Hindawi Publishing Corporation
EURASIP Journal on Advances in Signal Processing
Volume 2008, Article ID 453580, 7 pages
doi:10.1155/2008/453580

Research Article

A Robust Zero-Watermarking Algorithm for Audio

Ning Chen and Jie Zhu

The Department of Electronic Engineering, Shanghai Jiao Tong University, DongChuan Road no. 800, Shanghai 200240, China

Correspondence should be addressed to Ning Chen, chenning_750210@163.com.

Received 30 July 2007; Accepted 25 November 2007

Recommended by Mark Liao

In traditional watermarking algorithms, the insertion of watermark into the host signal inevitably introduces some perceptible quality degradation. Another problem is the inherent conflict between imperceptibility and robustness. Zero-watermarking technique can solve these problems successfully. Instead of embedding watermark, the zero-watermarking technique extracts some essential characteristics from the host signal and uses them for watermark detection. However, most of the available zero-watermarking schemes are designed for still image and their robustness is not satisfactory. In this paper, an efficient and robust zero-watermarking technique for audio signal is presented. The multiresolution characteristic of discrete wavelet transform (DWT), the energy compression characteristic of discrete cosine transform (DCT), and the Gaussian noise suppression property of higher-order cumulant are combined to extract essential features from the host audio signal and they are then used for watermark recovery. Simulation results demonstrate the effectiveness of our scheme in terms of inaudibility, detection reliability, and robustness.

1. INTRODUCTION

Unauthorized copying and distribution of digital data creates a severe problem in the protection of intellectual property rights. The embedding of digital watermark into multimedia content has been proposed to tackle this problem. However, currently available digital watermarking schemes mainly focus on image and video copyright protection and only a few audio watermarking techniques have been reported [1]. Comparing with the development of digital video and image watermarking, digital audio watermarking provides a special challenge because the human auditory system (HAS) is extremely more sensitive than human visual system (HVS) [2].

In traditional audio watermarking techniques, either in spatial domain, transform domain, or dual domain [3, 4], the embedding of watermark into the host audio inevitably introduces some audible quality degradation. Another problem is the inherent conflict between the imperceptibility and robustness. Then, zero-watermarking technique was proposed by some researchers to solve these problems [5–14]. Instead of embedding watermark into the host signal, the zero-watermarking approach just constructs a binary pattern based on the essential characteristics of the host signal and uses them for watermark recovery. An efficient zero-watermarking technique was presented in [5]. At first, the host image was rearranged randomly in the spatial domain and the result of which was divided into blocks according to the size of the watermark. Next, the variance of each block was compared with the average of all variances to generate a binary pattern. Finally, an exclusive or (XOR) operation was performed between the binary pattern and the binary watermark to obtain a secret key. For watermark recovery, a binary pattern was extracted from the test image first, and then the XOR operation was applied to the extracted binary pattern and the secret key to recover the binary watermark. In [6, 11], the property of the natural images that the vector quantization (VQ) indices among neighboring blocks tend to be very similar was utilized to generate the binary pattern. In [12], a scheme that combined the zero-watermarking with the spatial-domain-based neural networks was proposed, in which the differences between the intensity values of the selected pixels and the corresponding output values of the neural network model were calculated to generate the binary pattern. In [13], some low-frequency wavelet coefficients were randomly selected from the original image by chaotic modulation and used for character extraction. And in [14], two zero-watermarks were constructed from the host image. One was robust to signal process and central cropping, which was constructed from low-frequency coefficients in discrete wavelet transform domain and the other was robust to general geometric distortions as well as signal process, which was constructed from DWT coefficients of log-polar mapping of the host image. However, all these zero-watermarking techniques are designed for still image and their robustness against some signal processing manipulations or malicious attacks is not satisfactory. In this paper, a novel robust zero-watermarking technique for audio signal is proposed. The multiresolution characteristic of DWT, the energy compression characteristic of DCT, and the Gaussian noise suppression property of higher-order cumulant are combined to extract essential features from the host audio signal and they are then used for watermark recovery. Simulation results demonstrate the effectiveness of our algorithm in terms of inaudibility, detection reliability, and robustness against both common audio signal processing manipulations and malicious attacks provided by the practical audio watermarking evaluation tool, Stirmark for Audio v0.2 [15]. The remainder of this paper is organized as follows. In Section 2, the definition and properties of higher-order cumulant are reviewed. In Section 3, the proposed zero-watermarking method is described in detail. The simulation results and discussions are given in Section 4. And the conclusions are drawn in Section 5.

2. HIGHER-ORDER CUMULANT

The properties of higher-order statistics are becoming more and more thoroughly studied in the field of signal processing. One property of great interest is the fact that the cumulant of Gaussian signal disappears entirely at higher orders. Since many noise and interference signals have Gaussian distribution, this property offers the possibility that the higher-order statistics may be useful in signal recovery or interference mitigation [16]. In this paper, the higher-order cumulant is combined in the proposed algorithm to improve its robustness against Gaussian noise addition.

Let $\vec{v} = (v_1, v_2, \dots, v_k)$ and $\mathbf{x} = (x_1, x_2, \dots, x_k)$, where $(x_1, x_2, \dots, x_k)$ denotes a collection of random variables. The kth-order cumulant of these random variables is defined as the coefficient of $(v_1, v_2, \dots, v_k)$ in the Taylor series expansion (provided it exists) of the cumulant-generation function [17]

$$K(\vec{v}) = \ln E\{\exp(j\vec{v}\mathbf{x})\}. \tag{1}$$

Let $\{x(t)\}$ be a zero-mean kth-order stationary random process. The kth-order cumulant of this process, denoted $C_{k,x}(\tau_1, \tau_2, \dots, \tau_{k-1})$, is defined as the joint kth-order cumulant of the random variables $x(t), x(t+\tau_1), \dots, x(t+\tau_{k-1})$, that is,

$$C_{k,x}(\tau_1, \tau_2, \dots, \tau_{k-1}) = \operatorname{cum}(x(t), x(t+\tau_1), \dots, x(t+\tau_{k-1})). \tag{2}$$

Cumulant has the following important properties.

[CP1] If α_i $(i = 1, \dots, k)$ are constants and x_i $(i = 1, \dots, k)$ are random variables, then

$$\operatorname{cum}(\alpha_1 x_1, \dots, \alpha_k x_k) = \left(\prod_{i=1}^{k} \alpha_i\right) \operatorname{cum}(x_1, \dots, x_k). \tag{3}$$

[CP2] Cumulants are symmetric in their arguments, that is,

$$\operatorname{cum}(x_1, \dots, x_k) = \operatorname{cum}(x_{i_1}, \dots, x_{i_k}), \tag{4}$$

where $(i_1, \dots, i_k)$ is a permutation of $(1, \dots, k)$.

[CP3] Cumulants are additive in their arguments, that is,

$$\begin{aligned}\operatorname{cum}(x_0 + y_0, z_1, \dots, z_k) \\ = \operatorname{cum}(x_0, z_1, \dots, z_k) + \operatorname{cum}(y_0, z_1, \dots, z_k).\end{aligned} \tag{5}$$

[CP4] If α is a constant, then

$$\operatorname{cum}(\alpha + z_1, z_2, \dots, z_k) = \operatorname{cum}(z_1, \dots, z_k). \tag{6}$$

[CP5] If the random variables $\{x_i\}$ are independent of the random variables $\{y_i\}, i = 1, 2, \dots, k$, then

$$\begin{aligned}\operatorname{cum}(x_1 + y_1, \dots, x_k + y_k) \\ = \operatorname{cum}(x_1, \dots, x_k) + \operatorname{cum}(y_1, \dots, y_k).\end{aligned} \tag{7}$$

[CP6] If a subset of the k random variables $\{x_i\}$ is independent of the rest, then

$$\operatorname{cum}(x_1, \dots, x_k) = 0. \tag{8}$$

The cumulants of an independent, identically distributed random sequence are delta functions, that is to say, if $u(t)$ is such process, then $C_{k,u}(\tau_1, \tau_2, \dots, \tau_{k-1}) = \gamma_{k,u}\delta(\tau_1)\delta(\tau_2)\cdots\delta(\tau_{k-1})$, where $\gamma_{k,u}$ is the kth-order cumulant of the stationary random sequence $u(n)$.

Suppose $z(n) = y(n) + v(n)$, where $y(n)$ and $v(n)$ are independent, then from [CP5]

$$\begin{aligned}C_{k,z}(\tau_1, \tau_2, \dots, \tau_{k-1}) \\ = C_{k,y}(\tau_1, \tau_2, \dots, \tau_{k-1}) + C_{k,v}(\tau_1, \tau_2, \dots, \tau_{k-1}).\end{aligned} \tag{9}$$

If $v(n)$ is Gaussian (colored or white) and $k \geq 3$, then $C_{k,z}(\tau_1, \tau_2, \dots, \tau_{k-1}) = C_{k,y}(\tau_1, \tau_2, \dots, \tau_{k-1})$. This makes the higher-order cumulant quite robust to additive measurement noise, even if that noise is colored. In essence, cumulants can draw non-Gaussian signals out of Gaussian noise, thereby boosting their signal-to-noise ratios.

3. PROPOSED ZERO-WATERMARKING SCHEME

3.1. Fundamental theory

The wavelet transform is a time-scale analysis. Its multiresolution decomposition offers high-temporal localization for high frequencies while offering high-frequency resolution for low frequencies. So the wavelet transform is a very good tool to analyze the audio signal which is nonstationary. Cox et al. suggest that a watermark should be placed in perceptually significant regions of the host signal if it is to be robust

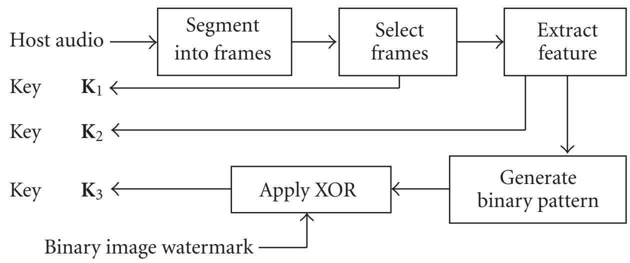


Figure 1: Embedding process.

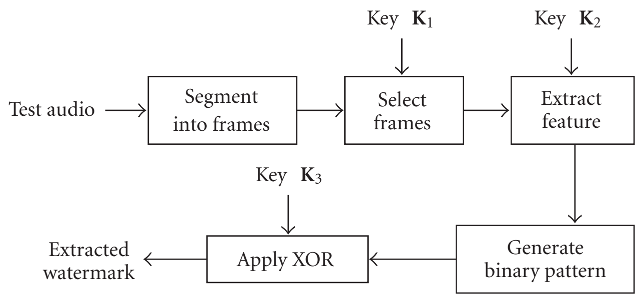


Figure 2: Extraction process.

[18]. In the proposed scheme, three-level wavelet decomposition is applied to get the low-frequency subband of the host audio, which is the perceptually significant region of it. The decorrelation, energy compaction, separability, symmetry, and orthogonality properties of discrete cosine transform lead to its widespread deployment in audio processing standard, for example, MPEG-1. To make the proposed scheme resist lossy compression operation such as Mp3 compression, DCT is performed on the obtained low-frequency wavelet coefficients. And considering the Gaussian signal suppression property of higher-order cumulant, the fourth-order cumulants of the obtained DWT-DCT coefficients are calculated to ensure the robustness of the proposed scheme against various noise addition operations. Finally, the essential features extracted based on DWT, DCT, and higher-order cumulant are used for generating binary pattern. Thus, any manipulations attempting to destroy the watermark will destroy the host audio signal first, so the high robustness of the proposed scheme is ensured. And since the essential features of different host audio signals are different, the detection reliability can also be achieved.

The block diagrams of embedding process and extraction process of the proposed zero-watermarking scheme are shown in Figures 1 and 2, respectively. In the embedding stage, the host audio signal is first segmented into equal frames according to the size of watermark and the frames with larger energy values are selected for watermark embedding. Next, DWT is performed on each selected frame to get its coarse signal, on which DCT is performed. Then, the higher-order cumulants of the obtained DWT-DCT coefficients are calculated and those elements with large absolute value are selected to generate a binary pattern. Finally, the watermark detection key is generated by applying XOR operation to the binary pattern and the binary-valued image watermark to be embedded. In the extraction stage, a binary pattern is calculated from the test audio signal first and then an estimated watermark is obtained by performing XOR operation between the obtained binary pattern and the watermark detection key.

3.2. *Embedding process*

Let $\mathbf{A} = \{a(i) \mid i = 0, \ldots, L_A - 1\}$ be the host audio signal and let $\mathbf{W} = \{w(i,j) \mid w(i,j) \in \{0,1\}\}$, where $i = 0, \ldots, M-1$, $j = 0, \ldots, N-1$, be the binary-valued image watermark to be embedded, then the watermark embedding procedure can be described as follows.

Step 1. At first, $\mathbf{A}$ is segmented into L frames, denoted as $\mathbf{F} = \{\mathbf{f}_i \mid i = 0, \ldots, L-1, L > 2MN\}$, and each frame has L_f samples. Next, the energy value of each frame is calculated and all the frames are rearranged in order of decreasing energy value. Then, the first T frames are selected for watermark embedding. And, the indices of the selected frames in $\mathbf{F}$, denoted as $\mathbf{I}_1$,

$$\mathbf{I}_1 = \{i(k) \mid i(k) \in \{0, \ldots, L-1\}, k = 0, \ldots, T-1\} \quad (10)$$

are saved as the first secret key $\mathbf{K}_1$.

Step 2. H-level wavelet decomposition is performed on each selected frame $\mathbf{f}_{i(k)}$ to get its coarse signal $\mathbf{A}^H_{i(k)}$ and detail signals $\mathbf{D}^H_{i(k)}, \mathbf{D}^{H-1}_{i(k)}, \ldots, \mathbf{D}^1_{i(k)}$. And, to take the advantage of low-frequency coefficient which has a high-energy value and is robust against various signal processing manipulations the DCT is only performed on $\mathbf{A}^H_{i(k)}$as follows:

$$\mathbf{A}^{HC}_{i(k)} = \mathrm{DCT}(\mathbf{A}^H_{i(k)}) = \left\{a^{HC}_{i(k)}(n) \mid n = 0, \ldots, \frac{L_f}{2^H} - 1\right\}. \quad (11)$$

Step 3. For each $\mathbf{A}^{HC}_{i(k)}$, calculate its fourth-order cumulant, denoted as $\mathbf{C}_{i(k)}$,

$$\mathbf{C}_{i(k)} = \left\{c_{i(k)}(n) \mid n = 0, \ldots, \frac{L_f}{2^{H-1}}\right\}. \quad (12)$$

Then, the elements in $\mathbf{C}_{i(k)}$ are rearranged in order of decreasing absolute value and the first P ($P = (M \times N)/T$) elements are selected to generate a new sequence $\mathbf{D}_{i(k)}$as follows:

$$\mathbf{D}_{i(k)} = \{d_{i(k)}(p) \mid p = 0, \ldots, P-1\}. \quad (13)$$

And the index of $d_{i(k)}(p)$ in $\mathbf{C}_{i(k)}$ denoted as $\mathbf{I}_2$,

$$\mathbf{I}_2 = \left\{i_{i(k)}(p) \mid i_{i(k)}(p) \in \left\{0, \ldots, \frac{L_f}{2^{H-1}}\right\}, p = 0, \ldots, P-1\right\}, \quad (14)$$

is saved as the second secret key $\mathbf{K}_2$.

Step 4. A binary pattern, denoted as $\mathbf{B}_{i(k)}$,

$$\mathbf{B}_{i(k)} = \{b_{i(k)}(p) \mid p = 0, \dots, P-1\}, \tag{15}$$

is generated with (16) as follows:

$$b_{i(k)}(p) = \begin{cases} 1, & \text{if } d_{i(k)}(p) \geq 0, \\ 0, & \text{otherwise.} \end{cases} \tag{16}$$

And, the watermark detection key $\mathbf{K}_3 = \{K_{i(k)}(p) \mid k = 0, \dots, T-1, p = 0, \dots, P-1\}$ is obtained by performing XOR operation between $\mathbf{B}_{i(k)}$ and the binary watermark $\mathbf{W}$ as follows:

$$\begin{gathered} K_{i(k)}(p) = b_{i(k)}(p) \oplus w(i,j), \\ k = 0, \dots, T-1, \quad p = 0, \dots, P-1, \\ i = \text{floor}\left(\frac{k \times P + p}{N}\right), \quad j = \text{mod}\left(\frac{k \times P + p}{N}\right). \end{gathered} \tag{17}$$

Finally, the host audio signal, the secret keys $(\mathbf{K}_1, \mathbf{K}_2, \mathbf{K}_3)$, and the corresponding digital timestamp are registered or associated with an authentication center for copyright demonstration.

3.3. Extraction process

The watermark recovery procedure can be carried out without the host audio as follows.

At first, the test audio signal $\widetilde{\mathbf{A}} = \{\widetilde{a}(i) \mid i = 0, \dots, L_A - 1\}$ is divided into L frames $\mathbf{F} = \{\mathbf{f}_i \mid i = 0, \dots, L-1\}$, from which T frames, denoted as $\mathbf{f}_{i(k)}, k = 0, \dots, T-1$, are selected with $\mathbf{K}_1$.

Next, H-level wavelet decomposition is performed on each selected frame to get its coarse signal $\widetilde{\mathbf{A}}^{H}_{i(k)}$, on which DCT is performed to get $\widetilde{A}^{HC}_{i(k)}$.

Next, for each $\widetilde{\mathbf{A}}^{HC}_{i(k)}$, calculate its fourth-order cumulant $\widetilde{\mathbf{C}}_{i(k)}$, from which P elements are selected with secret key $\mathbf{K}_2$ to get a new sequence $\widetilde{\mathbf{D}}_{i(k)}$:

$$\widetilde{\mathbf{D}}_{i(k)} = \left\{\widetilde{d}_{i(k)}(p) \mid k = 0, \dots, T-1, p = 0, \dots, P-1\right\}. \tag{18}$$

Then, the estimated binary pattern $\widetilde{\mathbf{B}}_{i(k)}$

$$\widetilde{\mathbf{B}}_{i(k)} = \left\{\widetilde{b}_{i(k)}(p) \mid k = 0, \dots, T-1, p = 0, \dots, P-1\right\} \tag{19}$$

is generated as follow:

$$\widetilde{b}_{i(k)}(p) = \begin{cases} 1, & \text{if } \widetilde{d}_{i(k)}(p) \geq 0, \\ 0, & \text{otherwise.} \end{cases} \tag{20}$$

Finally, XOR operation is performed between the estimated binary pattern and the watermark detection key $\mathbf{K}_3$ to obtain the estimated binary image watermark $\widetilde{\mathbf{W}}$.

4. SIMULATION RESULTS AND DISCUSSIONS

4.1. Simulation results

To demonstrate the feasibility of our scheme, the performance test, detection reliability test, and robustness test were illustrated for the proposed watermarking algorithm, and the proposed watermark detection results were compared with that of scheme [3] against various audio signal processing manipulations and malicious attacks provided by Stirmark for Audio v0.2 [15]. All of the audio signals used in this test were audio with 16 bits/sample, 44.1 KHz sample rate, and 28.73s long. The watermark to be embedded was a visually recognizable binary image of size 64×64 . The Haar wavelet basis was used, and three-level wavelet decomposition was performed. The frame length was fixed at 512 samples and in each selected frame 4 bits were embedded.

We used the signal-to-noise ratio (SNR) (21) to evaluate the quality comparison between the attacked audio and original audio:

$$\text{SNR}(\mathbf{A}, \widetilde{\mathbf{A}}) = 10 \log_{10}\left\{\frac{\sum_{i=0}^{L_A-1} a^2(i)}{\sum_{i=0}^{L_A-1}[a(i) - \widetilde{a}(i)]^2}\right\}. \tag{21}$$

The normalized cross-correlation (NC) (22) was adopted to appraise the similarity between the estimated watermark and the original one:

$$\text{NC}(\mathbf{W}, \widetilde{\mathbf{W}}) = \frac{\sum_{i=0}^{M-1}\sum_{j=0}^{N-1} w(i,j)\widetilde{w}(i,j)}{\sqrt{\sum_{i=0}^{M-1}\sum_{j=0}^{N-1} w^2(i,j)}\sqrt{\sum_{i=0}^{M-1}\sum_{j=0}^{N-1} \widetilde{w}^2(i,j)}}. \tag{22}$$

And, the bit error rate (BER) (23) was employed to measure the robustness of our algorithm,

$$\text{BER} = \frac{B}{M \times N} \times 100\%, \tag{23}$$

where B is the number of erroneously extracted bits.

(1) *Performance test:* a plot of the host audio signal is shown in Figure 3(a). The original watermark image and the extracted watermark image are displayed in Figures 3(b) and 3(c) (NC = 1), respectively.

(2) *Imperceptibility:* one of the main requirements of audio watermarking techniques is inaudibility of the embedded watermark. For the proposed scheme, this requirement is naturally achieved because the watermark is embedded into the secret key but not the host audio signal itself. Actually, the watermarked audio is the identical to the original one.

(3) *Detection reliability:* to examine whether the proposed technique has the undesired property to extract the watermark $\mathbf{W}$ from the audio signals with no embedded watermark. More specifically, we attempt to extract $\mathbf{W}$ from the nonwatermarked audio signals using the same keys needed to extract $\mathbf{W}$ from the host audio signal. The waveforms of the original host audio signal (Figure 4(a)) and another three pieces of audio signals (Figures 4(b)–4(d)), and their corresponding extracted watermarks (Figures 4(e)–4(h)) are

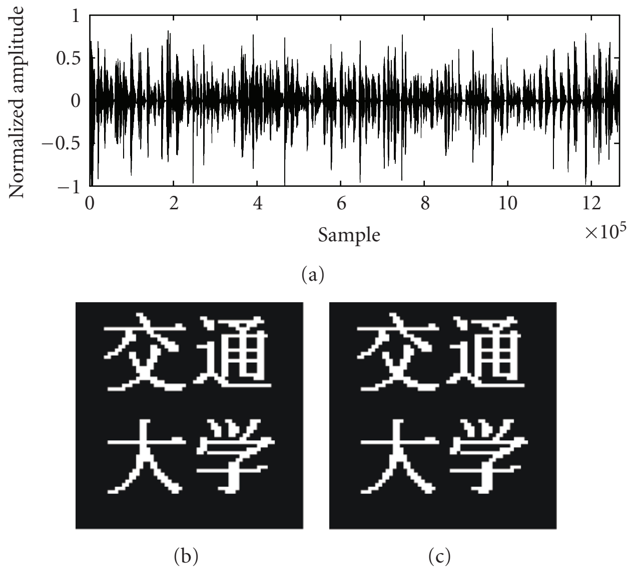


Figure 3: Watermark detection results. (a) Original host audio signal. (b) Original watermark. (c) Extracted watermark without being attacked (NC = 1).

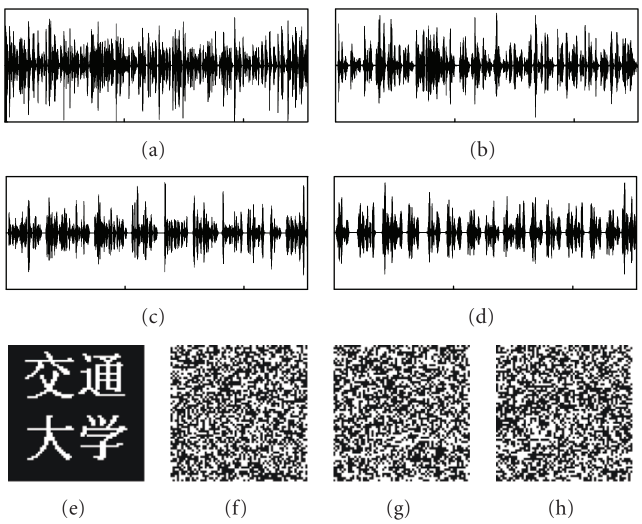


Figure 4: Audio signals and their extracted watermarks. (a) Original host audio. (b)–(d) Three audio signals without embedded watermarks. (e) Extracted watermark from (a). (f) Extracted watermark from (b). (g) Extracted watermark from (c). (h) Extracted watermark from (d).

shown in Figure 4. Furthermore, watermark detection results for 101 different audio signals (50 speech signals, 50 music signals, and the original host audio signal) are shown in Figure 5, the peek of which corresponds to the original host audio signal. It is clear that the proposed scheme detects correctly a watermark from the matched audio signal and keys, while avoiding false watermark detection from the unmatched audio signals.

(4) *Robustness:* another important requirement for watermarking techniques is robustness. The robustness of a watermarking algorithm measures its ability to correctly detect the watermark from the watermarked signal with nonmalicious and/or malicious attacks. In this paper, some commonly used audio signal processing manipulations, such as Mp3 compressing, requantizing, resampling, low-pass filtering, equalizing, amplitude amplifying, time delaying, echo-adding and noise-adding, and the malicious attacks provided by the practical audio watermarking evaluation tool Stirmark for Audio v0.2 [15] are utilized to estimate the robustness of the proposed scheme. The detection results including SNR, NC, and extracted watermark of the proposed scheme compared with those of scheme in [3] against various attacks are summarized in Table 1. And, the BER comparison of the proposed scheme and the scheme in [3] is shown in Figure 6.

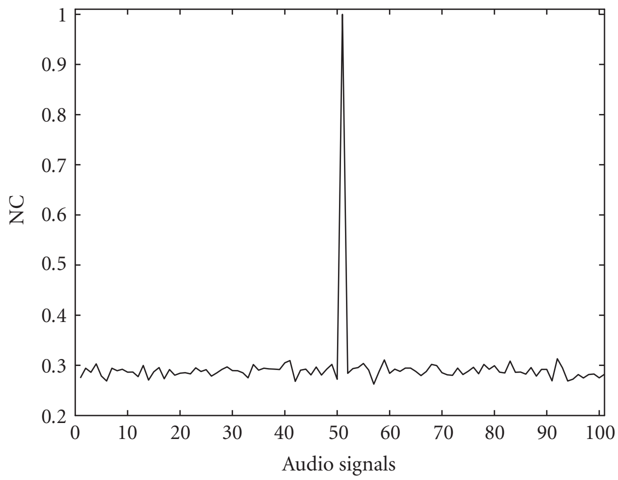


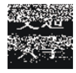
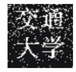
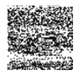
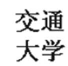
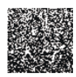
Figure 5: Detection reliability test result.

Experimental results show that our audio watermarking scheme not only introduces no distortion into the host audio, but also achieves great robustness against various attacks. The performance of it is better than that of the scheme in [3].

4.2. Discussions

From the experimental results, it can be seen that the proposed audio watermarking scheme possesses five essential properties of transparency, robustness, security, reliability, and blindness. It has transparency because it is lossless. For high-quality digital audio signal, for example, lossless is very important property. It is also robust. This is especially important as many available audio watermarking schemes are vulnerable to time-delaying and noise-addition attacks. It is secure. The security of the proposed technique is based on the host audio itself, the keys generated in watermark embedding stage, and the digital timestamp, which are registered in an authentication center. It is reliable because it can correctly extract watermark from the matched audio and keys, while avoiding false watermark estimation from the unmatched audio signals. It has blindness since the watermark recovery can be performed without the original audio. In practice, this is an essential property of the copyright protection scheme.

Table 1: Watermark detection results for various attacks.

No.	Attacks	SNR(our)	SNR([3])	NC(our)	NC([3])	Watermark(our)	Watermark([3])
a	MPEG layer 3 compression (48 Kbps)	+∞	17.70	1	1		
b	Requantization (8 ⟶ 16 ⟶ 8 bits/sample)	23.34	17.58	1	1		
c	Low-pass filtering (22.05 kHz)	76.54	17.70	1	1		
d	Equalization	11.71	10.75	0.97	0.43		
e	Addnoise-900	16.22	14.74	0.99	0.77		
f	Addbrumm-1100	12.30	13.46	0.98	0.94		
g	Addsinus	10.77	12.45	0.92	0.85		
h	Amplitude amplify (5 dB)	3.56	1.89	1	0.25		
i	Amplify	6.02	5.94	1	0.23		
j	Compressor	18.75	16.26	0.99	1		
k	Normalize	59.88	16.48	1	0.98		
l	Invert	−6.01	−6.06	0.997	0		
m	Real-reverse	29.74	16.96	1	0.998		
n	Zero-cross	20.27	15.96	0.97	0.37		
o	Delay (500 ms, 10%)	76.54	16.62	1	1		
p	Echo (100 ms, 10%)	26.93	17.70	1	1		
q	Smooth	25.08	17.87	0.999	0.88		
r	Stat1	20.24	15.22	0.94	0.50		
s	Stat2	32.04	16.94	1	1		
t	Resampling (44.1 ⟶ 22.05 ⟶ 44.1 kHz)	59.18	18.02	1	1		

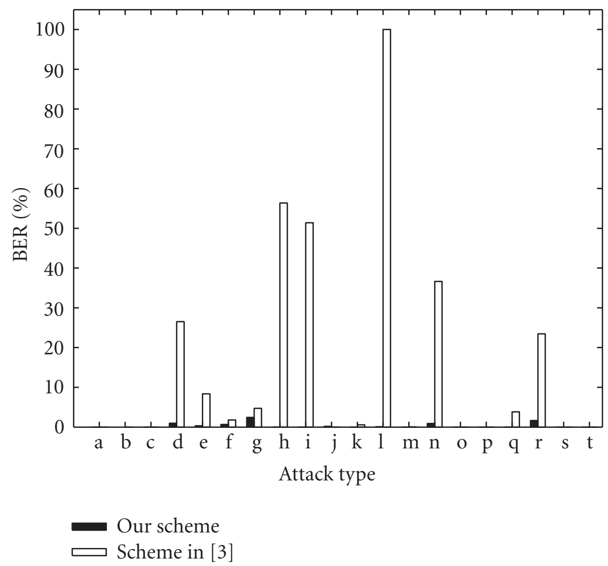


Figure 6: BER comparison between the proposed scheme and the scheme in [3] under various attacks.

5. CONCLUSIONS

Most of the currently available watermarking algorithms suffer in two points: one is the inevitable quality degradation introduced by the embedded watermark and the other is the inherent conflict between the imperceptibility and the robustness. To solve these problems, zero-watermarking technique is proposed. In this paper, an efficient and robust zero-watermarking algorithm for audio signal has been proposed. It achieves great detection reliability and robustness since it combines the multiresolution characteristic of DWT, the energy-compression characteristic of DCT, and the Gaussian noise suppression property of higher-order cumulant to extract essential characteristics from the host audio and uses them for watermark recovery. In addition, it guarantees the inaudibility because it hides the watermark into the secret key but not the host audio itself. Simulation results demonstrate the outstanding nature of our algorithm in terms of inaudibility, detection reliability, and robustness. Our future work will concentrate on introducing synchronization strategy into the proposed scheme to make it resist synchronization attacks such as random cropping and time-scale modification; on combining the proposed scheme with the available low-bit-rate audio coding standards to make it more fit for practical applications; and on embedding multiple watermarks into the same host audio to provide dual protection for it.

REFERENCES

[1] L. Wei, Y. Yi-Qun, L. Xiao-Qiang, X. Xiang-Yang, and L. Pei-Zhong, "Overview of digital audio watermarking," *Journal of Communication*, vol. 26, no. 2, pp. 100–111, 2005.

[2] X.-Y. Wang, Y.-R. Cui, H.-Y. Yang, and H. Zhao, "A new content-based digital audio watermarking algorithm for copyright protection," in *Proceedings of the 3rd International Conference on Information Security (SEC '04)*, vol. 85, pp. 62–68, Palo Alto, Calif, USA, 2004.

[3] X.-Y. Wang and H. Zhao, "A novel synchronization invariant audio watermarking scheme based on DWT and DCT," *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4835–4840, 2006.

[4] X.-Y. Wang, W. Qi, and P. Niu, "A new adaptive digital audio watermarking based on support vector regression," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 8, pp. 2270–2277, 2007.

[5] C.-C. Chang, K.-F. Hwang, and M.-S. Hwang, "A block based digital watermarks for copy protection of images," in *Proceedings of the 5th Asia-Pacific Conference on Communications (APCC '99)*, vol. 2, pp. 977–980, Beijin, China, October 1999.

[6] H.-C. Huang, F.-H. Wang, and J.-S. Pan, "Efficient and robust watermarking algorithm with vector quantisation," *Electronics Letters*, vol. 37, no. 13, pp. 826–828, 2001.

[7] T. Sun, W. Quan, and S.-X. Wang, "Zero-watermark watermarking for image authentication," in *Proceedings of the Signal and Image Processing*, pp. 503–508, Kauai, Hawaii, USA, August 2002.

[8] G. Horng, C. Chen, B. Ceng, and T. Chen, "Neural network based robust lossless copyright protection technique," http://www.csie.cyut.edu.tw/TAAI2002/TAAI2002PDF/Parallel.

[9] Q. Wen, T.-F. Sun, and S.-X. Wang, "Concept and application of zero-watermark," *Tien Tzu Hsueh Pao/Acta Electronica Sinica*, vol. 31, no. 2, pp. 214–216, 2003.

[10] S. Yang, C. Li, F. Sun, and Y. Sun, "Study on the method of image non-watermark in DWT domain," *Chinese Journal Image Graphics*, vol. 8A, no. 6, pp. 664–669, 2003.

[11] D. Charalampidis, "Improved robust VQ-based watermarking," *Electronics Letters*, vol. 41, no. 23, pp. 21–22, 2005.

[12] J. Sang, X. Liao, and M. S. Alam, "Neural-network-based zero-watermark scheme for digital images," *Optical Engineering*, vol. 45, no. 9, 2006.

[13] C. Hanqiang, X. Hua, L. Xutao, L. Miao, Y. Sheng, and W. Fang, "A zero-watermarking algorithm based on DWT and chaotic modulation," in *Independent Component Analyses*, vol. 6247 of *Proceedings of SPIE*, pp. 1–9, Orlando, Fla, USA, 2006.

[14] L. Jing and F. Liu, "Double zero-watermarks scheme utilizing scale invariant feature transform and log-polar mapping," in *Proceedings of the IEEE International Conference on Multimedia and Expo*, pp. 2118–2121, Las Vegas, Nev, USA, February 2007.

[15] "StirMark benchmark for audio," 2005, http://amsl-smb.cs.uni-magdeburg.de/.

[16] D. R. Green, "The utility of higher-order statistics in gaussian noise suppression," US Government Authored or Collected Report, Naval Postgraduate School, Memory, Calif, USA, 2003.

[17] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signla processing and system theory: theoretical resutlts and some applications," *Proceedings of IEEE*, vol. 79, no. 3, 1991.

[18] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1673–1687, 1997.