THE JOURNAL OF THE ACOUSTICAL SOCIETY OF KOREA VOL.25, NO.2E 2006. 6 pp. 56~59

Audio Watermark Design Using Hilbert Transform

Jingeol Lee*
*Department of Electronic Engineering, Paichai University
(Received April 7 2006; Accepted April 27 2006)

Abstract

A novel audio watermark design using Hilbert transform is proposed, and its superiority of performance over the existing design using the absolute values of the audio signal is demonstrated experimentally with measurements of bit error rate (BER) in correlation based decoder. (Classification No. 3.3)

Keywords: *Audio watermark, Hilbert transform, BER, Psychoacoustics*

I. Introduction

Watermarkings have attracted attention as means for copy right protection as digital contents flood into internet. In order to meet one of requirements on audio watermarks, that is, *imperceptibility*, the watermarks based on the psychoacoustic model, originally adopted in MPEG audio, or simply on the absolute values of the host audio signal have been developed [1-2]. In either way, the amplitude of the resultant watermarks varies with that of the host audio signal, which is justified by the fact that regions of the host audio signal with high amplitude provide high levels of the masking thresholds in human auditory system, which leads to high amplitude in the watermarks and vice versa. The watermark design using the psychoacoustic model leads to the maximum inaudible watermark level that can be embedded in the host audio signal. In that sense, the design using the psychoacoustic model is more accurate than that using the absolute value of host audio signal. However, the computational burden to the design using the psychoacoustic model is so high that it is not comparable to that to the design using the absolute value of host audio signal. On the other hand, the amplitude of the watermark using the absolute value of host audio signal concentrates on the value of 0, which is caused by the generalized Gaussian distribution of the audio signals [3-4], and thus makes the design disadvantageous in terms of bit error rate (BER) and robustness.

As stated earlier, the amplitude of the resultant watermark varies with that of the host audio signal. Based on this property, a novel audio watermark design using Hilbert transform is proposed in this paper. It is well known that Hilbert transform provides the complex envelopes of signals [5]. The envelopes of signals derived from Hilbert transform are utilized in designing the watermark, by which the amplitude of the watermark does not concentrates on the value of 0 as in the design using the absolute value of host audio signal, and the required computational burden is not so high as the design using the psychoacoustic model. The performance improvement of the proposed design over the design using the absolute value is demonstrated with BER measurements in correlation based decoder.

Corresponding author: Jingeol Lee (jingeol@pcu.ac.kr)
Department of Electronic Engineering Paichai University, Taejon 302-735 Korea

II. Watermark embedding and decoding

An audio signal $x(i)$ is divided into N_s frames composed of N samples

$$x_k(i) = x(kN+i) \qquad i = 0,\ldots,N-1,\ \ k = 0,\ldots,N_s-1. \tag{1}$$

The watermark embedding and the decoding based on the correlation normalized by the variance for each frame are as follows:

$$y_k(i) = x_k(i) + b_k t_k(i) s_k(i) \tag{2}$$

$$r_k = \sum_i \frac{y_k(i) t_k(i) s_k(i)}{\sigma_k^2} \tag{3}$$

where $b_k \in \{-1,1\}$ is an embedded bit, $t_k(i)$ is the masking thresholds, and $s_k(i)$ is the pseudorandom sequence whose values depend on a secret key. The decoded value of the bit b_k is determined by

$$\hat{b}_k = sign(r_k). \tag{4}$$

The masking thresholds associated with the existing watermark design using the absolute value of host audio signal [2] and with the proposed design using Hilbert transform are as follows, respectively:

$$t_k(i) = \alpha |x_k(i)| \tag{5}$$

$$t_k(i) = \beta A_k(i) \tag{6}$$

where $A_k(i)$ is the complex envelope of the signal given by Hilbert transform [5]

$$A_k(i) = \left(x_k^2(i) + (x_k(i) * h(i))^2\right)^{1/2} \tag{7}$$

$$h(i) = \frac{2}{\pi} \frac{\sin^2(\pi i/2)}{i}, \quad i \neq 0$$
$$= 0, \quad i = 0. \tag{8}$$

The existing watermark design is accomplished with the multiplication of the masking threshold given by (5) to a random sequence, followed by a filtering with the Hamming window. Since the main concerns in this paper is to examine the effect of the masking thresholds given by (5) and (6) on the performance of decoder, the filtering is excluded.

III. Experimental results

The test audio clips (sampling rate 44100 Hz, 30 sec long each) consist of a classic (Tchaikovsky symphony No. 6), a blues, a vocal, and a chorus. The pseudorandom sequence has two level discrete marginal distribution, $s_k(i) \in \{-1, 1\}$. A bit, $b_k \in \{-1,1\}$ is multiplied to the masking thresholds and random sequence, and embedded in a frame. The listening test reveals that $\alpha = 0.01$ in (5) corresponding to SNR=40 dB leads to transparent audio qualities. In order to compare the BER measurements for the existing method with those for the proposed method under the same level of SNR, the constant, $\beta = 0.007$ in (6) is experimentally determined. The SNR measurements for the proposed method are 40.0877 dB, 40.0880 dB, 40.0957 dB, 40.0877 dB for the classic, the blues, the vocal, and the chorus, respectively, which means that the powers of watermarks designed by the proposed method are a little bit lower than those designed by the existing method. The imperceptibility of watermarks designed by the proposed method is certified by the listening test and the psychoacoustic model based test. It turns out that the 7 subjects participating the test can not tell the difference between each original test audio and its watermarked audio by the proposed method. Fig. 1 shows the masking thresholds for a frame of the test audio of the vocal and the corresponding excitation levels of the watermark designed by the existing and the proposed method, respectively, which are produced by the psychoacoustic model 1 in MPEG audio. It shows that the similar levels of SNR of two watermarks lead to the similar excitation levels. It is well known that human can not perceive the presence of a target (watermark) as long as its excitation levels are below the masking thresholds of a masker (host audio), by which the imperceptibility is demonstrated with fig. 1 except the region around 10 KHz. The listening test confirms that this excess of the excitation level can be tolerable.

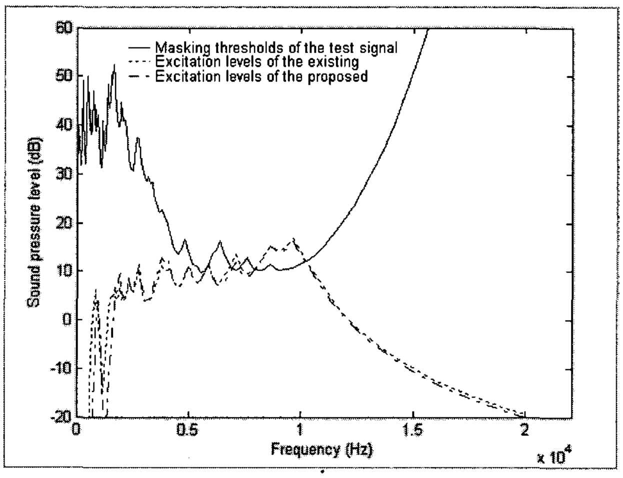


Fig. 1. Masking thresholds and excitation levels.

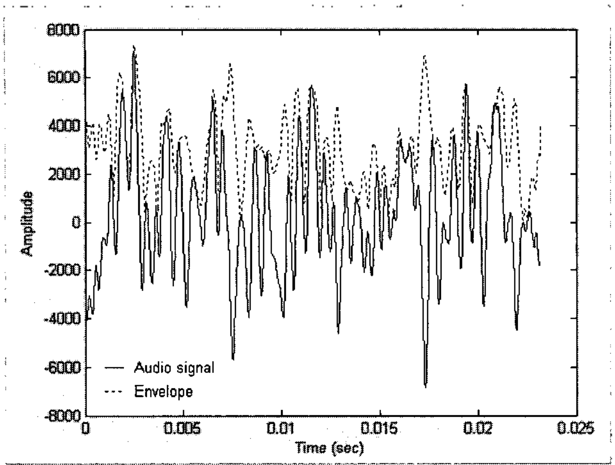
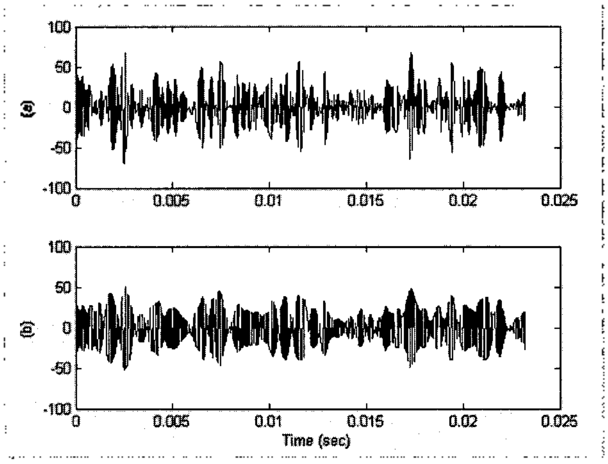
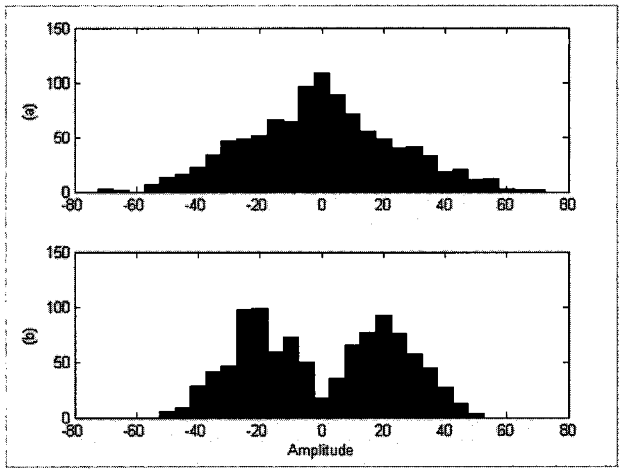


Fig. 2. Audio signal and its envelope.



(a) Watermark designed by the existing method.
(b) Watermark designed by the proposed method.
Fig. 3. Waveforms of the watermarks.

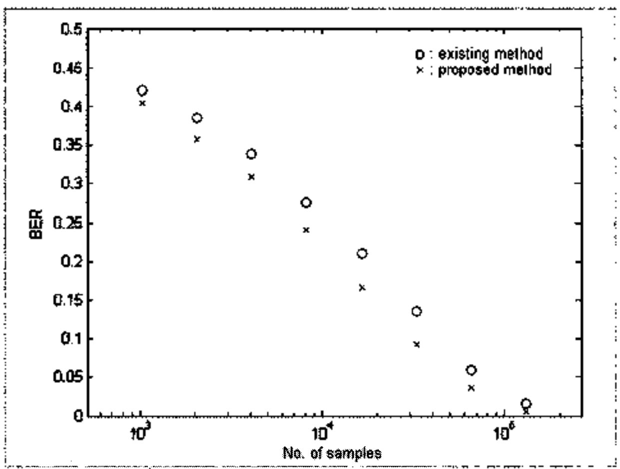
For the same segment of the audio used in the psychoacoustic model based test, fig. 2 shows the audio signal and its envelope derived from Hilbert transform of the signal, and fig. 3 and 4 shows the watermarks and their histograms designed by the



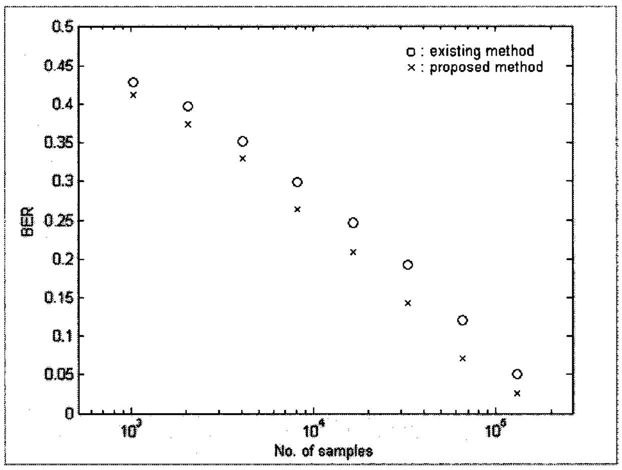
(a) Distribution of the watermark designed by the existing method.
(b) Distribution of the watermark designed by the proposed method.
Fig. 4. Distributions of the watermarks.

existing and the proposed method, respectively. As shown in fig. 3 and 4, the amplitude of the watermark using the existing method concentrates on the value of 0 whereas that of the watermark using the proposed method has the rarefaction on the value of 0.

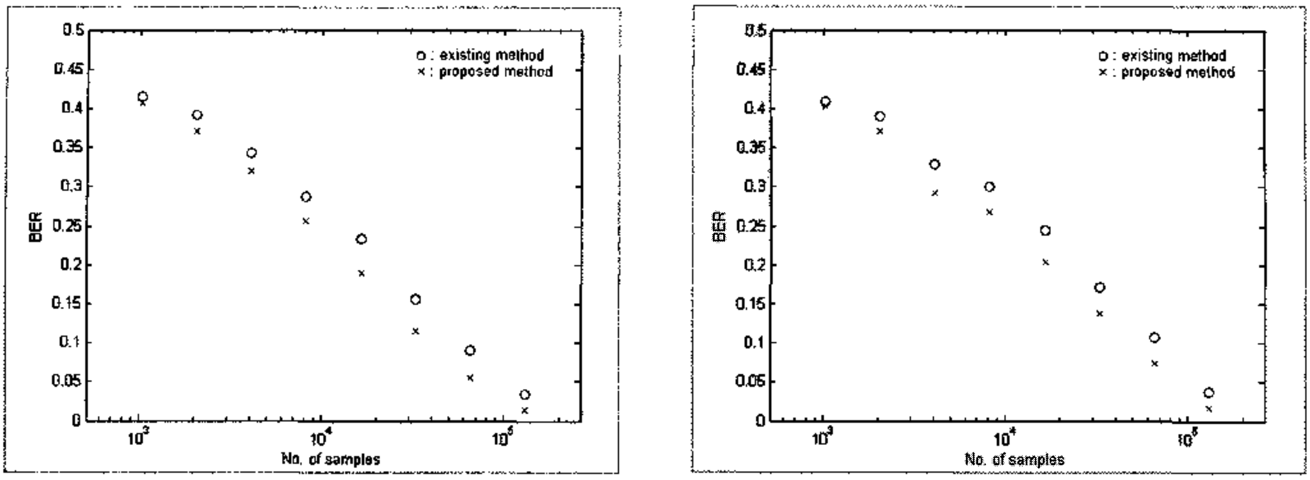
In order to check the effect of the distribution on the performance of decoder, the BER measurements are taken. The measurements for the test audios are presented in fig. 5 as the sample number of a frame is varied from 2^{10} to 2^{17} by doubling the sample number. The measurements for each of $b_k = 1$ and $b_k = -1$ are obtained using 1 random sequence for the sample number of 2^{10}, 2 sequences for $2^{11}, \cdots$ 128 sequences for 2^{17} in order to have the same number of the test bits of about 2500 for each of the sample number in a frame. The measurements demonstrate that the performance of the proposed method is superior to the existing method in terms of BER.



(a) Measurement for the classic.



(b) Measurement for the blues.



(c) Measurement for the vocal.

(d) Measurement for the chorus.

Fig. 5. BER measurements.

IV.. Conclusions

The proposed audio watermark design is based on the signal dependent watermark derived from Hilbert transform. The imperceptibility of the proposed method is certified by the listening and the psychoacoustic model based test. The superiority of performance over the existing method is demonstrated experimentally with measurements of BER in decoder, which is attributed to the distribution of the watermark designed by the proposed method in contrast to that designed by the existing method.

References

1. Mitchell D. Swanson, Bin Zhu, Ahmed H. Tewfik, and Laurence Boney, "Robust Audio Watermarking using Perceptual Masking", Signal Processing, **66** 337-355, 1998
2. Paraskevi Bassia, Ioannis Pitas, and Nikos Nikolaidis, Robust Audio Watermarking in the Time Domain, IEEE Trans. Multimedia, **3** (2) 232-241, 2001
3. Te-Won Lee, *Independent Component Analysis* (Kluwer Academic Publishers, Boston, 1998), Chap. 2 (36)
4. Jingeol Lee, "Optimal Parameter Estimation of the ML Test Based Audio Watermark Decoder", The Journal of the Acoustical Society of Korea, **25** (2) 56-60, 2006
5. Alan V. Oppenheim, Ronald W. Schafer, *Digital Signal Processing* (Prentice-Hall, Inc., Englewood, 1975), Chap. 7, 360-361

[Profile]

•Jingeol Lee

Jingeol Lee received the Bachelor of Engineering degree; the Master of Science degree from Korea University in 1981 and 1985, respectively; and the Doctor of Philosophy degree from University of Florida in 1994. From 1982 to 1990, he was with Agency for Defense Development, and from 1995 to 1996, he had worked for Samsung Electronics, where he was involved in research and development of military electronics throughout his job experiences. Since 1997 he has been with the Department of Electronic Engineering, Paichai University. His professional interests include watermarkings and acoustics.