

Research Article

An Effective Watermarking Method Based on Energy Averaging in Audio Signals

S. E. Tsai¹ and S. M. Yang²

¹Department of Computer Science and Information Engineering, Chang Jung Christian University, Tainan 701, Taiwan

²Department of Aeronautics and Astronautics, National Cheng Kung University, Tainan 701, Taiwan

Correspondence should be addressed to S. E. Tsai; sean@mail.cjcu.edu.tw

Received 19 June 2017; Revised 26 February 2018; Accepted 18 March 2018; Published 25 June 2018

Academic Editor: Elisa Francomano

Copyright © 2018 S. E. Tsai and S. M. Yang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Methods based on discrete cosine transform (DCT) have been proposed for digital watermarking of audio signals; however, the watermark is often vulnerable to data compression and signal processing. This paper presents an effective audio watermarking method by energy averaging of DCT coefficients such that an audio signal with watermark is robust to data processing. The method is to divide an audio signal into segments by three parameters defining the segment length, the segment sequence of watermark location, and the frequency range of DCT coefficients for watermark location. An error correcting code is also integrated to improve audio signal quality after watermarking. Experimental results show that the method is robust to data compression and many other kinds of signal processing. No original signal is required for decoding the watermark. Comparison of watermarking performance with a recent work validates that the watermarking method has better audio quality and higher robustness.

1. Introduction

Audio watermarking is currently at the forefront of technology development to detect illegal reproduction and redistribution of audio recordings. Because the human auditory system (HAS) is more sensitive than the human visual system, audio watermarking is more challenging than visual watermarking. A reliable digital audio watermarking shall have imperceptibility, data capacity, and robustness [1]. The watermark must be inaudible within the host audio to maintain audio quality. The watermark data capacity is the information embedded or hidden in the host audio without perceptible distortion. The watermark robustness is that the watermark must remain intact or identifiable through signal processing such as compression, time-scaling, filtering, and resampling performed on the watermarked audio.

Many watermarking methods for embedding digital patterns into audio signals in time domain or frequency domain have been proposed in open literature. Robert and Picard [2] developed a masking model to identify the location and strength of a watermark in audio watermarking. Ko et al.

[3] proposed a time-spread echo method by pseudonoise sequence for audio watermarking. Chen and Wu [4] also presented an echo hiding scheme to minimize echo amplitude in audio signals. The above time-domain methods aimed at exploiting the insensitivity of human ears to very short-delayed echo, but the watermarks suffer from poor immunity to processing such as channel noises, resampling, and filtering. Lie and Chang [5] proposed an embedding method based on the relative energy relations between three consecutive sample sections in the time domain. The method embedded one information bit in every three sections according to the relation between the energy difference.

By comparison, frequency domain approaches by fast Fourier transform (FFT), digital wavelet transform (DWT), and discrete cosine transform (DCT) are attractive alternatives, because the phase or amplitude of transform domain coefficients can be modified to carry desired watermark information. Takahashi et al. [6] proposed a watermarking method based on phase modulation, but the method suffered from low robustness in signal processing. Fallahpour and Megías [7] embedded the watermark in FFT domain to exploit the

translation-invariant property of the Fourier transform such that the distortions in time domain can be reduced. Another method was then proposed to embed the watermark in the lowest DWT coefficients by energy proportion to improve robustness; however, the watermark is limited in size due to audio quality concerns [8]. In audio signal, the DCT coefficients are vulnerable to operations such as MPEG-1 Layer-3 (MP3) compression, leading to poor audio quality. Some then considered audio watermarking in both DWT and DCT domains by singular value decomposition (SVD) to increase signal robustness [9–11], but the high computation load is of significant concern.

There are many audio watermarking methods using spread spectrum, especially in speech watermarking. Malik et al. [12] presented a watermarking method in frequency-selective spread spectrum by exploiting the features in HAS. Other spread spectrum audio watermarking schemes with empirical mode decomposition (EMD) were proposed for HAS, but the watermarking robustness for filtering attacks remains a challenge [13, 14]. Tiwari and Jain [15] proposed a method by measuring the masking threshold in spread spectrum, and Kang et al. [16] also considered a watermark of embedding synchronization codes to improve robustness. But the embedding capacity of both is again relatively limited. Nematollahi et al. [17] presented a speech watermarking to embed watermark into line spectral frequency by least significant bit. Sarreshtedari and Akhaee [18] also applied self-embedding speech signals to protect against channel coding, but the original speech signal is required to extract the watermark.

A fast DCT algorithm has recently been developed to achieve higher computation efficiency [19–21]. By using the algorithm, a watermarking method based on energy averaging of DCT coefficients is presented in this paper for digital audio signals. The watermark, instead of being directly embedded in the DCT coefficients, is converted into binary bits and applied to tune the energy of audio signal in frequency domain, thus making it difficult, if not impossible,

to detect the locations of watermark bits. The audio watermarking is therefore highly undetectable and imperceptible. In addition, an error correcting code is integrated to enhance the quality of retrieved watermark and to resist common signal processing such as cropping, time shifting, filtering, compression, and resampling. The paper is organized as follows. Section 2 describes the embedding algorithm including the fast DCT algorithm, signal segmentation, and energy averaging. Section 3 introduces the error correcting code and the bit error rate. The experimental results are in Section 4. Finally, Section 5 summarizes the effectiveness and efficiency of the proposed audio watermarking method.

2. Signal Segmentation and Energy Averaging

The basic structure of an audio encoder and decoder in MPEG is shown in Figure 1. The 8-point DCT and a power-law quantization matrix are used to transform an audio signal into a bit stream. Consider a spatial audio signal, $s = [s(0), s(1), \dots, s(7)]^T$ being transformed into the frequency domain, $f = [f(0), f(1), \dots, f(7)]^T$, where

$$f(k) = \frac{1}{2} C(k) \sum_{i=0}^7 s(i) \cos\left(\frac{(2i+1)k\pi}{16}\right) \quad (1)$$

with $C(0) = 1/\sqrt{2}$, and $C(k)=1$, otherwise. Therefore $f(2)$ and $f(6)$ can be evaluated by

$$\begin{aligned} & \begin{bmatrix} f(2) \\ f(6) \end{bmatrix} \\ &= A \begin{bmatrix} (b-d)(s(0)-s(3)-s(4)+s(7)) \\ d(s(0)+s(1)-s(2)-s(3)-s(4)-s(5)+s(6)+s(7)) \\ -(b-d)(s(1)-(2)-s(5)+s(6)) \end{bmatrix} \end{aligned} \quad (2)$$

where $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$. $f(1)$, $f(3)$, $f(5)$, and $f(7)$ become

$$\begin{bmatrix} f(7) \\ f(5) \\ -f(1) \\ f(3) \end{bmatrix} = A^* \begin{bmatrix} A \begin{bmatrix} (e+a-f+c)(s(0)-s(7)) \\ (f-c)(s(0)-s(7)+s(6)-s(1)) \\ (a-e-f+c)(s(6)-s(1)) \\ (-a-c)(s(0)-s(7)+s(3)-s(4)) \end{bmatrix} \\ c(s(0)-s(7)+s(3)-s(4)+s(6)-s(1)+s(2)-s(5)) \\ A \begin{bmatrix} (e-c)(s(6)-s(1)+s(2)-s(5)) \\ (e-a-f-c)(s(3)-s(4)) \\ (f+c)(s(3)-s(4)+s(2)-s(5)) \\ (a+e-f-c)(s(2)-s(5)) \end{bmatrix} \end{bmatrix} \quad (3)$$

where $a = 0.488$, $b = 0.463$, $c = 0.416$, $d = 0.192$, $e = 0.098$, $f = 0.278$, and $A^* = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$ with I the 2×2 identity matrix. It has been shown that the fast DCT (FDCT) algorithm needs only 12 multiplications [20].

In digital audio watermarking, because human ears are extremely sensitive to the tone of a sound, it is futile to embed a watermark by adjusting the frequency. However, the energy level of an audio signal is vague to HAS, and

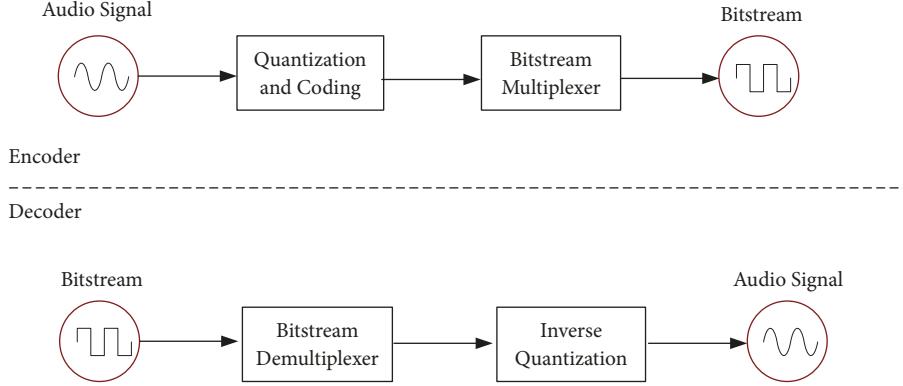
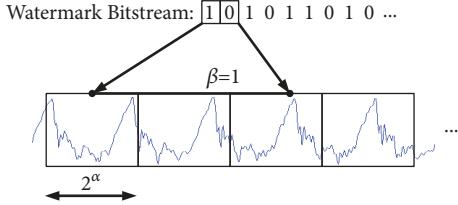


FIGURE 1: The structure of encoding/decoding audio signal in MPEG.

FIGURE 2: Illustration of an audio signal with segment length 2^α and $\beta = 1$.

that opens a window for implanting a digital watermark in an audio signal. An energy averaging method is proposed to incorporate watermark bits (0 and 1) in the host audio signal. Consider that a host signal is characterized by three parameters: α , β , and γ . They represent the segment, the segment sequence, and the frequency range for marking the watermark, respectively. The three adjustable parameters (α , β , and γ) are the “watermark keys” that constitute a security mechanism in guarding against unauthorized decoding of a watermarked signal. Parameter α determines the 2^α length in segmenting the host signal; for example, if $\alpha = 8$, then each segment of an audio signal contains 256 samples. Parameter β defines the sequence of segments selected to mark a watermark bit. If $\beta = 0$, it indicates every segment is selected; if $\beta = 1$, then every other segment is selected; i.e., segments 1, 3, 5, etc. are selected. If $\beta = 2$, then selected segments 1, 5, 9, etc. are selected. Parameter γ identifies the frequency range to identify the watermark; for example, $\gamma = 5$ represents that the watermark bit information is at the 32th (2^5) frequency coefficient of the selected segments.

Figure 2 illustrates the first two parameters, where a host audio signal is divided into segments, each of length 2^α (except the last segment) with every other segment selected to mark a watermark bit when $\beta = 1$. After each segment is transformed into frequency domain by the fast DCT algorithm in (2) and (3), parameter γ identifies the frequency range for the watermark bits. There are some analyses and suggestions of the three watermark keys:

- (1) Adjusting α means changing the block size; thus the larger α is, the more complex it is. However, the larger

the block size, the higher the precision. Moreover, the adjustment of block size would also affect the frequency location of γ .

- (2) Adjusting β means changing the coefficient block to embed watermark; thus it affects the system’s payload. The more the watermark bits embed, the lower the SNR.
- (3) Adjusting γ means changing the frequency location to embed watermark; thus the higher the frequency, the better the fidelity. However, according to the characteristics of the frequency coefficients, the lower the frequency, the better the robustness.

Figure 3 illustrates the proposed audio watermarking method by using three keys (α , β , and γ). Because the energy of an audio signal is concentrated in the low frequency coefficients after DCT, embedding watermark bits directly in the DCT coefficients may have good robustness but the audio quality is unacceptable. If DCT coefficients were replaced by a watermark bit, the quality of the encrypted signal will be seriously degraded. To maintain audio quality, a watermark shall not directly replace the low frequency coefficients. The coefficients, equivalent to the energy level of the segment, shall be modified by energy averaging. The location in the segment of energy averaging then marks a watermark bit.

Consider a host signal with each segment further divided by three subsegments of the same length L . The energy of each subsegment is

$$E = \sum_{i=2^\gamma}^{L-1} f(i)^2 \quad (4)$$

where $f(i)$ is the DCT coefficient of subsegments and i is the coefficient length. An effective watermarking in digital audio signal shall provide a balance between audio quality and robustness. The human ear is sensitive to the frequency changes but somewhat inert to amplitude change of an audio signal. Let E_{\max} , E_{med} , and E_{\min} be the maximum, median, and minimum energy of the three subsegments. In order to mark the watermark bits, the subsegment energy can be tuned by $E_{\max}(1 - d)$, $E_{\text{med}}(1 + d)$, and $E_{\min}(1 - d)$, where d is the threshold. In this paper, we set the value of d at five

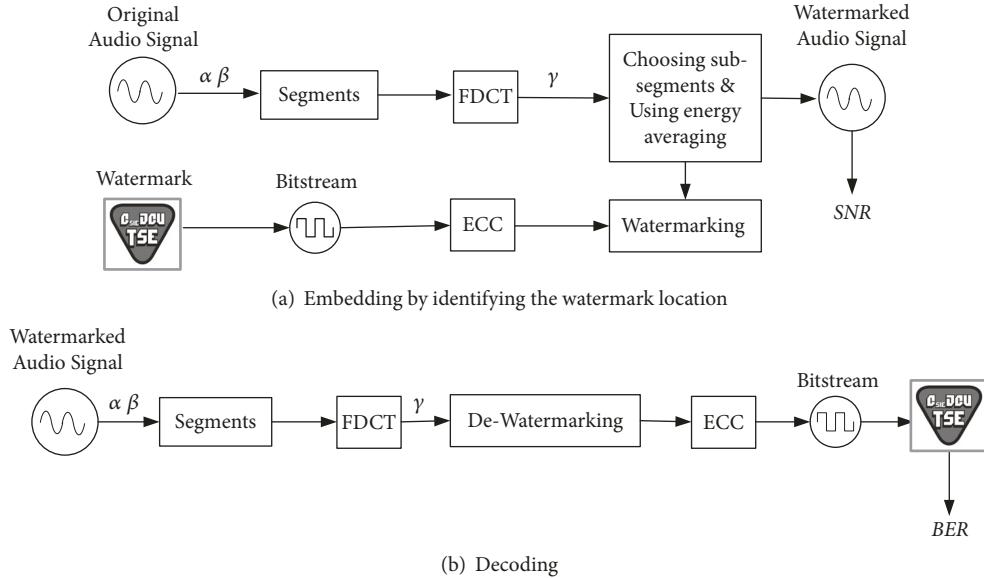


FIGURE 3: The watermarking method integrated with an error correcting code (ECC) with the signal-to-noise ratio (SNR) and bit error rate (BER) to evaluate the watermarking performance.

percent ($d = 5\%$). Other values of d are possible, where d is an adjustable parameter and the robustness of watermark and the host audio quality are related to the value of d . Define

$$E_1 = (E_{\max} - E_{\text{med}}) - d(E_{\max} + E_{\text{med}}); \quad (5a)$$

$$E_2 = (E_{\text{med}} - E_{\min}) - d(E_{\text{med}} + E_{\min}). \quad (5b)$$

Energy averaging on $(E_1 - E_2)$, if necessary, is to match the watermark bit (0 or 1) such that E_1 and E_2 are always close to each other and the modification of the signal energy will be minimal and undetectable to HAS. If and only if the watermark bit is 1, then conduct energy averaging to $(E_1 - E_2) \geq 0$ if necessary. Conversely, when $(E_1 - E_2) < 0$, one needs to decrease E_{med} and E_{\min} and increase E_{\max} by the threshold d until $(E_1 - E_2) \geq 0$. Similarly, if the watermark bit is 0, one needs to ensure $(E_1 - E_2) < 0$, by energy averaging if necessary.

It should be noted that the original audio signal is not required to decode the watermark; only the parameters α , β , and γ and the subsegment length L are needed. α and β are to locate the segments containing the watermark information in an audio signal; then γ and L are to locate the three adjacent subsegments. The watermarking method based on energy averaging is shown in Figure 4. All watermark bits in the modified audio signal can be obtained by inverse DCT as illustrated in Figure 3. After all the segments are processed, the decoded watermark bit stream can be transformed to recover watermark information.

3. Error Correcting and Performance Measure

Error correcting code (ECC), an integral part in many digital communication systems, is to protect the watermarked signal against transmission and storage errors. Cyclic code by shifting a codeword cyclically to another codeword is applied to obtain higher decoding accuracy. It is generated by a shift

register and thus systematic for providing proper checking bits. Encoding and syndrome calculation of a cyclic code can thus be implemented easily by employing simple shift registers with feedback connections. Its inherent algebraic structure also allows efficient decoding. The effectiveness of audio watermarking method is evaluated by the audio quality and robustness of a watermarked signal against operations such as compression, filtering, cropping, resampling, additive noise, and echo. Audio quality refers to the imperceptibility of signal after energy averaging. It is vital that the watermark is undetectable in all applications and the quality of host signal is not perceivably distorted. Two measures are used to evaluate the watermarking performance. The first is the signal-to-noise ratio (SNR) calculated by the original signal amplitude over the noise amplitude of the watermarked signal

$$\text{SNR} = 10 \log \left(\frac{\sum_{i=0}^{n-1} s^2(x)}{\sum_{i=0}^{n-1} [s(x) - s_w(x)]^2} \right) \quad (6)$$

where $s(x)$ is the input sequence of the original signal, $s_w(x)$ is the output sequence of the watermarked signal after energy averaging, and n is the signal length. An SNR higher than 30 dB is good because it is beyond the human capability to detect the difference between the two (original versus watermarked) signals [22]. The second is the bit error rate (BER) defined as the number of error bits after decoding the watermark over the total number of watermark bits,

$$\text{BER} = \frac{\text{Bit Errors}}{\text{Total Watermark Bits}}. \quad (7)$$

It is generally acceptable to have a BER no more than 0.2 before any attack on the audio signal with watermarking [22].

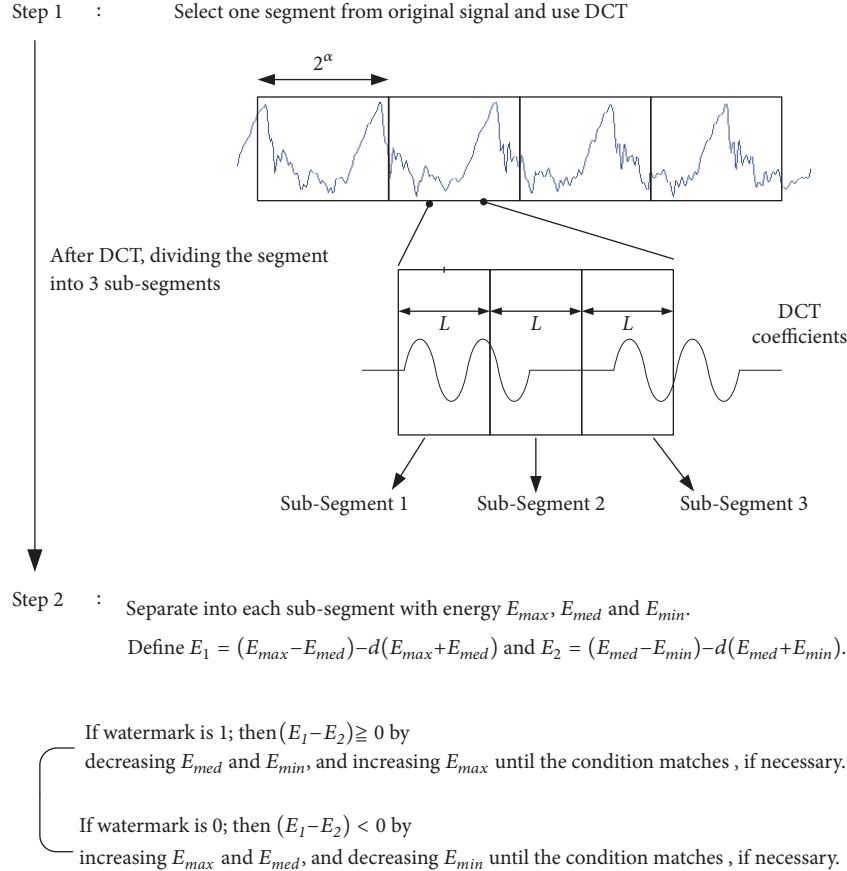


FIGURE 4: The energy averaging in audio signal watermarking.

4. Experimental Results

To ensure the proposed audio watermarking method is effective and robust, three audio signals, classical piano, vocal, and symphonic, are used to evaluate the performance. The audio signals are all in 16 bits, mono, 44.1 kHz sampling rate and 60 seconds in length. The input audio signal can be in the Windows PCM, .wav, or MP3 format. The watermark is an 8 bit of size 32×32 image.

Based on the energy of DCT coefficients, this technique provides a higher robustness. Table 1 shows the results of watermarked pop music with keys (9, 1, and 6) under the MP3 compression in the first mode (without ECC). The results indicate that the higher L is, the better BER is. However, with the increase of L, the SNR is decreased. Therefore, the balance of SNR and BER depends on the requirements of the application.

As for the parameter analyses of d, the robustness of watermark and the host audio quality are related to the value of d. Moreover, the parameter d is suggested by considering the data shift after common manipulation. The suggestion is listed as the Table 2. In this experiment, the keys are (9, 0, and 6), L = 50, and d = 1~ 15% for symphonic music. Thus, it is logical to keep d low, usually less than 10%, as proven during technique development. Consequently, the modification of the audio signal will be very low. This

TABLE 1: The effects of the subsegments length of averaging DCT coefficients to increase the robustness. (pop music).

L	SNR	Time (sec)	BER after MP3 compression
1	47.32	0.12	0.43
5	38.76	0.88	0.38
10	36.79	2.14	0.35
20	36.72	5.87	0.31
30	34.98	9.98	0.27
40	35.79	13.30	0.24
50	35.67	18.29	0.21
100	33.34	32.85	0.20
150	30.27	59.93	0.20

article suggests that d should be higher than 2% in order to maintain a robust level of watermark. From the results, d = 5% of the watermarking method makes the embedding time, SNR, and BER preferable. Therefore, 5% is suggested in the watermarking criteria. Any other percentage that yields good effectiveness is acceptable.

For a long audio signal, α and β should be large enough. Different α and β will change the number of selected segments to mark the watermark, and the number of selected segments should be higher than or equal to the total number

TABLE 2: The influence of the parameter in the watermarking criteria.

d	Watermark embedding time (s)	SNR	BER after MP3
1%	3.42	48.79	0.35
2%	3.11	42.40	0.24
5%	2.93	41.78	0.18
10%	2.79	37.96	0.17
15%	2.77	34.81	0.17

TABLE 3: The SNRs and BERs validating the performance of the watermarking method with different parameters on classical piano music.

(α, β, γ)	SNR (dB)	BER
(9, 0, 6)	43.27	0
(9, 0, 7)	54.16	0
(9, 1, 6)	45.76	0
(9, 1, 7)	56.32	0
(10, 0, 6)	36.52	0
(10, 0, 7)	44.64	0
(10, 1, 6)	39.67	0
(10, 1, 7)	47.81	0

of watermark bits for complete watermarking. When α and β are small, the number of selected segments may be larger than the number of watermark bits; the same watermark may be used more than once to the host audio signal. For the case of $d = 5\%$ and $L = 50$, the effects of different parameters α , β , and γ on watermarking performance are presented in Table 3. The watermarking method is shown to be excellent to classical piano music. The SNRs are higher than 36 dB and the BERs are all 0 for the eight sets of $(\alpha, \beta$, and γ) parameters, indicating superb quality and perfect decoding. Among all the tests, increasing the segment length from 512 ($\alpha = 9$) to 1024 ($\alpha = 10$) leads to about 6 to 10 dB decrease in SNR. Of the tests in Table 3, the SNR is increased by about 2 dB when β is from 0 to 1. For $\gamma = 6$ or 7, i.e., energy averaging up to the 64th or 128th coefficient, the SNR is increased by about 6 to 10 dB.

Table 4 shows the robustness of the proposed watermarking method against MP3 compression, low pass filtering, time-scaling, and resampling. For all signals subject to MP3 compression at 128 kbit/s, the BERs are lower than 0.21, indicating that the extracted watermarks remain satisfactory to verification. The BERs of the watermark are improved substantially to within 0.05 by integrating ECC. On pop music when the watermarked signals pass through a low pass filter of 4 kHz bandwidth, the BERs are lower than 0.26. This is the special case of ECC offering marginal help. Nevertheless, ECC is shown to be helpful to BER on all types of music.

The most challenging audio signal processing is time-scaling of preserving pitch and/or tempo. The watermarking method integrated with cyclic code is shown robust. With ECC, the BERs of the watermarked signal after time-scaling are lower than 0.22 and the watermarks are still identifiable.

Without ECC, the BERs at 0.31 would have been unacceptable. When resampling the watermarked signal (44.1 kHz to 22.05 kHz or to 11.025 kHz) and then upsampling back to 44.1 kHz, ECC can improve the BERs to within 0.13. From the above tests, the retrieved watermark remains acceptable.

The effectiveness of the watermarking method is further validated by comparing with that of Kang et al. [16]. For vocal and symphonic music of the same data rate, Table 5 shows that the watermarking method with ECC is superior. The audio watermarking method based on energy averaging is more effective with higher robustness. The results in Tables 3 and 4 illustrate that the watermarking method is robust against common audio signal manipulations. Integration with ECC is helpful to guard against malicious data attacks.

Compared to another method in Lie and Chang [5] that also used average concept, the information bits were embedded in every three sections according to the relation between the energy differences in the time domain. In their method, one section is 1024 samples and threshold is 0.05, similar to our study with $(\alpha, \beta, \gamma) = (10, 0, 7)$ and $d = 5\%$. The comparison between average concept in time domain and DCT domain is shown in Table 6. From the results, the average concept applied to DCT domain not only provides better audio signal quality (better SNR) but also has better robustness (better BER) in MP3 compression and decompression, low pass filtering, time-scaling, and resampling and requantization.

5. Conclusions

- (1) This paper presents an efficient and robust digital audio watermarking method based on energy averaging in discrete cosine transform. By averaging the energy of three adjacent subsegments, an audio signal can be encrypted and decoded effectively and efficiently. No original audio signal is required in decoding the watermark.
- (2) The audio watermarking method employs a set of three keys $(\alpha, \beta$, and γ) to provide security and robustness. Each defines the segment length, the sequence of the segments for watermarking, and the DCT coefficients of a segment for averaging. Error correcting codes (ECC) also integrate the watermarking method so as to increase decoding effectiveness. It is shown that the cyclic code can provide higher accuracy and better robustness of the decoded watermark. Implementation shows that the watermarking method can maintain audio signal quality and achieve smaller bit error rate (BER).
- (3) Verifications on classical, vocal, symphonic, and pop music illustrate that the audio watermarking method is robust against common audio signal manipulations. In the experimental results, d should be kept in the range of 2% ~10 % in order to keep audio quality and maintain a robust level of watermark. This article provides a set of recommended values ($L = 50$ and $d = 5\%$) to facilitate subsequent implementation; then it suggests that $\alpha = 6 \sim 11$, $\beta = 0 \sim 3$, and γ is less

TABLE 4: The low BER illustrating the effectiveness of the watermarking method with error correcting code (ECC) under signal processing attack.

Music Type	Classical piano music		Vocal music		Symphonic music		Pop music	
			Algorithm		w/o	w/cc	w/o	w/cc
Attack	w/o	w/cc	w/o	w/cc	w/o	w/cc	w/o	w/cc
MP3 compression (128 kbytes/s)	0.19	0.05	0.21	0.04	0.18	0.05	0.21	0.03
Low-pass filter (4 kHz)	0.06	0.06	0.22	0.21	0.09	0.08	0.26	0.26
Time-scaling (preserve pitch)	0.21	0.13	0.26	0.19	0.22	0.13	0.28	0.18
Time-scaling (preserve tempo)	0.13	0.07	0.18	0.12	0.14	0.07	0.17	0.11
Time-scaling (preserve both)	0.21	0.16	0.31	0.21	0.22	0.16	0.25	0.22
Resampling (44.1 kHz to 22.05 kHz)	0.12	0.06	0.04	0.02	0.13	0.07	0.05	0.03
Resampling (44.1 kHz to 11.025 kHz)	0.19	0.11	0.07	0.03	0.21	0.13	0.08	0.05

w/o: audio watermarking with energy averaging only.

w/cc: audio watermarking with energy averaging and ECC (cyclic code).

TABLE 5: The BER illustrating the robustness of this method compared to [15] under several signal processing attacks.

Music Type	Vocal music		Symphonic music	
	Method		Kang et al. [15]	This work
Attack	Kang et al. [15]	This work	Kang et al. [15]	This work
MP3 compression (128 kbytes/s)	0.16	0.04	0.14	0.05
MP3 compression (64 kbytes/s)	0.21	0.11	0.20	0.11
Low-pass filter (4 kHz)	0.16	0.09	0.15	0.08
Resampling (44.1 kHz to 22.05 kHz)	0.06	0.02	0.08	0.07
Resampling (44.1 kHz to 11.025 kHz)	0.09	0.03	0.11	0.13

TABLE 6: Comparison between energy concept applying in time domain [5] and DCT domain.

Music Type	Symphonic music	
	Domain	Lie and Chang [5] (Time domain)
Attack (BER)	The work (DCT domain)	Lie and Chang [5] (Time domain)
No Attack	SNR: 41.47 BER: 0	SNR: 34.26 BER: 0.02
MP3 Compression (128 kbytes/s)	0.05	0.29
MP3 Compression (64 kbytes/s)	0.11	0.45
Low Pass Filter (4KHz)	0.08	0.11
Time-Scaling (Preserve Tempo)	0.14	0.46
Resampling (44.1 kHz to 22.05 kHz)	0.08	0.12
Resampling (44.1 kHz to 11.025 kHz)	0.11	0.22

than α . The SNRs are higher than 36 dB and the BERs are all 0, indicating excellent audio quality and perfect watermark decoding. To guard against digital signal processing such as MP3 compressing, decompression, low pass filter, time-scaling, and resampling, integration with error correcting code is shown to achieve smaller bit error rate. The performance of the watermarking method is compared with a previous work [19, 20], and the results validate that the method is effective on all music types and robust to signal processing.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

- [1] G. Hua, J. Huang, Y. Q. Shi, J. Goh, and V. L. L. Thing, "Twenty years of digital audio watermarking - A comprehensive review," *Signal Processing*, vol. 128, pp. 222–242, 2016.
- [2] A. Robert and J. Picard, "On the use of masking models for image and audio watermarking," *IEEE Transactions on Multimedia*, vol. 7, no. 4, pp. 727–739, 2005.
- [3] B.-S. Ko, R. Nishimura, and Y. Suzuki, "Time-spread echo method for digital audio watermarking," *IEEE Transactions on Multimedia*, vol. 7, no. 2, pp. 212–221, 2005.
- [4] O. T.-C. Chen and W.-C. Wu, "Highly robust, secure, and perceptual-quality echo hiding scheme," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 16, no. 3, pp. 629–638, 2008.
- [5] W.-N. Lie and L.-C. Chang, "Robust and high-quality time-domain audio watermarking based on low-frequency amplitude modification," *IEEE Transactions on Multimedia*, vol. 8, no. 1, pp. 46–59, 2006.
- [6] A. Takahashi, R. Nishimura, and Y. Suzuki, "Multiple watermarks for stereo audio signals using phase-modulation techniques," *IEEE Transactions on Signal Processing*, vol. 53, no. 2, part 2, pp. 806–815, 2005.
- [7] M. Fallahpour and D. Megías, "High Capacity FFT-Based Audio Watermarking," in *Communications and Multimedia Security*, vol. 7025 of *Lecture Notes in Computer Science*, pp. 235–237, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.
- [8] S.-T. Chen, H.-N. Huang, C.-J. Chen, and G.-D. W, "Energy-proportion based scheme for audio watermarking," *IET Signal Processing*, vol. 4, no. 5, pp. 576–587, 2010.
- [9] A. Al-Haj, "An imperceptible and robust audio watermarking algorithm," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2014, no. 1, 2014.
- [10] H.-T. Hu, H.-H. Chou, C. Yu, and L.-Y. Hsu, "Incorporation of perceptually adaptive QIM with singular value decomposition for blind audio watermarking," *EURASIP Journal on Advances in Signal Processing*, vol. 2014, no. 1, 2014.
- [11] B. Y. Lei, I. Y. Soon, and Z. Li, "Blind and robust audio watermarking scheme based on SVD-DCT," *Signal Processing*, vol. 91, no. 8, pp. 1973–1984, 2011.
- [12] H. Malik, R. Ansari, and A. Khokhar, "Robust audio watermarking using frequency-selective spread spectrum," *IET Information Security*, vol. 2, no. 4, pp. 129–150, 2008.
- [13] K. Khaldi and A.-O. Boudraa, "Audio watermarking via EMD," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 21, no. 3, pp. 675–680, 2013.
- [14] N. El Hamdouni, A. Adib, S. D. Larbi, and M. Turki, "A blind digital audio watermarking scheme based on EMD and UIZA techniques," *Multimedia Tools and Applications*, vol. 64, no. 3, pp. 809–829, 2013.
- [15] A. Tiwari and L. Jain, "Digital Audio Watermarking using Frequency Masking Technique," *International Journal of Computer Applications*, vol. 126, no. 4, pp. 1–7, 2015.
- [16] X. Kang, R. Yang, and J. Huang, "Geometric invariant audio watermarking based on an LCM feature," *IEEE Transactions on Multimedia*, vol. 13, no. 2, pp. 181–190, 2011.
- [17] M. A. Nematollahi, C. Vorakulpipat, and H. Gamboa Rosales, "Semifragile Speech Watermarking Based on Least Significant Bit Replacement of Line Spectral Frequencies," *Mathematical Problems in Engineering*, vol. 2017, 2017.
- [18] S. Sarreshtedari, M. A. Akhaee, and A. Abbasfar, "A Watermarking Method for Digital Speech Self-Recovery," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 23, no. 11, pp. 1917–1925, 2015.
- [19] S. Tsai, Y. Lin, and S. Yang, "Effective watermarking based on energy balancing in audio signals," in *Proceedings of the 2012 International Conference on Anti-Counterfeiting, Security and Identification (2012 ASID)*, pp. 1–5, Taipei, Taiwan, August 2012.
- [20] S. E. Tsai and S. M. Yang, "A fast DCT algorithm for watermarking in digital signal processor," *Mathematical Problems in Engineering*, vol. 2017, Article ID 7401845, 7 pages, 2017.
- [21] S. E. Tsai, K. C. Liu, and S. M. Yang, "An Efficient Image Watermarking Method Based on Fast Discrete Cosine Transform Algorithm," *Mathematical Problems in Engineering*, vol. 2017, pp. 1–10, 2017.
- [22] M. A. Nematollahi, S. Al-Haddad, and F. Zarafshan, "Blind digital speech watermarking based on Eigen-value quantization in DWT," *Journal of King Saud University—Computer and Information Sciences*, vol. 27, no. 1, pp. 58–67, 2015.

