MODIFIED SPREAD SPECTRUM AUDIO WATERMARKING ALGORITHM

Libo Zhang[1], Sridhar Krishnan[1], and Heping Ding[2]
[1]Department of Electrical and Computer Engineering, Ryerson University, Toronto, Canada, {libo, krishnan}@ryerson.ca
[2]Institute for Microstructural Sciences, National Research Council of Canada, Ottawa, Canada, heping.ding@nrc-cnrc.gc.ca

1. INTRODUCTION

Digital audio watermarking is a technique to embed some information, i.e. the watermark, into the host audio signal in an imperceptible and robust way. The watermark can be used for copyright protection, usage monitoring, etc. The imperceptibility requires that the watermark is embedded without causing perceptual distortion and the robustness requires that the watermark can be extracted even after the watermarked signal is distorted by different signal processing manipulations and/or intentional attacks.

Watermarking can be thought of as a telecommunication means to transmit the watermark signal over the medium of the host signal. The imperceptibility limits the power of the watermark signal to be lower than the hearing threshold and thus the communication is characterized by a low SNR. Many robust audio watermarking algorithms are based on Spread Spectrum (SS) technique due to its robustness to interference and low-energy requirement [1].

In the following, the conventional SS audio watermarking will be reviewed first and then a modification will be presented to increase the robustness significantly.

2. CONVENTIONAL SS WATERMARKING

The projection of one vector $\vec{u}$ onto another vector $\vec{v}$, i.e. *normalized correlation,* is defined as,

$$u_v = v_u = \vec{u} \bullet \vec{v} \equiv \frac{1}{N}\sum_{i=1}^{N} u_i v_i \qquad (1)$$

where N is the length of the vectors $\vec{u}$ and $\vec{v}$.

In the standard SS watermarking scheme, a watermark bit $b = \{\pm 1\}$ is spread by a normalized pseudo-random sequence $\vec{w}$, satisfying $w \bullet w = 1$ and of length N, to generate the spread watermark $b \cdot w$ and this spread watermark is embedded into the host signal vector $\vec{x}$ of the same length. The result is the following watermarked signal,

$$y = x + \alpha \cdot (b \cdot w) \qquad (2)$$

where the perceptual factor $\alpha > 0$ controls the perceptibility of the watermark. The watermark signal power is as follows,

$$W = \| y - x \|^2 = (y - x) \bullet (y - x) = \alpha^2 \qquad (3)$$

At the receiver, the received signal $\vec{r}$, which is the watermarked signal $\vec{y}$ corrupted by an additive transmission noise $\vec{n}$, is projected onto the sequence $\vec{w}$ as

$$c = r \bullet w = (y + n) \bullet w = \alpha \cdot b + (x + n) \bullet w \qquad (4)$$

When the spreading factor N is large enough, the second term decreases so the first term dominates and thus *sign(c)* can be used to determine the embedded watermark bit *b*.

Assuming $\vec{x} \sim N(0, \sigma_x^2)$ and $\vec{n} \sim N(0, \sigma_n^2)$, it can be easily shown that $c \sim N(\alpha \cdot b, \frac{\sigma_x^2 + \sigma_n^2}{N})$ and thus the corresponding Bit Error Rate (BER) is as follows,

$$BER = Q(\frac{\mu}{\sigma}) = Q\left(\sqrt{\frac{N \cdot \alpha^2}{\sigma_x^2 + \sigma_n^2}}\right) = Q\left(\sqrt{N \cdot WNR \cdot \frac{1}{SNR + 1}}\right) \qquad (5)$$

where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^{+\infty} e^{-\frac{u^2}{2}} du$ is normally called the complementary error function, $WNR = \frac{\alpha^2}{\sigma_n^2}$ denotes the power ratio of the watermark to the noise, and $SNR = \frac{\sigma_x^2}{\sigma_n^2}$ is the power ratio of the host signal to the noise.

It can be seen from (5) that both the host signal and the attacks act as noise in watermark extraction. Normally the former impacts the extraction more severely since $\sigma_x^2 >> \sigma_n^2$. For this reason a large spreading factor N is often needed, which results in a very small embedding rate.

3. MODIFIED SS WATERMARKING

Geometrically, the vector $\vec{w}$ and a normalized orthogonal vector $\vec{v}$, satisfying $w \bullet v = 0$ and $v \bullet v = 1$, determine a plane. The projection of the host signal onto this subspace can be represented by the point (x_w, x_v), which maps onto the line w=v at the point $(\frac{x_w + x_v}{2}, \frac{x_w + x_v}{2})$. If the watermark is embedded in the direction of $\vec{w}$ from this point, as shown in Fig. 1, the extraction can be eased. This is shown as follows.

The distortion vector due to this mapping can be expressed as $(\frac{x_v - x_w}{2}, \frac{x_w - x_v}{2})$. The watermarked signal is thus

$$\vec{y} = \vec{x} + \frac{x_v - x_w}{2} \cdot \vec{w} + \frac{x_w - x_v}{2} \cdot \vec{v} + \alpha_{MSS} \cdot (b \cdot \vec{w}) \qquad (6)$$

The watermark signal power in this case is

$$W = E[\| \vec{y} - \vec{x} \|^2] = E[(\vec{y} - \vec{x}) \bullet (\vec{y} - \vec{x})] = \alpha_{MSS}^2 + \frac{\sigma_x^2}{N} \qquad (7)$$

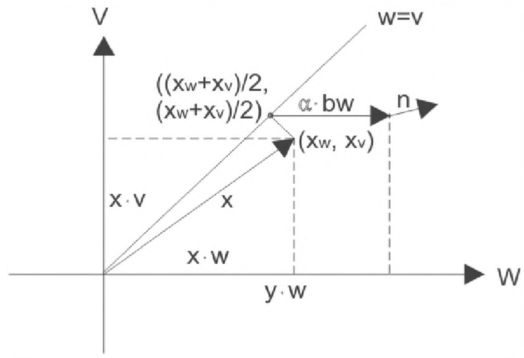

Fig. 1. Proposed watermark embedding algorithm

Based on the above explanation, it can be easily shown that the sign of the following expression can be use to determine the embedded watermark bit,

$$cc = r \bullet w - r \bullet v = \alpha_{MSS} \cdot b + (w - v) \bullet n \qquad (8)$$

To illustrate the performance improvement, we compare BERs of the modified and the conventional SS systems with the same watermark signal power, i.e., α_{MSS} is determined as per

$$\alpha_{MSS}^2 + \frac{\sigma_x^2}{N} = \alpha^2 \quad \Rightarrow \quad \alpha_{MSS} = \sqrt{\alpha^2 - \frac{\sigma_x^2}{N}} \qquad (9)$$

Thus the BER of the modified system can be expressed as

$$BER = Q(\frac{\mu}{\sigma}) = Q\left(\sqrt{\frac{N \cdot \alpha_{MSS}^2}{2\sigma_n^2}}\right) = Q\left(\sqrt{N \cdot WNR \cdot \frac{N - SWR}{2N}}\right) \qquad (10)$$

where $SWR = \frac{\sigma_x^2}{\alpha^2}$ denotes the power ratio of the host signal to the watermark.

This modification improves the conventional scheme under most situations. For example, when SWR=*25dB* and SNR=*5dB*, there is an improvement of about *3dB* at BER of 10^{-5}, as shown in Fig. 2. With the spreading factor *N=4000* (i.e. *N**WNR≈*16* dB), the BER of the conventional SS is 9.675×10^{-4}, while that of the modified SS is 8.864×10^{-6}, an improvement of the order of *2*.

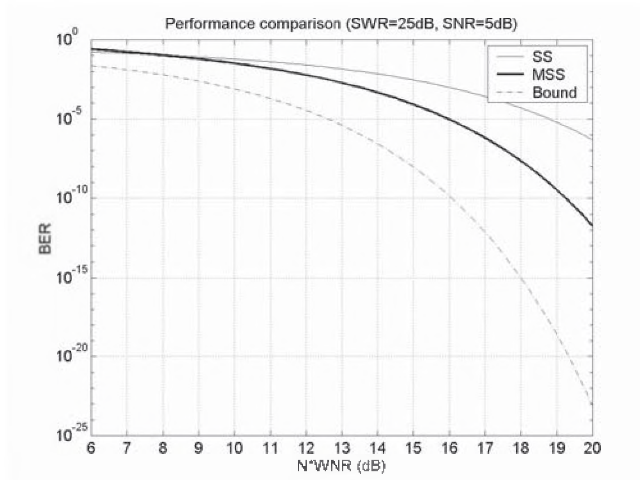

Fig. 2 Bit Error Rate (BER) of the proposed algorithm (Bound: theoretical limit, with host known at receiver)

4. SIMULATION

Audio pieces of different genres, sampled at *44.1* kHz, are tested against the proposed watermarking system. The watermark is spread with the factor *N=4000*, which corresponds to the embedding rate of *11* bps, and the amplitude is controlled by SWR=*32*dB. To evaluate the subjective fidelity, *8* listeners were asked to report dissimilarities between the host and the watermarked using a *5*-point impairment scale (5: imperceptible, 4: perceptible but not annoying, 3: slightly annoying, 2: annoying, 1: very annoying). The average mean opinion score (MOS) was *4.73* with deviation of *0.29*.

To evaluate the robustness, the following attacks are simulated,

- Re-sampling: subsequently down and up sampling.
- Low pass: cutoff at *4 kHz*.
- MP3: compressed at *48kbs*.
- Noise: *-20dB* white Gaussian noise addition.

The mean BER across all the above attacks is *0.075%*. Compared with the conventional SS, the modified SS improves the BER by an averaged order of *2*.

In this paper, we proposed a blind, transparent and simple audio watermarking algorithm. This algorithm can be further improved by introducing more mapping grids.

REFERENCES

1. D. Kirovski and H. S. Malvar, *Robust Spread-Spectrum Audio Watermarking,* IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 1345-1348, 2001.