A Robust Audio Digital Watermarking Algorithm based on Psychoacoustics Model

Meghali Deokate[1], Sanjay Ganorkar[2]

[1]Sinhgad College of Engineering, Pune University, Pune, India
meghali_one@rediffmail.com

[2]Sinhgad College of Engineering, Pune University, Pune, India
srganorkar.scoe@sinhgad.edu

Abstract: *Digital watermark technology is now drawing attention as a new method of protecting digital content from unauthorized copying. A novel audio watermarking algorithm is used to protect against unauthorized copying of digital audio. Watermarking scheme includes a psychoacoustics model to ensure that the watermarking does not affect the quality of the original sound. The watermarking scheme is robust against common signal processing attacks and it introduces no audible distortion after watermark insertion.*

Keywords: Audio watermarking, Digital watermark, Copyright Protection, Psychoacoustics Model

1. Introduction

The Music industry is looking for reliable solutions to problems associated with the protection and misuse of audio data. Restriction of access to the data is not manageable in applications where the user needs to access the data in order to enjoy it. New methods like digital watermarking, and encryption amongst others have been proposed to meet these needs.

Digital watermark technology is now drawing attention as a new method of protecting digital content from unauthorized copying of digital content. A digital watermark was a signal added to the original digital data, which can later be extracted or detected. The watermark was intended to be permanently embedded into the digital data so that authorized users can easily access it. The watermark should not degrade the quality of the digital data. Some of the issues that must be kept in mind while testing the efficiency of an audio watermarking algorithm are discussed here.

1.1 Inaudibility

The presence of the audio watermark should not compromise with the quality of the given audio signal. The acceptability of the impairments depends on the intended audience. Some people argue that any alteration of the original sound is unacceptable. The inaudibility of the watermark is of most priority for an audio watermarking scheme.

1.2 Robustness

The watermark data should be detectable even after the audio has been subjected to a wide variety of distortions introduced by broadcast, audio compression algorithms and Internet distribution, home recording devices and other manipulations. A robust watermark is one that is able to maintain its integrity and is detectable despite such manipulations.

1.3 Efficiency

The embedding and detection of the watermark data should be achievable at an optimum cost in terms of computation and equipment.

1.4 Security

One of the key issues for any kind of covert technology is Security. The watermarks need to be designed to be resistant to attempts at forgery, alteration, erasure and decoding by unauthorized parties. Different methods can be used to make sure that the algorithm is secure. One of the methods is the use of a "key" while embedding the watermark. To make the algorithm more secure the "key" should be unique so as to protect against proliferation.

Several digital watermark algorithms have been proposed, but most watermark algorithms focus on image and video. Only a few audio watermark algorithms have been reported.

This paper is organized as follows. Section 2 introduces literature survey. Section 3 presents the proposed watermarking scheme, including the embedding and detection process. We show the experimental results of the proposed algorithm in section 4 and conclude the paper in section 5.

2. Literature Survey

Bender et al. [9] proposed several watermarking techniques, which include the following: spread-spectrum coding, which uses a direct sequence spread-spectrum method; echo coding, which employs multiple decaying echoes to place a peak in the cepstrum domain at a known location; and phase coding, which uses phase information as a data space. Unfortunately, these watermarking algorithms cause perceptible signal distortion .Swanson et al. [1] presented an audio watermarking algorithm that exploits temporal and frequency masking by adding a perceptually shaped spread-spectrum sequence. However,

the disadvantage of this scheme is that the original audio signal is needed in the watermark detection process Bassia et al. [3] presented a watermarking scheme in the time domain. They embedded a watermark in the time domain of a digital audio signal by slightly modifying the amplitude of each audio sample. The characteristics of this modification were determined both by the original signal and the copyright owner key. Solana Technology [4] proposed a watermark algorithm based on spread-spectrum coding using a linear predictive coding technique and fast Fourier transform (FFT) to determine the spectral shape. In the watermark detection process, the watermarked audio signal is whitened before correlation using the rake receiver and show low robustness. Furthermore, they have a relative high complexity in the detection process.

In this paper the watermark embedding scheme accomplishes perceptual transparency after watermark embedding by exploiting the masking effect of the human auditory system, as in [7]. This approach can be applied to other signals, such as image and video. For images, this can be accomplished by adopting a human visual model [2]. In the proposed scheme does not need the original audio signal to extract the watermark information.

Linnartz et al. [8] and Depovere et al. used a similar approach in image watermarking. They applied the whitening procedure in watermark detection to remove correlation in pixels using a low pass filter. We achieved this by removing the audio spectrum in the watermarked audio using a linear prediction analysis, which is a frequently used technique in speech signal processing.
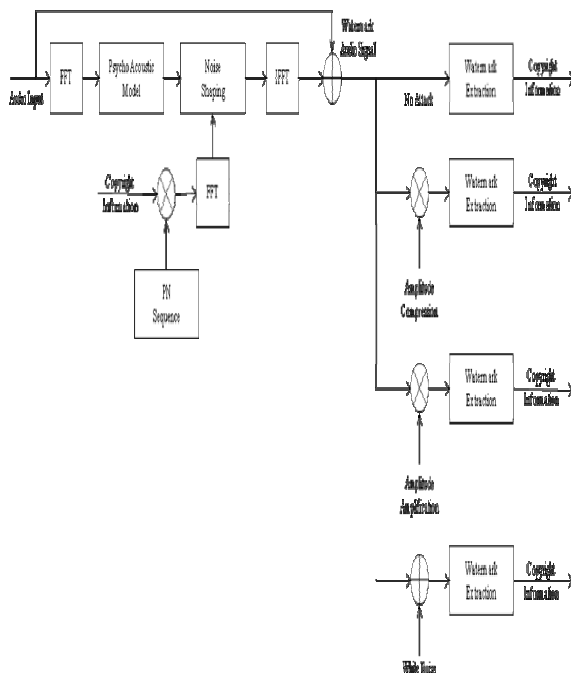
Figure 1: Block diagram of audio watermark embedding and detection procedure

3. Watermarking Scheme

3.1 Watermark Embedding

Our watermark-embedding scheme is based on a direct sequence spread-spectrum (DSSS) method. The auxiliary information that is to be embedded is modulated by a pseudo noise (PN) sequence. The spread-spectrum signal is then shaped in the frequency domain and inserted into the original audio signal. The embedding strength determines the energy of the watermark signal and can be considered as the energy of the noise added to the audio signal. Our goal was to design the weighting function so that the watermark energy was maximized subject to a required maximal acceptable distortion. The strength of the embedded watermark signal depends on the human perceptual characteristics of the audio signal. We accomplished our goal with an embedding scheme that exploits the masking effect of the human auditory system. We intended to adapt the watermark so that the energy of the watermark was maximized and the auditory artifact was kept as low as possible. We used the masking model Audio Psychoacoustics Model. Simultaneous masking is a frequency domain phenomenon where a low-level signal can be made inaudible by a simultaneously occurring stronger signal. The low-level signals below this threshold will not be audible. The filtering is then applied to the PN sequence using the filter coefficients obtained from the all-pole modeling to shape the watermark signal to be imperceptible in the frequency domain. This masking threshold is the limit for maximizing the watermark energy while keeping the perceptual audio quality.

In our embedding process, we repeatedly apply an embedding operation on short segments of the audio signal. Each one of these segments is called a frame. A diagram of the audio watermark-embedding scheme is shown in Figure1.

Using the concept of the Psychoacoustic model and DSSS method, our watermark embedding works as follows:

a) Calculate the masking threshold of the current analysis audio frame using the Psychoacoustic model with an analysis audio frame size samples and FFT.
b) Generate the PN sequence
c) Perform the FFT to the copyright information, which is modulated by the PN sequence.
d) Using the masking threshold, shape the watermark signal to be imperceptible in the frequency domain.
e) Compute the inverse FFT of the shaped watermark signal.
f) Create the final watermarked audio signal by adding the watermark signal to the original audio signal in the time domain.

As shown in figure 2, we have got the hopped signal from the Copyright information (original bit sequence) and PN sequence.
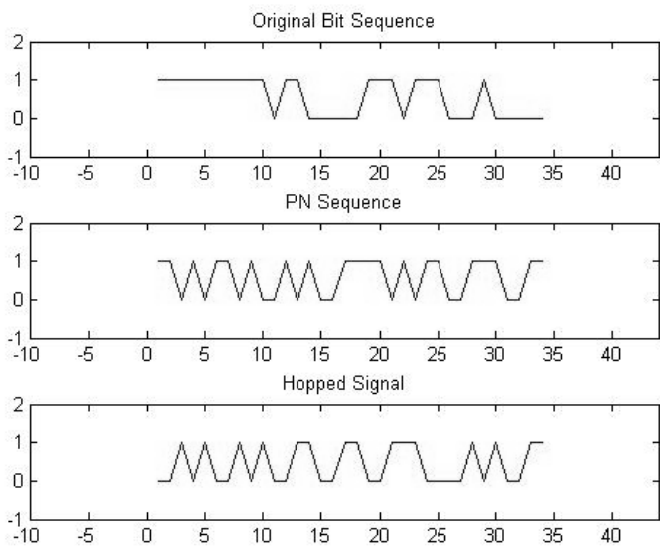
Figure 2: Different signals (Original Bit Sequence, PN Sequence, and Hopped Signal)

Psychoacoustics Model

Psychoacoustics is the study of the perception of sound. Through experimentation, psycho acousticians have established that the human ear presents several limitations. In particular, when two tones, close to each other in frequency, are played simultaneously, frequency masking may, if one of the tones is sufficiently loud, it masks the other one. Psychoacoustic models generalize the frequency-masking effect to non-tonal.
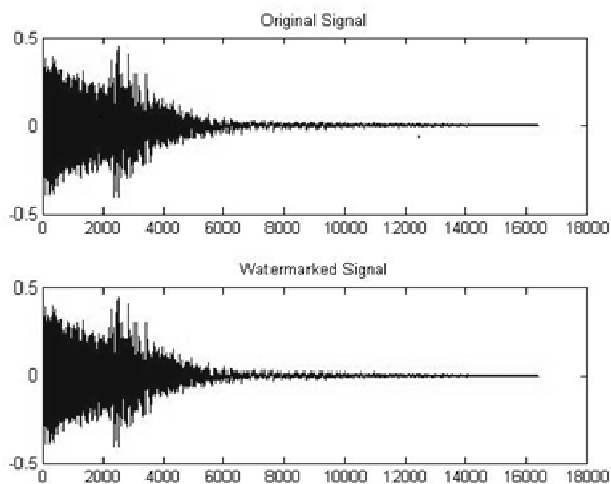

Figure 3: Comparison of Original and watermark audio signal

3.2 Watermark Detection

When designing a watermark detection system, we need to consider the desired performance and robustness of the system. The watermark should be able to be detected under common signal processing operations, such as digital-to-analog and analog-to-digital conversion, linear and nonlinear filtering, compression, and scaling. Furthermore, in most applications, watermark extraction processes do not have to access the original signal. In fact,

not having access to the original signal is essential in real environments such as broadcasting. The watermark detection procedure does not require access to the original audio signal to detect the watermark signal.
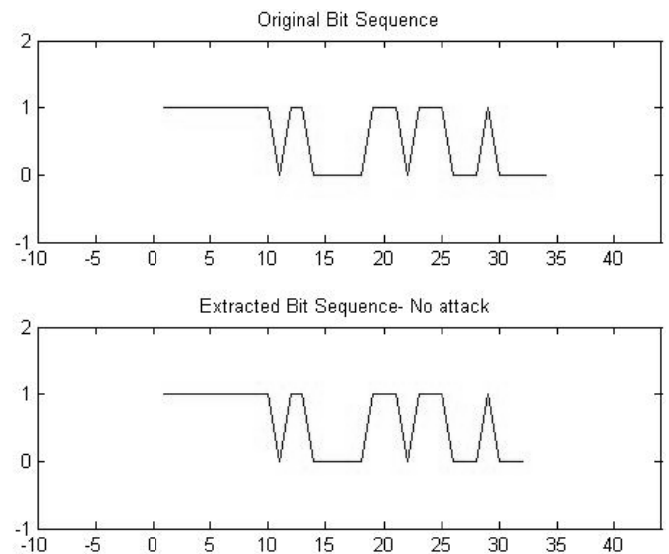
Figure 4: Extracted Original information with no attack

4. Experimental Results

The audio signals used in the experiments were sampled at 44.1 kHz with a 16 bits/sample. Information was embedded into the audio at a rate of 128 bits per 15 seconds.

4.1 Robustness Test

To evaluate the performance of the proposed watermarking algorithm, we tested its robustness. The detailed robustness test procedure is as follows.

4.1.1 Amplitude compression (Deamplication attack)

The amplitude of the audio signal is compressed with a non-linear gain factor. As shown in figure 5, after deamplication attack, we extract the original bit sequence.
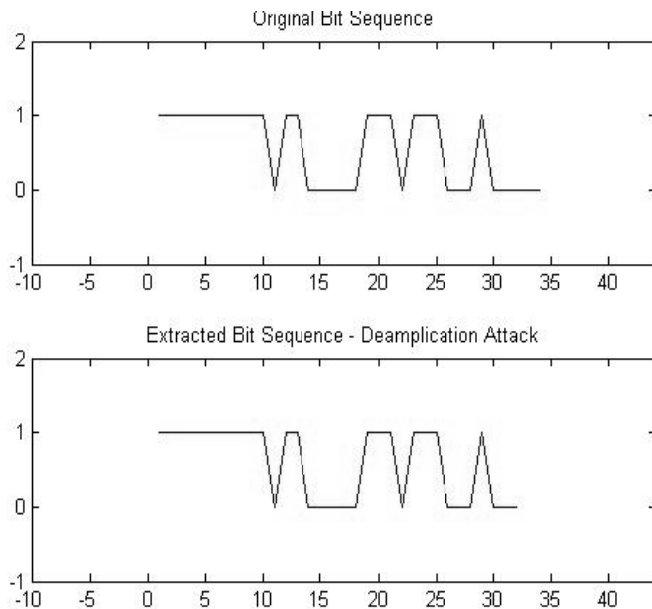
Figure 5: Extracted Original information with amplitude compression (Deamplication) attack
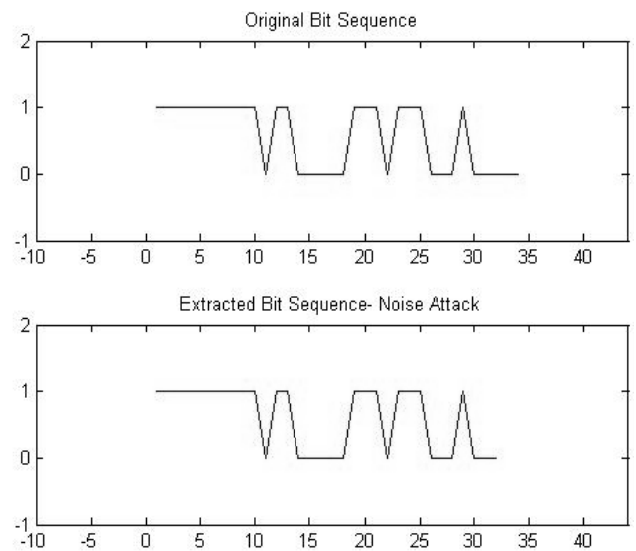
4.1.2 Amplitude Amplification

The amplitude of the audio signal is amplified with a non-linear gain factor.
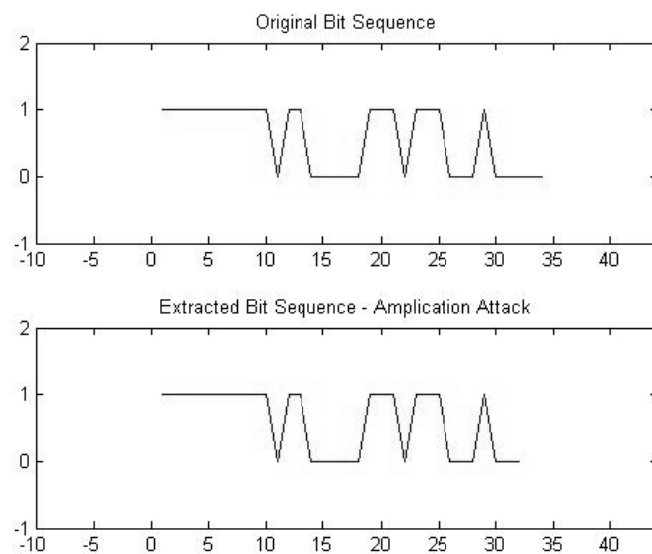
Figure 6: Extracted Original information with amplitude amplification attack

4.1.3 Noise addition

White noise with a constant level of 36 dB was added to the watermarked audio under the averaged power level of the audio signal.

Figure7: Extracted original information with noise attack

Table I shows the robustness test's detection results for the various attacks. We obtained a perfect detection result for the mix down, echo addition, and amplitude compression.

Table 1: Robustness Test: Detection Result for Various Attacks

Type of Attack	Bit Error Rate (%)
Amplitude Compression	0
Amplitude Amplification	0
Noise Attack	0

5. Conclusion

In this paper, we described a new algorithm for digital audio watermarking. The proposed watermark embedding scheme accomplishes perceptual transparency after watermark embedding by exploiting the masking effect of the human auditory system. This embedding scheme adapts the watermark so that the energy of the watermark is maximized under the constraint of keeping the auditory artifact as low as possible.

Our detection procedure extracts copyright information without access to the original signal. Experimental results showed that our watermarking scheme is robust to common signal processing attacks such as amplitude compression, amplitude amplification and noise as shown in figure 5, 6 and 7.

References

[1] M. Swanson, B. Zhu, A. Tewfik, and L. Boney, "Robust Audio Watermarking Using Perceptual Masking," Signal Processing, vol. 66, no. 3, May 1998, pp. 337-355

[2] C. Podilchuk and W. Zeng, "Image-Adaptive Watermarking Using Visual Models," IEEE J. on Selected Areas in Comm., special issue on Copyright and Privacy Protection, vol. 16, no. 4, May 1998, pp. 525-539.

[3] P. Bassia, I. Pitas, and N. Nikolaidis, "Robust Audio Watermarking in the Time Domain," IEEE Trans. On Multimedia, vol. 3, no. 2, June 2001, pp. 232-241.

[4] C. Lee, K. Moallemi, and R. Warren, "Method and Apparatus for Transporting Auxiliary Data in Audio Signals," U.S. Patent 5, 822,360, Oct. 1998.

[5] Yuval Cassuto, Michael Lustig and Shay Mizrachy, Real Time Digital Watermarking System for Audio Signals Using Perceptual Masking, Internal Report, Israel Institute of Technology, 2001.

[6] M Arnold. Audio watermarking: features, applications and algorithms.IEEE International Conference on Multimedia and Expo, New York, USA, 2000, pp:1013-1016.

[7] Tao Zhi, Zhao Heming, Gu Jihua. Digital Audio Watermarking Based on Psychoacoustic Model and the Critical Band Wavelet Transform [J]. Acoustics, 2006,31:114-119

[8] M. Swanson, B. Zhu, A. Tewfik, and L. Boney, "Robust Audio

[9] Watermarking Using Perceptual Masking," Signal Processing, vol.66, no. 3, May 1998, pp. 337-355.

[10] J.P.M.G. Linnartz, A.A.C. Kalker, G. Depovere, "On the Design of a Watermarking System: Considerations a Rationales," Information Hiding Workshop, Preliminary Proc. and Springer Verlag, LNCS 1768, Sept. 1999, pp. 254-269.

[11] W. Bender, D. Gruhl, and N. Morimoto, "Techniques for Data Hiding," Proc. SPIE, vol. 2420, San Jose, CA, Feb. 1995, p. 40.

Author Profile

Ms. M. H. Deokate was born on July 1, 1984. She received her BE degree in the field of Electronics and Telecommunication in 2006 from Shivaji University, Kolhapur. She is pursuing ME in Electronics Engineering at SCOE, affiliated to University of Pune. She is working as Lecturer at E & TC department of Sou. Venutai Chavan Polytechnic, Pune since July 2007. She has paper published in Conference. She is life member of ISTE, New Delhi and also a fellow of IETE, New Delhi.

S. R. Ganorkar was born on August 6; 1965.He has completed his ME in Adv. Electronics Engineering. His research interests are in Artificial Neural Network and Image Processing. He has 24 years of experience, 13 year in Industrial and 11 years of teaching experience. He is presently working as Associate Professor at E & TC department at Sinhgad College of Engineering, Pune. He has published 12 papers in International journal, 13 papers in International conference and 40 papers in national conference. He is life member of ISTE, New Delhi. He is also a fellow of IETE, New Delhi.