

A NEW DCT-BASED WATERMARKING METHOD FOR COPYRIGHT PROTECTION OF DIGITAL AUDIO

Pranab Kumar Dhar^{*1}, Mohammad Ibrahim Khan² and Saif Ahmad³

Department of CSE, Chittagong University of Engineering and Technology, Bangladesh

¹pranab_cse@yahoo.com

²muhammad_ikhancuet@yahoo.com

³saeef.cse05@yahoo.com

ABSTRACT

Digital watermarking plays an important role for protecting digital contents from unauthorized copying. This paper proposes a new audio watermarking method based on Discrete Cosine Transformation (DCT) for copyright protection. In our proposed watermarking method, the original audio is transformed into DCT domain. The absolute values of DCT coefficients are divided into an arbitrary number of segments and the energy of each segment is calculated. Watermarks are then embedded into the selected peaks of the highest energy segment. Watermarks are extracted by performing the inverse operation of watermark embedding process. Simulation results indicate that our proposed watermarking method is highly robust against various kinds of attacks such as noise addition, cropping, re-sampling, re-quantization, MP3 compression, and echo, and achieves similarity values ranging from 13 to 32. In addition, our proposed method shows SNR (signal-to-noise ratio) values ranging from 13 dB to 24 dB.

KEYWORDS

Copyright protection, Digital Watermarking, Discrete Cosine Transform, Sound Contents

1. INTRODUCTION

Recent improvement in computational world and the proliferation of the Internet have facilitated the production and distribution of unauthorized copies of copyrighted digital contents. Consequently, the area of effective and efficient copyright protection is a field of interest of the worldwide scientific communities. Digital audio watermarking has drawn extensive attention for copyright protection of audio data. A digital audio watermarking is a process of embedding watermarks into audio signal to show authenticity and ownership. Audio watermarking should meet the following requirements : (a) *Imperceptibility*: the digital watermark should not affect the quality of original audio signal after it is watermarked; (b) *Robustness*: the embedded watermark data should not be removed or eliminated by unauthorized distributors using common signal processing operations and attacks; (c) *Capacity*: capacity refers to the numbers of bits that can be embedded into the audio signal within a unit of time; (d) *Security*: security implies that the watermark can only be detectable by the authorized person. All these requirements are often contradictory with each other. However, it should satisfy the important properties such as imperceptibility and robustness.

In this paper, we propose a new watermarking method using Discrete Cosine transformation (DCT) for audio copyright protection. The watermarks are embedded into the selected prominent peaks of the highest segment of absolute DCT coefficients. Experimental results indicate that the proposed watermarking method provides strong robustness against several kinds of attacks such as noise addition, cropping, re-sampling, re-quantization, MP3

* Corresponding author

compression, and echo and achieves similarity values ranging from 13 to 32. In addition, our proposed method achieves SNR (signal-to-noise ratio) values ranging from 13 dB to 24 dB.

The rest of this paper is organized as follows. Section 2 provides a brief description of previous works related to audio watermarking. Section 3 introduces our proposed watermarking method including watermark embedding process and watermark detection process. Section 4 discusses the performance of our proposed method in terms of imperceptibility and robustness. Finally, section 5 concludes this paper.

2. RELATED RESEARCH

A significant number of watermarking techniques have been reported in recent years in order to create robust and imperceptible audio watermarks. Lie *et al.* [1] propose a method of embedding watermarks into audio signals in the time domain. The proposed algorithm exploits differential average-of-absolute-amplitude relations within each group of audio samples to represent one-bit information. It also utilizes the low-frequency amplitude modification technique to scale the amplitudes in selected sections of samples so that the time domain waveform envelope can be almost preserved. In [2], authors propose a blind audio watermarking system which embeds watermarks into audio signal in time domain. The strength of the audio signal modifications is limited by the necessity to produce an output signal for watermark detection. The watermark signal is generated using a key, and watermark insertion depends on the amplitude and frequency of audio signal that minimizes the audibility of the watermarked signal. Ling *et al.* [3] introduce a watermarking scheme based on non-uniform discrete Fourier transform (NDFT), in which the frequency points of embedding watermark are selected by the secret key. Zeng *et al.* [4] describe a blind watermarking method which embeds watermarks into DCT coefficients by utilizing quantization index modulation technique. In [5], the authors propose a watermarking system which embeds synchronization signals in time domain to resist against several attacks. Pooyan *et al.* [6] introduce an audio watermarking method which embeds watermarks in wavelet domain. The watermarked data is then encrypted and combined with a synchronization code and embedded into low frequency coefficients of the sound in wavelet domain. The magnitude of quantization step and embedding strength is adaptively determined according to the characteristics of human auditory system. Wang *et al.* [7] proposes a blind audio watermarking scheme using adaptive quantization against synchronization attack. In addition, the multiresolution characteristics of discrete wavelet transform (DWT) and the energy compression characteristics of discrete cosine transform (DCT) are combined in this scheme to improve the transparency of digital watermark. Watermark is then embedded into low frequency components by using adaptive quantization according to human auditory system. In [8], authors propose a watermarking system in cepstrum domain in which a pseudo-random sequence is used as a watermark. The watermark is then weighted in the cepstrum domain according to the distribution of cepstral coefficients and the frequency masking characteristics of human auditory system. Liu *et al.* [9] propose a blind watermarking method which takes the advantages of the attack-invariant feature of the cepstrum domain and the error-correction capability of BCH code to increase the robustness as well as imperceptibility of audio watermarking.

In Cox's method [10] watermarks are embedded into the highest m DCT coefficient of the whole sound excluding the DC component by the following equation:

$$v'_i = v_i(1 + \alpha x_i) \quad (1)$$

where, m is the length of the watermark sequence, v_i is a magnitude coefficient into which a watermark is embedded, x_i is a watermark to be inserted into v_i , α is a scaling factor, and v'_i is an adjusted magnitude coefficient. The watermark sequence is extracted by performing the inverse operation of (1) represented by the following equation:

$$x_i^* = \left(\frac{v_i^*}{v_i} - 1\right) / \alpha \quad (2)$$

3. PROPOSED WATERMARKING METHOD

In this section, we give an overview of our basic watermarking method which consists of watermark embedding process and watermark detection process. In this implementation, a watermark consists of a sequence of real numbers $\mathbf{X} = \{x_1, x_2, x_3, \dots, x_n\}$. We create a watermark where each value of x_i is chosen independently according to $N(0,1)$ where $N(\mu, \sigma^2)$ denotes a normal distribution with mean μ and variance σ^2 .

2.1. Watermark Embedding Process

The proposed watermark embedding process is shown in Figure 1. The embedding process is implemented in the following seven steps:

1. The original audio signal is transformed into DCT domain to calculate the DCT coefficients.

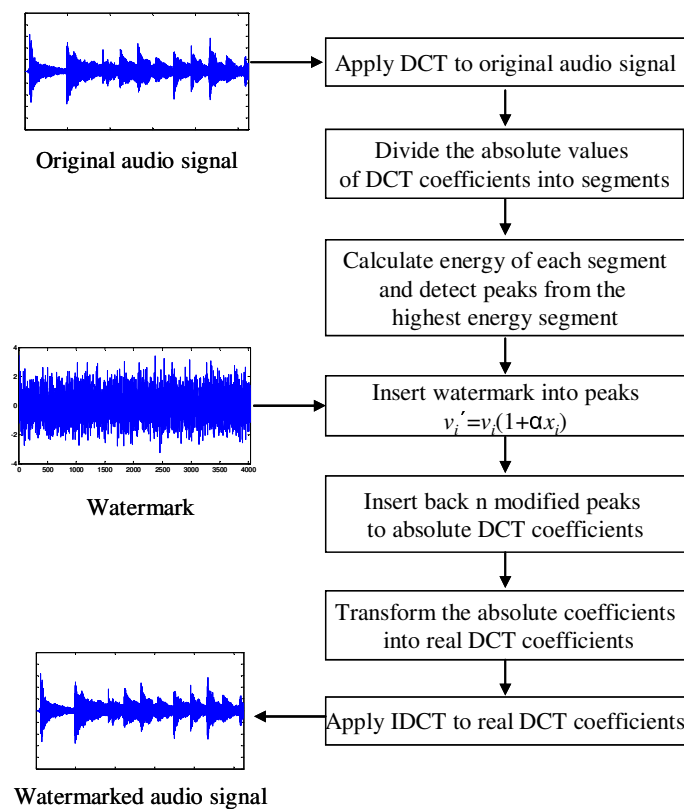


Figure 1. Watermark embedding process

2. Absolute values of the DCT coefficients are divided into an arbitrary number of segments.
3. Energy of each segment is then calculated. Mathematically, the energy is calculated using the following equation:

$$E = \sum_i |v(i)|^2 \quad (3)$$

4. Find the most prominent peaks from the highest energy segment using a peak detection algorithm.
5. The watermark is then embedded into the selected N peaks of the highest energy segment, where N is the length of watermark. This ensures that the watermark is located at the most significant perceptual components. When we insert the watermark X into V to obtain V', we specify a scaling parameter α , which determines the extent to which X alters V, shown in the equation (1) [10].
6. Insert back the modified peak into the highest energy segment of absolute DCT coefficients and transform these absolute coefficients to real DCT coefficients.
7. Apply an inverse Discrete Cosine Transformation (IDCT) to the real DCT coefficients to obtain the watermarked audio signal.

2.2. Watermark Detection Process

The proposed watermark detection process is shown in Figure 2.

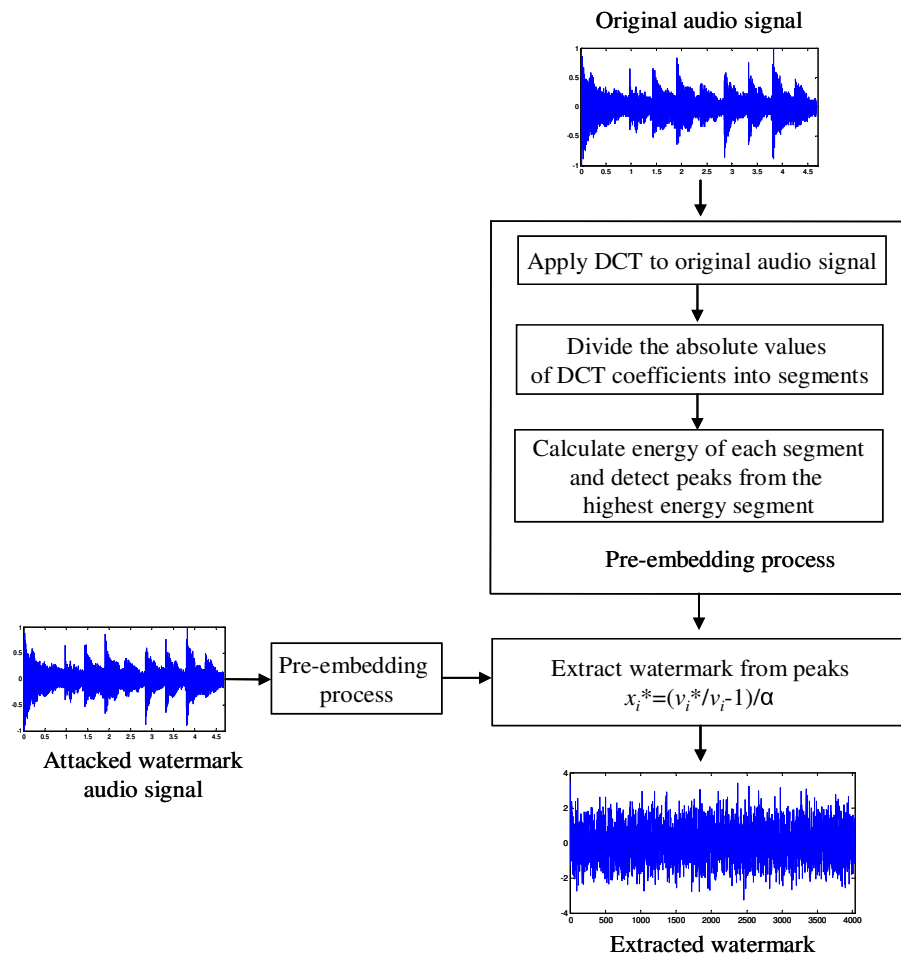


Figure 2. Watermark detection process

The detection process is implemented in the following three steps:

1. The attacked watermarked audio signal is transformed into DCT domain.
2. Extract the highest prominent peaks from the absolute DCT coefficients which are located at the same position in the embedding process above.
3. The watermark sequence $X^* = x_1^*, x_2^*, x_3^*, \dots, x_n^*$ is then extracted by performing the inverse operation of (1) represented by the equation (2)

4. SIMULATION RESULTS AND DISCUSSION

In this section, we evaluate the performance of our watermarking method for different types of 16 bit mono audio signals sampled at 44.1 KHz as shown in Figure 3. The sound files are: (a) Symphony: a clip of Mozart’s symphony play, (b) Gayageum: a clip of an instrumental solo, (c) Human Voice: a human voice providing recitation of a Bengali poem ‘Bidrohi’, and (d) Classical: a clip of a classical song. Each audio file contains 262,144 samples with duration of 6 seconds.

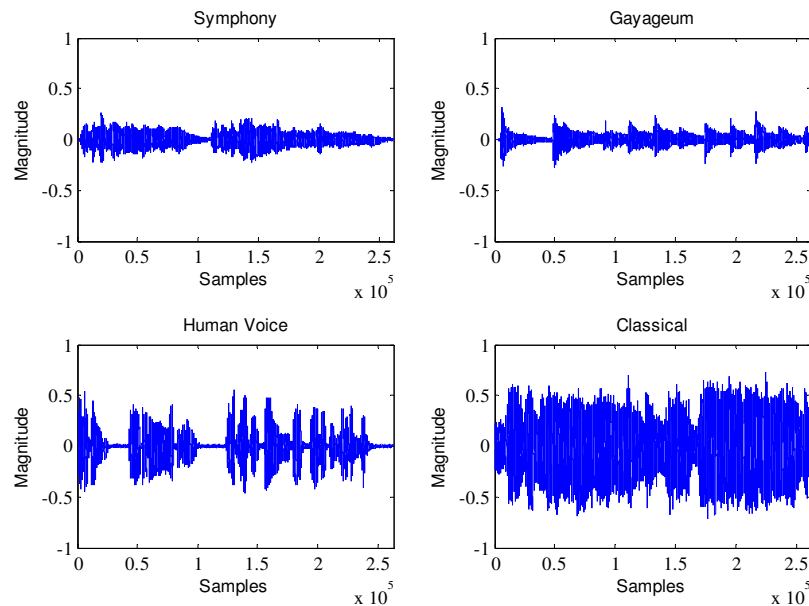


Figure 3. Original sounds used in this study

In order to evaluate the performance of the proposed watermarking method in terms of watermark detection, the correlation coefficient between the original watermark X and the extracted watermark X^* is calculated by the following similarity $SIM(X, X^*)$ equation:

$$SIM(X, X^*) = \frac{X \cdot X^*}{\sqrt{X^* \cdot X^*}} \quad (4)$$

It is highly unlikely that X^* will be identical to X . To decide whether X and X^* match, we determine whether the $SIM(X, X^*) > T$, where T is a detection threshold. In this study, the selected detection threshold (T) value is 6 [10].

In order to evaluate the quality of watermarked signal, the following signal-to-noise ratio (SNR) equation is used:

$$SNR = 10 \log_{10} \frac{\sum_{n=1}^N S^2(n)}{\sum_{n=1}^N [S(n) - S^*(n)]^2} \quad (5)$$

where $S(n)$ and $S^*(n)$ are original audio signal and watermarked audio signal respectively.

Figure 4 shows a qualitative evaluation of the original audio with a watermarked audio in which the watermarks are imperceptible using the proposed method.

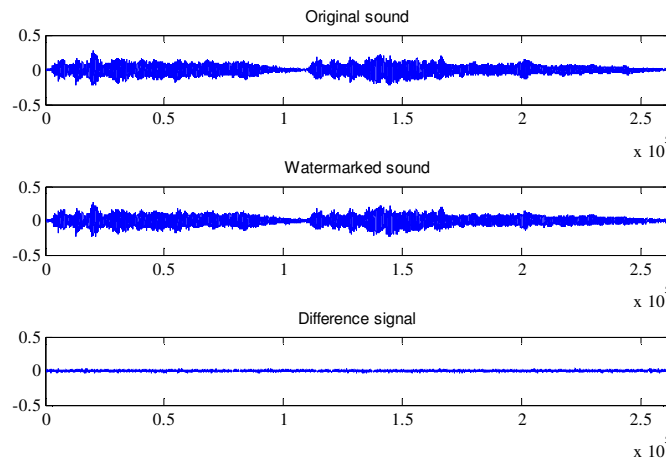


Figure 4. Imperceptibility of the watermarked audio using the proposed method for the audio signal ‘Symphony’

Figure 5 shows the segmentation process applied to the absolute DCT coefficients of the audio signal ‘Symphony’. In this study, the selected number of segments for DCT coefficients is 20.

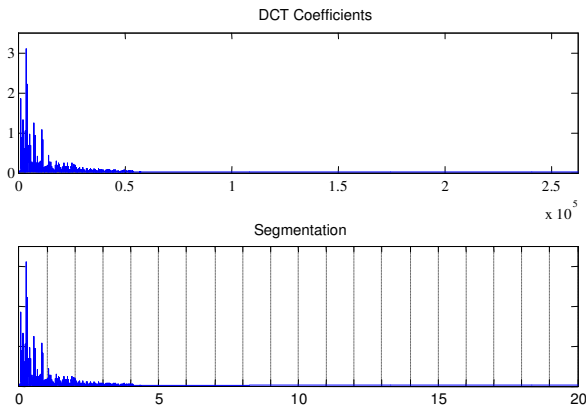


Figure 5. Segmentation of absolute DCT coefficients for the audio signal ‘Symphony’

Figure 6 shows the peak detection of the highest energy segment in DCT domain for original sound ‘Symphony’. In our proposed method, watermarks are embedded into the selected prominent peaks of highest energy segment which provides high robustness against different kinds of attacks as well as good SNR values for watermarked audio signals.

4.1. Imperceptibility Test

Informal listening using head set reveals that the watermark embedded into the original audio signal using the proposed watermarking method does not affect the quality of the sound, which ensures the imperceptibility of the embedded watermark. Table 1 shows the SNR results of the proposed watermarking method for different values of α . Our proposed method achieves SNR values ranging from 13 dB to 24 dB for different watermarked sounds.

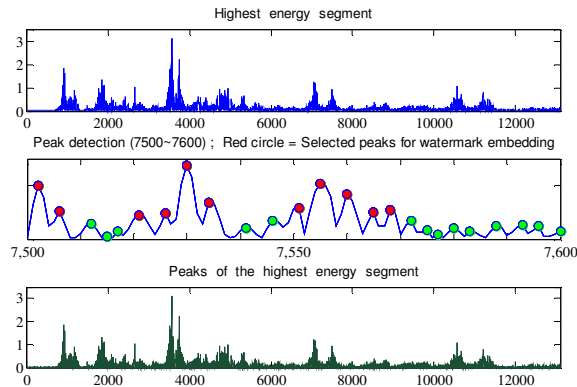


Figure 6. Peak detection of the highest energy segment in DCT domain

Table 1. SNR results of the proposed watermarking method

Types of Signal	SNR		
	$\alpha=0.1$	$\alpha=0.2$	$\alpha=0.3$
Symphony	21.5905	17.5699	14.0481
Gayageum	22.7490	16.7284	13.2066
Human Voice	23.8578	17.8373	14.3153
Classical	24.9609	18.9403	15.4185

2.5 Robustness Test

The robustness of a watermarking method is defined as the ability of watermark detector to extract the embedded watermark after common signal processing operations and attacks. Robustness is measured in terms of the similarity function.

Table 2 shows the performance of our proposed method in terms of similarity when no attack is applied to four different types of watermarked audio signals.

Table 2. Watermark detection results of the proposed method without attacks

Types of Signal	SIM		
	$\alpha=0.1$	$\alpha=0.2$	$\alpha=0.3$
Symphony	31.5060	31.5060	31.5060
Gayageum	31.5060	31.5060	31.5060
Human Voice	31.5060	31.5060	31.5060
Classical	31.5060	31.5060	31.5060

In order to test the robustness of our proposed method, six different types of attacks, summarized in Table 3, were performed to the watermarked audio signal.

Table 3. Attacks used in this study for the watermarked sound

Attacks	Description
Noise addition	Additive white Gaussian noise (AWGN) is added to the watermarked audio signal.
Cropping	We removed 10% samples from the beginning of the watermarked signal and then replaced these samples by

	the original signal.
Re-sampling	The watermarked signal originally sampled at 44.1 kHz is resampled at 22.050 kHz, and then restored by sampling again at 44.1 kHz.
Re-quantization	The 16 bit watermarked audio signal is quantized down to 8 bits/sample and again re-quantized back to 16 bits/sample.
MP3 Compression	MPEG-1 layer 3 compression with 128 kbps is applied to the watermarked signal.
Echo	Decay factor of 25% and the delay time of 100 milliseconds are applied to the watermarked audio signal.

Figure 7 shows the response of the watermark detector to 1000 randomly generated watermarks where correct watermark is at the 520th position against Gaussian noise attack for $\alpha=0.3$ using the proposed method.

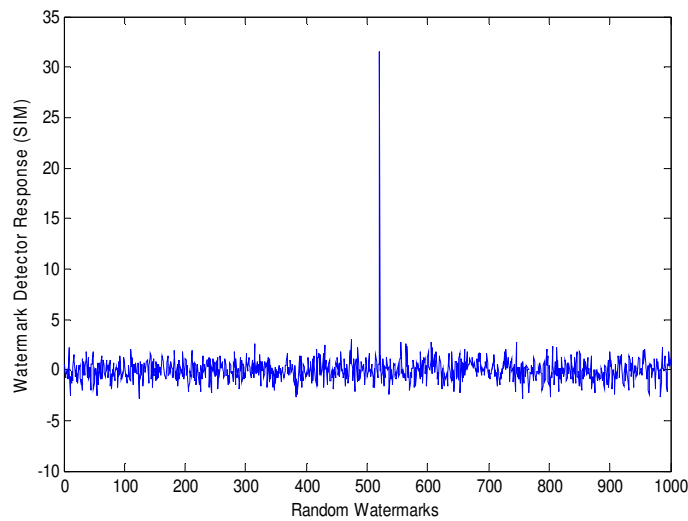


Figure 7. Watermark detector response using the proposed method against white Gaussian noise attack

Figure 8 illustrates the watermark detector response for the audio signal ‘Symphony’ against different kinds of attack for $\alpha=0.3$ using the proposed method.

Tables 4, 5, 6, 7, 8 and 9 show the similarity results of the proposed watermarking method in terms of robustness for different values of α against several kinds of attacks applied to four different types of watermarked audio signal ‘Symphony’, ‘Gayageum’, ‘Human Voice’, and ‘Classical’ respectively.

Overall, our proposed watermarking method provides good results in terms of robustness against several attacks such as noise addition, cropping, resampling, re-quantization, MP3 compression, and echo as well as provides good SNR values for different watermarked sounds.

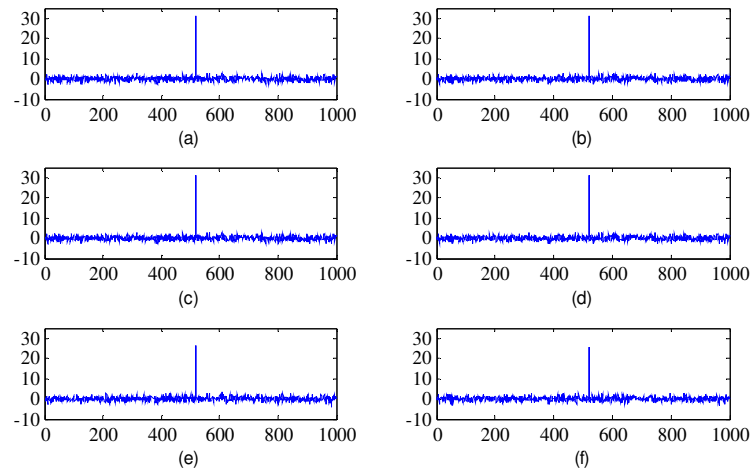


Figure 8. Watermark detector response of the proposed method for the audio file ‘Symphony’ for different types of attack: (a) Noise addition (b) Cropping (c) Re-quantization (d) Re-sampling (e) MP3 compression (f) Echo

Table 4. Similarity results of the proposed method against noise attack

Types of Audio Signal	SIM		
	$\alpha = 0.3$	$\alpha = 0.2$	$\alpha = 0.1$
Symphony	31.2963	31.0412	29.7736
Gayageum	31.1618	30.7437	28.7173
Human Voice	31.4734	31.4328	31.2153
Classical	31.4902	31.4706	31.3647

Table 5. Similarity results of the proposed method against cropping

Types of Audio Signal	SIM		
	$\alpha = 0.3$	$\alpha = 0.2$	$\alpha = 0.1$
Symphony	30.9362	30.9362	30.9361
Gayageum	31.0210	31.0209	31.0210
Human Voice	31.0930	31.0930	31.0931
Classical	31.0234	31.0234	31.0234

Table 6. Similarity results of the proposed method against re-quantization

Types of Audio Signal	SIM		
	$\alpha = 0.3$	$\alpha = 0.2$	$\alpha = 0.1$
Symphony	31.4914	31.4699	31.3497
Gayageum	31.4827	31.4471	31.2568
Human Voice	31.5038	31.5008	31.4815
Classical	31.5051	31.5043	31.4988

Table 7. Similarity results of the proposed method against re-sampling

Types of Audio Signal	SIM		
	$\alpha = 0.3$	$\alpha = 0.2$	$\alpha = 0.1$
Symphony	31.5060	31.5060	31.5060
Gayageum	31.5060	31.5060	31.5060

Human Voice	31.5060	31.5060	31.5060
Classical	31.5060	31.5060	31.5060

Table 8. Similarity results of the proposed method against MP3 compression

Types of Audio Signal	SIM		
	$\alpha = 0.3$	$\alpha = 0.2$	$\alpha = 0.1$
Symphony	26.6607	22.7742	13.9254
Gayageum	25.4869	21.5381	13.5082
Human Voice	27.7913	24.8040	16.3729
Classical	25.3012	21.1650	12.9137

Table 9. Similarity results of the proposed method against echo

Types of Audio Signal	SIM		
	$\alpha = 0.3$	$\alpha = 0.2$	$\alpha = 0.1$
Symphony	25.4980	21.5384	13.3723
Gayageum	25.5508	21.5217	13.2385
Human Voice	25.2057	21.0583	12.7199
Classical	26.0238	22.3288	14.4196

3. CONCLUSIONS

In this paper, we have presented a new watermarking method in DCT domain for copyright protection of audio data. Experimental results indicate that our proposed watermarking method shows strong robustness against several kinds of attacks such as noise addition, cropping, re-sampling, re-quantization, MP3 compression, and echo attack and achieves similarity values ranging from 13 to 32. Moreover, our proposed method achieves SNR values ranging from 13 dB to 24 dB for different watermarked sounds. These results demonstrate that our proposed watermarking method can be a suitable candidate for audio copyright protection.

REFERENCES

- [1] W. N. Lie and L. C. Chang, "Robust and High-Quality Time-Domain Audio Watermarking Based on Low-Frequency Amplitude Modification," *IEEE Transaction on Multimedia*, vol. 8, no. 1, pp. 46-59, February, 2006.
- [2] P. Bassia, I. Pitas and N. Nikolaidis "Robust Audio Watermarking in the Time domain," *IEEE Transaction on Multimedia*, vol. 3, no. 2, pp. 232-241, June, 2001.
- [3] L. Xie, J. Zhang and H. He, "Robust Audio Watermarking Scheme Based on Nonuniform Discrete Fourier Transform," in *Proceedings of IEEE International Conference on Engineering of Intelligent System*, pp. 1-5, 2006.
- [4] G. Zeng and Z. Qiu, "Audio Watermarking in DCT: Embedding Strategy and Algorithm," in *Proceedings of 9th International Conference on Signal Processing (ICSP'09)*, pp. 2193-2196, 2008.
- [5] J. Huang, Y. Wang, and Y. Q. Shi, "A Blind Audio Watermarking Algorithm with Self-Synchronization," in *Proceedings of IEEE International Symposium on Circuits and Systems, (ISCAS'02)*, vol. 3, pp. 627-630, 2002.
- [6] M. Pooyan and A. Delforouzi, "Adaptive and Robust Audio watermarking in Wavelet Domain," in *Proceedings of International Conference on International Information Hiding and Multimedia Signal Processing (IIH-MSP 2007)*, vol. 2, pp. 287-290, 2007.

- [7] X. Y. Wang, H. Zhao, "A novel Synchronization Invariant Audio Watermarking Scheme Based on DWT and DCT," *IEEE Transaction on Signal Processing*, vol. 54, no. 12, pp 4835-4840, December 2006.
- [8] S. K. Lee and Y. S. Ho, "Digital audio watermarking in the cepstrum domain," *IEEE Transactions on Consumer Electronics*, vol. 46, no. 3, pp. 744-750, 2000.
- [9] S. C. Liu and S. D. Lin, "BCH code based robust audio watermarking in the cepstrum domain," *Journal of Information Science and Engineering*, vol. 22, pp. 535-543, 2006.
- [10] I. Cox, J. Killian, F. Leighton, and T. Shamoan, "Secure Spread Spectrum Watermarking for Multimedia," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1673-1687, 1997.

Authors



Pranab Kumar Dhar received the B.S. degree in Computer Science and Engineering from Chittagong University of Engineering and Technology (CUET), Chittagong, Bangladesh in 2004. He received M.S. degree from the School of Computer Engineering and Information Technology at University of Ulsan, South Korea in 2010. Since 2005, he has been serving as a faculty member in the Department of Computer Science and Engineering at Chittagong University of Engineering and Technology (CUET), Chittagong, Bangladesh. His research interest includes Digital Watermarking, Sound Synthesis, Digital Image Processing, Multimedia Systems, and Digital Signal Processing.



Mohammad Ibrahim Khan received the B.S. degree in Electrical and Electronic Engineering from Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh in 1999. He received M.S. degree in Computer Science and Engineering from the same University in 2002. He received his Ph.D degree in Computer Science and Engineering from Jahangirnagar University in 2010. Since 1999, he has been serving as a faculty member in the Department of Computer Science and Engineering at Chittagong University of Engineering and Technology (CUET), Chittagong, Bangladesh. His research interest includes Digital Image Processing, Graph Theory, Cryptography, Digital Watermarking, Multimedia Systems, and Digital Signal Processing.



Saeef Ahmad received the B.S. degree in Computer Science and Engineering from Chittagong University of Engineering and Technology (CUET), Chittagong, Bangladesh in 2010. His research interest includes Digital Watermarking, Multimedia Systems, and Digital System Design.