THAM 16.5

DIGITAL AUDIO WATERMARKING IN THE CEPSTRUM DOMAIN

Sang-Kwang Lee and Yo-Sung Ho

Kwangju Institute of Science and Technology
1 Oryong-Dong Puk-Gu, Kwangju, 500-712, Korea

ABSTRACT

Watermarking is a technique used to label digital media by hiding copyright or other information into the underlying data. The watermark must be imperceptible and undetectable by the user and should be robust to various types of distortion. In this paper, we propose a new algorithm for digital audio watermarking in the cepstrum domain, where we insert a watermark into the cepstral components of the audio signal using techniques analogous to spread spectrum communications, hiding a narrow band signal in a wideband channel.

INTRODUCTION

In recent days, there has been significant interest in watermarking. This is primarily motivated by a need for providing copyright protection to digital contents, such as audio, image and video. Digital representation of copyrighted materials offers various advantages; however, the fact that an unlimited number of perfect copies can be illegally produced is a serious threat to the right of content owners. Watermarking is used for owner identification, royalty payments, and authentication by determining whether the data has been altered in any manner from its original form.

A watermark must be embedded in the data such that it is imperceptible by the user. Moreover, the watermark should be inaudible or statistical invisible to prevent unauthorized detection and/or removal. The watermark should also have similar compression characteristics as the original signal and be robust to any manipulations or signal processing operations on the host data, e.g., filtering, resampling, compression, noise, cropping, A/D-D/A conversions, etc. The watermark should also be embedded directly in the data, not in the header, and be self-clocking for ease of detection in the presence of cropping and time-scale change operations. The watermark should be characteristic of the author, but a private should not be able to detect the watermark by comparing several signals belonging to the same author. The signal should be degraded when the watermark is removed through any unauthorized means.

Compared to image and video signals, audio signals are represented by much less samples per time interval. This alone indicates that the amount of information that can be embedded robustly and inaudibly is much lower than for visual signal. An additional problem in audio watermarking is that the human audible system (HAS) is much more sensitive than the human visual system (HVS), and that inaudibility is much more difficult to achieve than invisibility for images.

Boney *et al.* [1] propose a spread-spectrum approach for audio watermarking. They use a pseudo-random sequence that is filtered in several stages in order to exploit long-term and short-term masking effects of HAS. Baassia and Pitas [2] apply a very straightforward time-domain spread-spectrum watermarking method to audio signals. They report robustness against audio compression, filtering and resampling.

In this paper, we propose a spread spectrum technique to insert a watermark into the cepstral component of the audio signal.

WATERMARK EMBEDDING

Frequency-domain watermarking was first intorduced by Cox *et al.* [3], who have independently developoed perceptually adaptive methods based on modulation. They draw parallels between their technology and spread spectrum communication since the watermark is spread over a set of visually important frequency components.

In their watermark embedding scheme, the watermark is spread over several frequency bins so that the energy in any bin is very small and certainly undetectable. Spreading the watermark throughout the audio spectrum ensures high security against unintentional or intentional attack. The watermarked signal does not produce any perceptual distortion.

Let us assume an audio signal of N samples $x = x(1),\ldots,x(N)$ and a pseudo-random sequence $w = w(1),\ldots,w(N)$. The watermarked sample $z(n)$ is represented as

$$z(n) = f\big(x(n), w(n)\big) \qquad (1)$$

where $f(\cdot)$ is a function to embed a watermark.

0-7803-6301-9/00 $10.00 © 2000 IEEE

In the frequency domain, Eq. (1) can be express as

$$Z(\omega) = F(\omega) \text{ (Fourier domain)} \qquad (2)$$

where $Z(\omega)$ and $F(\omega)$ are Fourier transforms of $z(n)$ and $f(x,w)$, respectively, taken over a finte length of size N.

The complex cepstrum $c_z(n)$ of $z(n)$, is defined by

$$c_z(n) = F^{-1}\{\log(Z(\omega)\}$$
$$= F^{-1}\{\log(F(\omega))\} \text{ (cepstrum domain)} \quad (3)$$

where $F^{-1}\{\cdot\}$ denotes the 1-D inverse Fourier transform. When we insert $w(n)$ into $x(n)$, we specify a scaling parameter α that determines the extent to which $w(n)$ alters $x(n)$. In order to embed a watermark, we modify Eq. (3) as

$$c_z(n) = c_z(n) + \alpha(c_z(n))w(n) \qquad (4)$$

where $c_x(n)$ is the complex cepstrum $x(n)$. We can view multiple scaling parameters $\alpha(\cdot)$ as a relative measure of how much one must alter $w(n)$ to change the perceptual quality of $x(n)$.

Fig. 1 shows the watermark insertion process.

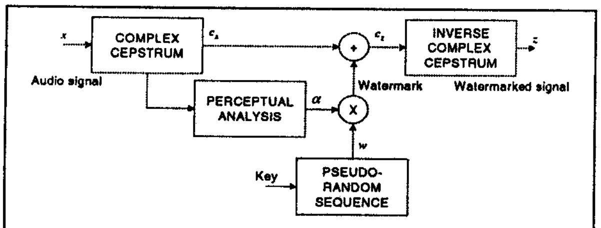

Fig. 1 Watermark insertion process.

Since the complex cepstrum takes the logarithm, it is robust against the differences in scale. Also, the multiple scaling parameters are used for the appropriate perceptual system.

WATERMARK DETECTION

To verify the presence of the watermark, we measure the cross-correlation between the recovered watermark $w^*(n)$ and the original watermark $w(n)$. The cross-correlation is given by

$$r(m) = \sum_{n=1}^{N} \alpha^{-1}(c_x(n))w^*(n)w(n+m)$$
$$= \sum_{n=1}^{N} w(n)w(n+m), \; 0 \le m \le N-1 \qquad (5)$$

where $w^*(n) = c_z(n) - c_x(n) = \alpha(c_x(n))w(n)$.

Ideally, a pseudo-random sequence should have an autocorrelation function that has correlation properties similar to the white noise. Figure 2 shows the watermark detector response to 1000 randomly genertated watermarks. The detection algorithm gives no false alarm.

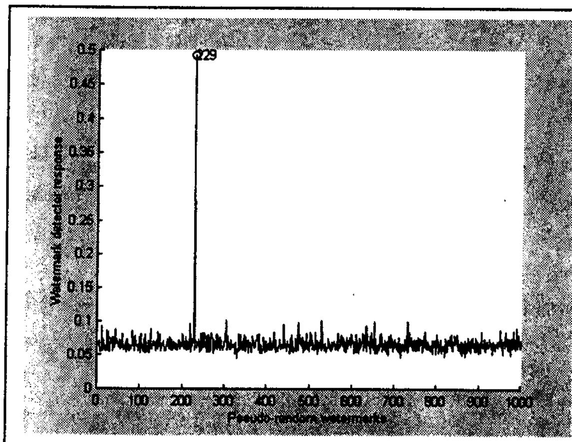

Fig.2 Watermark detector response (Key is 229).

CONCLUSIONS

In this paper, we propose a new algorithm for digital audio watermarking in the cepstrum domain. We insert a watermark into the cepstral components of the audio signal using a spread spectrum technique. Our watermarking scheme is robust against the differences in scale. Also, it considers the perceptual capacity of each frequency.

ACKNOWLEGEMENT

This work was supported in part by KOSEF through UFON, and in part by MOE through BK21.

REFERENCES

[1] L. Boney, A. H. Twefik and K. N. Hamdy, "Digital watermarks for audio signals," in *Proc. Europ. Signal Processing Conf.*, Trieste, Italy, Sept. 1996.

[2] P. Bassia and I. Pitas, "Robust audio watermarking in the time domain," in *Proc. Europ. Signal Processing Conf.*, Rhodes, Greece, Sept. 1998.

[3] I. Cox, J. Kilian, T. Leighton and T. Shamoon, "Secure spread spectrum watermarking for images, audio and video," in *Proc. IEEE Int. Conf. Image Processing*, Lausanne, Switzerland, Sept. 1996.

[4] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*, Prentice-Hall, Englewood Cliffs, NJ, 1978.