SPREAD SPECTRUM AUDIO WATERMARKING SCHEME BASED ON PSYCHOACOUSTIC MODEL

[1]Yüksel Tokur [2]Ergun Erçelebi

e-mail: tokur@gantep.edu.tr e-mail: ercelebi@gantep.edu.tr

[1]*Gaziantep University, MYO, 27310, Gaziantep, Turkey*

[2]*Gaziantep University, Electrical & Electronics Engineering, 27310,Gaziantep, Turkey*

Index Terms: Audio watermarking, psychoacoustic model

ABSTRACT

In this paper, we present a novel audio watermarking scheme using direct sequence spread spectrum (DSSS) method by which we can embed a text message as a watermark into an audio signal imperceptibly. The watermark embedding and extraction are based on the psychoacoustic model in the frequency domain. Experimental results show good robustness of the approach against amplitude compression, echo addition, noise addition, resampling, requantization, filtering and MP3 compression attacks.

I. INTRODUCTION

Over the last few years, audio watermarking has become an issue of significant interest. This is primarily motivated by a need to provide copyright protection to digital audio content. Digital watermarking is a technique to embed copyright or other information into the underlying data. The embedded data should be perceptually inaudible to maintain the quality of the host signal.

There are a number of desirable characteristics that a watermark should exhibit. These include that it should be difficult to notice, robust to common signal processing attacks, resistant to malicious attack of third parties, efficient to implement the system and detectable without original signal. Among them the first and most important problem that all watermarking schemes need to address is that of inserting data in the audio signal without deteriorating its perceptual quality. The early emphasis was on this requirement; since the applications were not concerned with signal distortions or intentional tampering that might remove a watermark. However, as watermarks are increasingly used for purposes of copyright control, robustness to common signal processing and resistance to tampering have become important considerations. To improve robustness the embedded watermarks should have larger energies. The importance of perceptual modeling and the needed to embed a signal in perceptually significant regions of an audio have been recognized. However, this requirement conflicts with the need for the watermark to be imperceptible.

Most watermark algorithms focus on image and video. But, only a few audio watermark algorithms have been reported. Bender et al. [1] 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 and show low robustness. Furthermore, they have a relative high complexity in the detection process. Swanson et al. [2] 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. Neubauer et al. [3, 4] introduced the bit stream watermarking concept. The basic idea of this method was to partly decode the input bit stream, add a perceptually hidden watermark in the frequency domain, and finally quantize and code the signal again. As an actual application, they embedded the watermarks directly into compressed MPEG-2 AAC (Advanced Audio Coding) bit streams so that the watermark could be detected in the decompressed audio data. Bassia et al. [5] 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 [6] proposed a watermark algorithm based on spread-spectrum coding using a linear predictive coding technique and FFT to determine the spectral shape.

Section 2 of paper presents basic idea. Section 3 discusses watermarking embedding. Section 4 covers watermark extractions. Section 5 presents experimental results. Section 6 shows conclusion.

II. BASIC CONCEPT

Our watermark-embedding scheme is based on a direct sequence spread spectrum (DSSS) method. An audio watermark must be inaudible in the watermarked audio signal. To do this, we used the psychoacoustic model based on many studies have shown that average human does not hear all the frequencies with the same effect. In fact, the psychoacoustic model is used to provide away of the determining which portion of a signal are inaudible and indiscernible to average human. The inner ear performs a short-time critical band analysis where the frequency to place transform occurs along the basilar membrane. The power spectra are not represented on a linear frequency scale but on limited frequency bands called critical bands. The auditory system can roughly be described as a band-pass filter bank. Simultaneous masking is a frequency domain phenomenon where a low-level signal can be made inaudible by a simultaneously occurring stronger signal. Such masking is the largest in the critical band in which the masker is located, and it is effective to a lesser degree in neighboring bands. The masking threshold can be measured, and low-level signals below this threshold are inaudible.

The proposed audio watermark scheme determines the portion of the audio signal that is masked by the human auditory system. In other words, this portion of the audio signal is inaudible by the average human. Then, the audio watermark is embedded in this portion in the frequency domain. Afterwards, we transform it to the time domain by the inverse fast Fourier transform and then it is added to the original audio signal. In the proposed audio watermarking, the watermark is formed in a row vector of the binary by applying a text to an input of the encoder.

III. WATERMARK EMBEDDING

A diagram of audio watermark embedding scheme is shown in Fig.1. 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 total frame size of 1280 samples and a 1280-point FFT (5 frames x256 samples =1280samples). Global masking threshold is shown in Fig.3.

b) Generate the watermark with using the encoder. The input text message is converted to binary. The watermark is applied to input of the binary encoder, the output of the encoder is a row vector of the binary representation of the text.

c) Using the masking threshold, shape the watermark signal to be imperceptible in the frequency domain.

d) Compute the inverse FFT of the shaped watermark signal.

e) Create the final watermarked audio signal by adding the watermark signal to the original audio signal in the time domain.

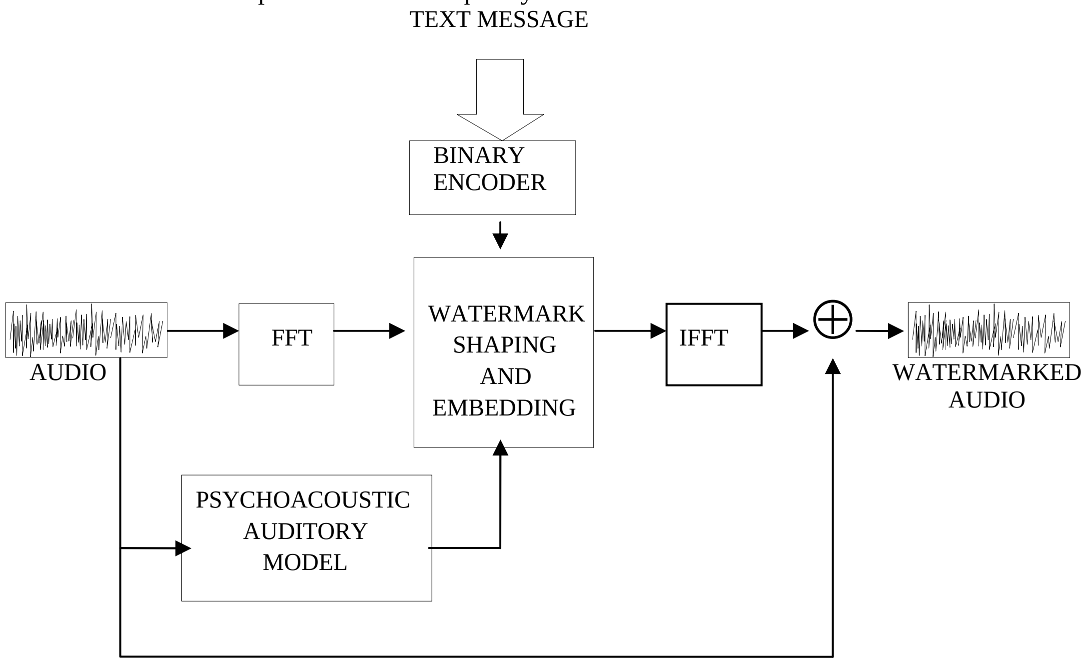

Fig. 1 The block diagram of the watermark embedding

IV. WATERMARK EXTRACTION

On designing a watermark extraction system, we need to consider the desired performance and the robustness of system. The proposed watermark extraction procedure does not require access to the original audio signal to extract the watermark signal.

Here, the psychoacoustic model is used to determine the inaudible portion of the audio signal that is called noise locations. Threshold compared with magnitude of the noise locations is computed by taking the average of the maximum value and the minimum value of noise locations. Then, the binary values are set to 1 if the magnitudes of the noise locations are greater than the threshold value. Otherwise, the binary values are cleared. The obtained binary values are decoded into the text message by the decoder. Then, the message strings can be extracted and it is easily seen that the extracted text matches the embedded text. The function encoder performs the embedding process while the decoder performing the extraction process.

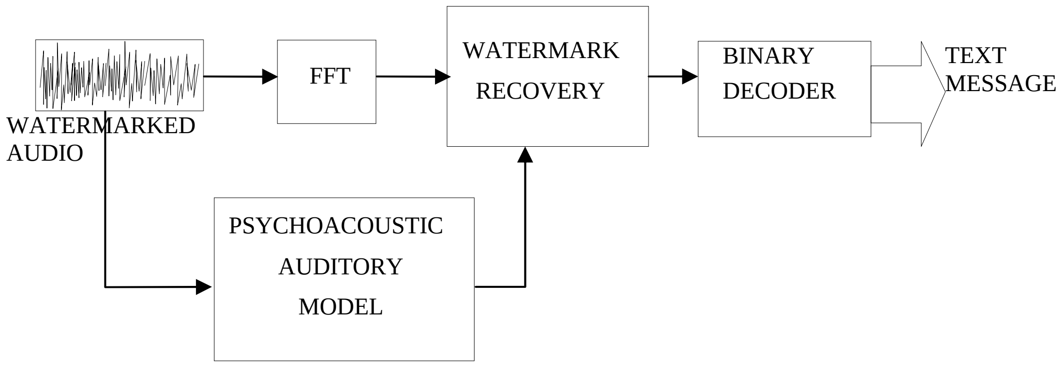

Fig.2 The block diagram of the watermark extraction

V. EXPERIMENTAL RESULTS

The audio segments used in experiments are stereo audio signals, 10-20 seconds long, sampled at 44.1 KHz with 16 bits resolution. Additional 96 bits were embedded in ten audio sequences.

Amplitude compression (comp 2:1 for above -20 dB, flat 1:1 for below -20 dB)

Echo addition: We tested watermarked signal echo addition with a delay 100ms and decay 50% respectively.

MPEG 1 Layer III audio compression: The robustness against MPEG 1 audio Layer III compression has been tested by using a compression rate 96 kbps for the watermarked signal.

Resampling: The original audio signals were sampled with a sampling a sampling rate of 44.1 kHz. The sampling rate of the watermarked audio data was reduced to 22.050 kHz and resampled to the original rate of 44.1 kHz. This causes audible distortions especially in audio tracks carrying high frequencies.

Requantization: Audio tracks sampled at 8-bit are often used in games and multimedia applications. We therefore tested the process of quantization of a 16-bit watermarked audio signal to 8-bit and back to 16-bit. This increases the incoherent background noise of the audio track due to the rounding errors during the processing.

Filtering: To test the robustness against filtering an fft filter was applied to a watermarked signal by the fft size of 8192 and windowing function of Hamming.

Noise addition: White noise with a constant level of 50 dB was added to the watermarked audio under the averaged power level of the audio signal.

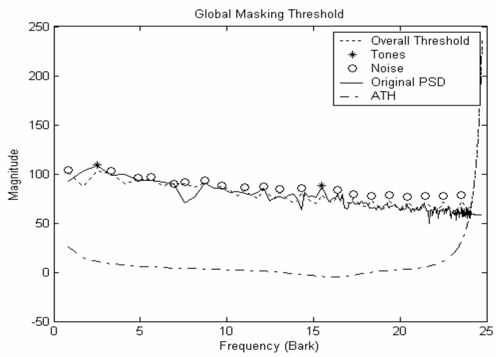

Fig.3 Global Masking Threshold

Table 1 Bit error for different attacks

Attack Type/ Bit Errors	Bit Errors		Bit Embedded
Attack Type	B1	B2	
Dynamic compression	3/96	0	192
Echo addition	13/96	3/96	192
Noise addition	8/96	2/96	192
Resampling	3/96	3/96	192
Requantization	2/96	7/96	192
MP3 Compression	30/96	30/96	192
Filtering	32/96	23/96	192
TOTALS	91	68	1344
BER (Bit Error Rate)	6.77 %	5.05 %	

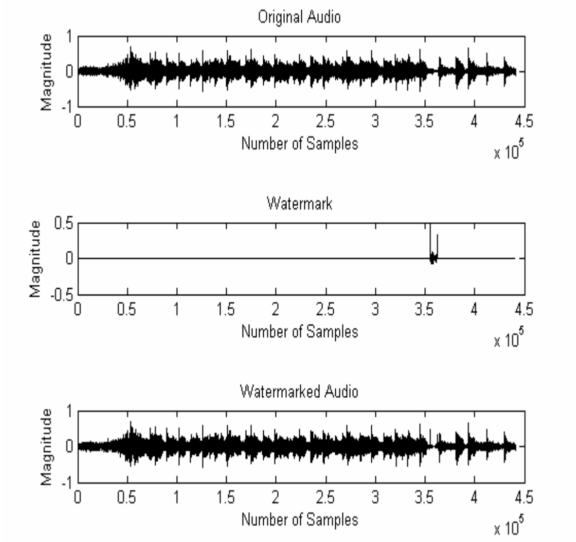

Fig.4 (a) Original audio signal (b) watermark signal (c) watermarked audio signal

VI. CONCLUSIONS

In this work, a novel audio watermarking scheme using direct sequence spread spectrum method, with the scope of copyright protection, has been proposed and tested. The watermark embedding and extraction are based on the psychoacoustic model in the frequency domain. Perceptual transparency of the algorithm is confirmed with listening tests and experimental results demonstrate that proposed technique is robust to amplitude compression, echo addition, noise addition, resampling, requantization, filtering and MP3 compression attacks. Finally, this technique does not require the original signal during extraction.

REFERENCE

1. W. Bender, D. Gruhl, and N. Morimoto, “Techniques for Data Hiding,” *Proc. SPIE*, vol. 2420, San Jose, CA, Feb. 1995, p. 40.
2. 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.
3. C. Neubauer and J. Herre, “Digital Watermarking and its Influence on Audio Quality,” *105th AES Convention, Audio Engineering Society preprint 4823, San Francisco,* Sept. 1998.
4. C. Neubauer and J. Herre, “Advanced Audio Watermarking and its Applications,” *109th AES Convention, Audio Engineering Society preprint 5176, Los Angeles*, Sept. 2000.
5. 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.
6. 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.
7. L. Boney, A. Tewfik , and K. Hamdy , *"Digital Watermarks for Audio Signals” IEEE Int. Conf. on Multimedia Computing and Systems* ,pp. 473-480, June 1996.

8. In-Kwon Yeo and H. J. Kim , “Modified patchwork algorithm: the novel audio watermarking scheme,” *IEEE Trans. Speech and Audio Processing*, vol. 11, no.4, pp. 381–386, July 2003.

9. Erçelebi E. and Tokur Y., “The Watermarking of Images in Frequency Domain Using Geometric Figures”, IJCI Proceedings of Intl. XII. Turkish Symposium on Artificial Intelligence and Neural Networks, Vol.1, No .1, pp. 447-449, 2003.

10. Tokur Y. and Erçelebi E., “Frequency-domain Wavelet Based Multiresolution Digital Watermarking,” *ELECO’2003*, *International Conference on Electrical and Electronics Engineering*, pp.180-184, Bursa, December 2003.