented into frames according to the coding algorithm. All the frames are decoded from compressed domain to uncompressed domain. Then the feature extraction model and the psychoacoustic model are applied to each decoded frame to calculate the features of the audio content and masking threshold in each frame. According to the features and masking threshold, a pre-designed filter bank is used to select the candidate frames suitable for embedding watermark. The watermark will be embedded into these selected frames using an adaptive multiple bit hopping and hiding scheme depicted in Figure5. The embedded frames will be re-encoded to generate the coded frames using the coding algorithm. Finally, The re-encoded frames and the non-embedded frames will be reconstructed to generate the watermarked compressed audio. Compared with the embedded scheme in wholly uncompressed domain, this scheme can not only get the same performance in audibility and robustness but also embed the watermark much faster. It is suitable for on-line embedding and distribution.

Figure5 illustrates the block diagram of detailed watermark embedding scheme for decoded frames from the compressed audio. Since audio coding is a lossy processing, the embedded watermark must exist after audio compression. Furthermore, the embedded watermark must not affect the audio quality perceptually. In order to satisfy these requirements, the embedding scheme fully considers the human auditory system and the features of audio content. For the decoded frames from the original compressed audio which will be selected to embed watermark, feature parameters are extracted from each selected frame to represent the characteristics of the audio content in that frame. In the meantime, each selected frame will pass through a psychoacoustic model to determine the ratio of the signal energy to the masking threshold. Based on the feature parameters and masking threshold, the embedding scheme for each selected frame is designed. The watermark is embedded into these frames using a multiple-bit hopping and hiding method. The watermarked audio frame will be compressed to generate the compressed audio frame.

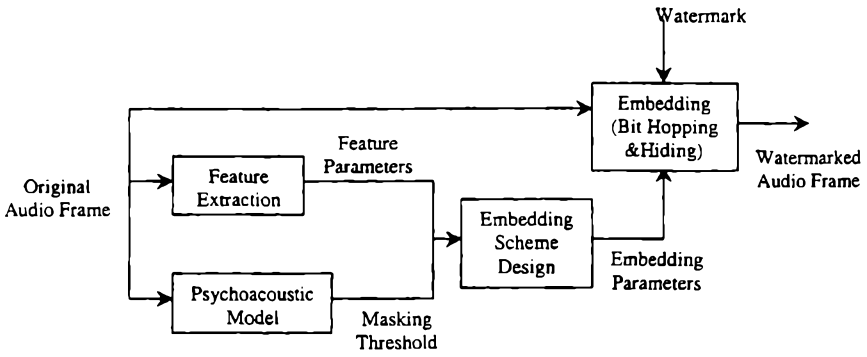

Figure5: Watermark Embedding Scheme for Single Frame

2.3 Extraction Scheme

In order to correctly detected the watermark from a compressed audio, the frames embedded watermark must be extracted at first. Figure6 illustrates how to extracted the frames including watermark from a compressed audio. This process is similar to the watermark embedding scheme to select candidate frames to embed watermark. The watermarked compressed audio is first segmented into frames according to the coding algorithm. These frames are decoded and each decoded frame is analyzed by the feature extraction model and the psychoacoustic model. According to the calculated feature parameters and masking threshold, a filter bank is applied to select the frames including watermark information. The watermark will be detected from these frames using the extraction scheme depicted as Figure7.

Figure7 illustrate the block diagram of watermark extraction from the selected frames. For each incoming frame, we examined the magnitude (at relevant locations in each audio frame) of the autocorrelation of the embedded signal's cepstrum. From the diagram of autocorrelation of the cepstrum, the bits of a watermark in each frame can be found according to a "power spike" at each delay of the embedded bits. Since we use multiple-bit hopping method to embed the bits into the frames, for detected bits in each frame, they will pass through a matched filter bank that can map the bits into the actual code (1 or 0). Finally, the watermark is recovered by correlate the detected codes with the original watermark.

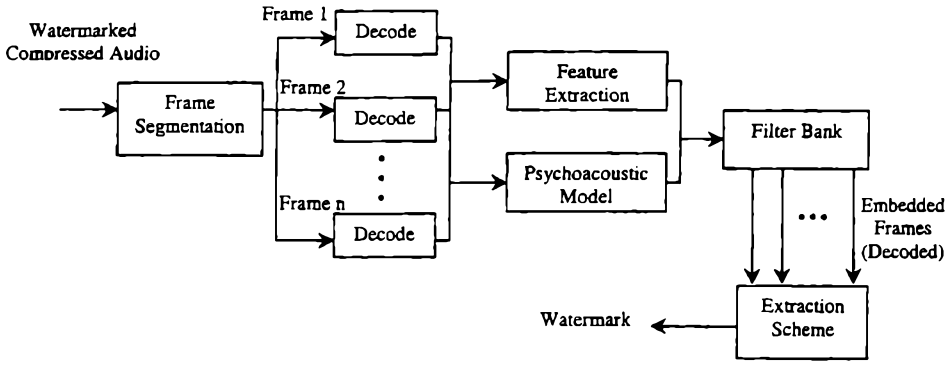

Figure6: Frames and Watermark Extraction Scheme

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.