Muteness-Based Audio Watermarking Technique

Khalid A. Kaabneh and Abdou Youssef

Dept. of Computer Science, The George Washington University
Washington, D.C., 20008, USA.
E-mail: kaabneh@seas.gwu.edu youssef@seas.gwu.edu

Abstract

Audio watermarking is a promising approach to copyright protection of audio data, especially music and songs. Several watermarking techniques have been developed and commercialized. The watermarks produced by those techniques can withstand a number of single attacks such as MPEG, resampling, filtering, and quantization. However, the watermarks are easily destroyed if subjected to chopping or multiple attacks.

In this paper, we present a new audio watermarking technique that provides watermark robustness not only to single attacks but also to multiple attacks and chopping.

1. Introduction

In today's digital world, there is a great wealth of information, which can be accessed in various forms: text, images, audio, and video. It is easy to ensure the copyright of "analog documents" and protect the author ("author" will also be used to denote composer, artist, designer, etc.) from having his/her work stolen or copied by adding a form of owner's signature to the analog documents. In digital data, a watermark, or an invisible stamp, can be used to provide proof of "authorship" of a signal in a court of law. Watermarks can be used also to provide proof of ownership and track illegal copies of the signal.

A watermark must meet several requirements. First, the watermark embedded in data must be imperceptible. Second, the watermark should have robustness against attacks on the signal such as lossy compression, filtering, resampling, noise, chopping, and A/DD/A conversions. Third, the watermark should support multiple watermarks, i.e. multiple users. Fourth, a "pirate" should not be able to detect (identify) the watermark by comparing several signals belonging to the same author. Finally, the signal should be degraded beyond use when the watermark is destroyed through unauthorized means.

The focus of this research is the development and the study of a novel watermarking technique for audio data, especially music and songs. Data hiding in audio signals is especially challenging because (1) data hiding must not be audible since otherwise it will mask the original audio signal and it will be easily tampered with and removed, (2) the human auditory system (HAS) operates over a wide dynamic range between 20 Hz to 20 kHz, making it difficult to embed outside this range, (3) there is a limited area of embedding the data. Therefore, it is a challenging task to manipulate the features of the audio signal without being detected.

In text watermarking [1], a technique was introduced called the open space methods. The methods exploit inter-sentence spacing, end-of-line spaces, and inter-word spacing in justified text, to encode data by inserting spaces at the end of lines. The data are encoded allowing for a predetermined number of spaces at the end of each line.

Our audio watermarking approach bears certain similarities with the open space text watermarking. The audio counterparts of text spaces are periods of silence. In this research we code watermarks and insert them in those periods of silence.

In almost any audio signal there are several mute or silence periods. For example, in a speech clip, the person tends to have normal periods of silence between words or phrases. In almost any song, there are also silence periods. Similarly, listening to piano music quickly reveals many short periods of silence. All those mute or silence periods are considered a vital part of any song or speech; they cannot be neglected or omitted. Omitting those features can cause a major irregularity in the audio clip. By undetectably increasing the length of each mute

0-7695-1080-9/01 $10.00

period, we can have an opening area to embed our watermark.

To that effect, we studied (1) what constitutes effective periods of silence, (2) what can be inserted without distorting the audio quality of a signal, and (3) what makes a watermark that meets the requirements outlined earlier.

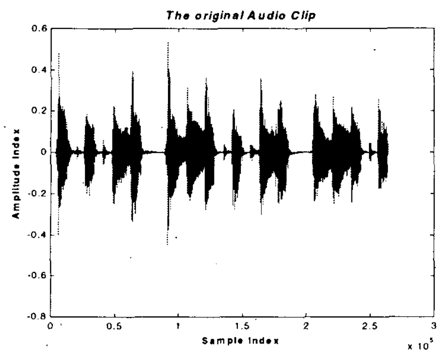

Figure 1. An audio clip containing mute periods.

2. Related Work

A watermarking method for MPEG encoded audio, presented in [2], embeds the watermark directly into the MPEG audio bits streams rather than going through expensive decoding / encoding process in order to apply watermarking schemes in the uncompressed data domain. The main idea of this scheme is to embed the watermark bits into the Scale Factors of the MPEG audio streams.

Another watermarking scheme developed for audio signals works by introducing an echo into a host audio signal [3]. It has been proved that HAS cannot distinguish between the original signal and the echo, if the delay time between the two signals is less than 1msec [4]. The echo hiding technique hides information in sound by introducing an echo characterized by a small delay τ and small amplitude α.

The audio watermarking scheme in the time domain offers copyright protection to an audio signal by modifying its temporal characteristics [5]. The amount of modification embedded is limited by the necessity that the output signal must not be perceptually different from the original one.

3. Our Technique

In any audio signal (music or speech), a mute period offers the following advantages:

- A mute period is an integral part of any audio signal. It cannot be omitted since it represents an integral part of the audio signal.
- It occurs randomly in an audio signal, which is generated by the music process.
- A mute period represents a real time interval that will not be decreased when compressed.

The audio watermarking technique presented in this paper is a muteness-based, which offers the following features:

- It extends the mute periods in an audio signal without any perceptual difference to the average human auditory system.
- The extension of mute periods carries the same amplitude so it will blend with the original signal and will not attract any attention.
- It does not require the original signal to extract the watermark.

The audio watermarking process can be described in the following diagram,

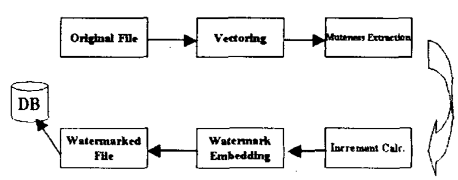

Figure 2. Watermark Process.

4. Watermark Embedding

The watermark embedding scheme proposed in this paper modifies the original audio signal, which is represented as a 16-bit or 8-bit sample sequences at 44.1 kHz stereo. First, we search for all the mute periods in an audio signal that falls within a certain predetermined threshold. The length of any mute period *M(i)* (where i is the index of mute periods in a song) is calculated by the number of audio samples it has. Figure 3 shows the result for a mute periods search in an audio clip.

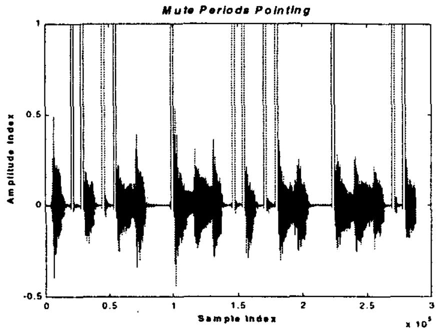

Figure 3. A mute periods search in an audio clip.

Second, we embed (code) the watermark by deciding for each mute period $M(i)$ whether or not to extend the mute period, and by how much (say $\delta(i)$). Now, let $M_o(i)$ represent the original size of the i-th mute period, and $M_w(i)$ the size of the mute period after being extended. Then,

$$M_w(i) = M_o(i) + \delta(i) \qquad (1)$$

5. Watermark Detection

Our method does not require the original signal for its detection. This way the owner of the data does not have to keep double copies of both original and watermarked products. In order to detect the watermark in an audio signal, the owner needs to have only the original length of all mute periods from the audio file $M_o(i)$. From the distributed copy (the watermarked copy), we can extract all the mute periods $M_w(i)$. Using formula (2), and knowing $M_o(i)$, we can recover the values of $\delta(i)$.

$$\delta(i) = M_w(i) - M_o(i) \qquad (2)$$

6. Audio Watermark Robustness to Signal Manipulation.

To test the robustness of the presented watermarking method, we have used a collection of audio signals. The audio files are 16-bit stereo sampled at 44.1 kHz (CD quality).

6.1 Robustness to MP3 audio compression.

We have tested robustness of the watermarking method described above, using MP3 compression. 16-bit signed stereo 44.1 kHz watermarked audio signals were encoded at different kbps rates.

In order to test our recovery rate for the inserted samples, we have used the formula described in echo hiding technique [6]. The formula measures the percentage of recovered (retrieved) number of samples to the inserted ones.

Recovery Rate % = Samples decoded / Samples placed (3)

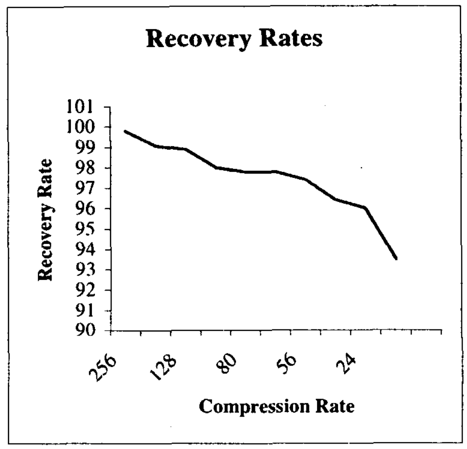

Figure 3. Plot of Recovery Rates for Different Compression Ratios.

Figure 3 above shows the average recovery rate for compressed files at different bit-rates. The test included 256kbps down to 16kbps. We can notice that at a maximum compression rate for an audio files achievable by MP3 we still got 93.5% recovery rate of the inserted samples.

6.2 Robustness to Chopping.

Since the mute periods are randomly distributed throughout the audio signal, it would be very difficult to extract a piece without any mute period. In three other cases the watermark can be detected because the watermark is fully embedded after each mute period.

6.3 Robustness to Low Pass Filtering.

A number of watermarked audio signals sampled at 44.1 kHz were filtered by a 25th order Hamming Low-Pass filter with cut-off frequency 2205Hz. In this experiment, we have 100% success in watermark detection. The reason for this high success is that we are dealing with time domain watermarking technique; so low-pass filtering has no effect on the duration of the audio signal.

6.4 Robustness to Resampling and Requantization.

To test for robustness against resampling, we resampled Watermarked audio signals at 44.1kHz down to 22.05kHz and 11.025kHz, and back again to their initial sampling frequency. Although the above processing caused noticeable distortion in relation to the original signals, the watermarks remained easily detectable, and the recovery rate is about 98.2%.

To test for robustness against quantization, we quantized 16-bit signals down to 8-bits and quantized them back to 16-bit signals. In each case, we tested if the watermark can be recovered. The watermarks resisted the requantization process and we got recovery rate of 94.4 %. The lower recovery rate reflects the miss-representation of samples during requentization process.

<u>*Test Results.*</u>

	Recovery Rate %
LP/Filter	100
Quantization	94.4 *
Resampling	98.4

(*) There was a noticeable distortion to the quantized audio signal.

7. Robustness to Multiple Attacks

Our watermarking technique was tested on a number of multiple attacks, which included MPEG, resampling, quantization, resampling, and chopping. Multiple attacks imply a combination of the above attacks together. The watermarking technique gave back encouraging results implying that this technique can resist a multiple attacks.

8. Conclusions

The watermarking technique presented above embeds a watermark by capitalizing on the vital role of mute periods in an audio signal. The watermarking technique embeds a watermark in the time domain of an audio signal by extending the mute periods. The extraction process does not require the original audio signal. Our audio watermarking technique is perceptually undetectable and resists MPEG compression plus other signal manipulations, such as resampling low- pass filtering, requantization, and chopping. It also resists double attacks. Further work is being conducted to enhance the robustness to more multiple attacks.

9. References

[1] J. Brassil, S. Low, N. Maxemchuk, and L. O'Gorman. Electronic marking and identification techniques to discourage document copying. *Proceedings of Infocom '94*, pages 1278-1287, June 1994.

[2] D. Gruhl, A. Lu, and W. Bender. Echo Hiding. In R. Anderson, editor, *Proceedings of the First International Workshop in Information Hiding*. Lecture Notes in Computer Science, pages 295-315, Cambridge, UK, May 30 – June 1, 1996.

[3] F. Mintzer, "*Opportunities for Watermarking Standards.*" Communications of the ACM, July 1998. Vol. 41, No. 7. p. 57-64.

[4] F. Petitcolas, & Ross J. Anderson. "*On the Limits of Steganography.*" IEEE Journal of Selected Areas in Communications, May 1998.

[5] J. Brassil, S. Low, N. Maxemchuk, and L. O'Gorman. Document Marking and Identification using Both Line and Word Shifting. *Proceedings of Infocom '95*, Boston, Massachusetts, April 1995.

[6] ISO/IEC IS11172-3, "*Coding of moving pictures and associated audio for digital*

storage media up to about 1.5 Mbits/s – Audio part", Nov. 1992.

[7] P. Bassia and I. Pitas. "*Robust Audio Watermarking in the Time Domain*". University of Thessalonika, Thessalonika, Greece, 1999.

[8] L. Qiao and K. Nahrstedt. *Non-Invertible Watermarking Method For MPEG Encoding Audio.* IS&T/SPIE Conference on Security and Watermarking of Multimedia Contents, San Jose, CA, January 1999. Pages 194-202.