Audio Zero-Watermarking Based on Discrete Hartley Transform and Non-Negative Matrix Factorization

Pranab Kumar Dhar†*, Emtiaj Hasan†, and Tetsuya Shimamura††

†Department of Computer Science and Engineering, Chittagong University of Engineering and Technology (CUET), Chittagong-4349, Bangladesh

††Graduate School of Science and Engineering, Saitama University, Saitama 338-8570, Japan

*Corresponding author

Summary

This paper proposes a zero-watermarking method based on discrete Hartley transform (DHT) and non-negative matrix factorization (NMF) for audio signal. Initially, the DHT is applied to the original audio signal to obtain the DHT coefficients. Then, the DHT coefficients are divided into an arbitrary number of frames. The coefficients of each frame are arranged into a square matrix and NMF is applied to each matrix. The summation of the coefficients of each weight matrix is calculated and the fractional part of this summation is selected. The fractional part is converted into integer value and modulation operation is performed to generate a binary pattern. XOR operation is applied between the watermark data and generated binary pattern. This operation generates a secret key, which is used later to extract the watermark. The proposed method does not embed any watermark into original audio, indicating that it is imperceptible. Experimental results show that the proposed method is robust against various attacks such as noise addition, re-sampling, re-quantization, low-pass filtering, echo, MP3 compression and so on. Moreover, it shows superior performance than the state-of-the-art watermarking methods in terms of robustness.

Key words:

Copyright protection; Discrete Hartley transform; Non-negative matrix factorization; Zero watermarking

1. Introduction

The rapid developments of the digital world and internet, transmission and distribution of digital data have become an extremely simple task. This has become a serious threat for multimedia content owners. As a result, the issue of digital right management (DRM) of multimedia data has attracted a lot of attention. The most common solution of DMR is digital watermarking. It is a process of embedding watermark into the digital data to show ownership and authenticity. This method has various applications such as data authentication, copyright protection, broadcast monitoring, data indexing, and so on.

Several robust and imperceptible audio watermarking algorithms have been introduced in the literature. A detail survey on various audio watermarking methods can be found in [1]-[2]. Most audio watermarking methods utilize either the time domain [3]-[4] or the transform domain such as discrete Fourier transform (DFT) [5], discrete wavelet transform (DWT) [6]-[7], discrete cosine transform (DCT) [8]-[9]. Among these methods, time domain techniques are simple and easy to implement. However, they have low robustness against various attacks compared with other techniques. On the other hand, transform domain methods can provide high robustness against attacks. The combination of various transforms such as DWT, DCT, Log-polar transform (LPT), Cartesian-polar transform with singular value decomposition (SVD) have been proposed in [10]-[12]. The watermarking algorithms explained earlier have some limitations. One is the insertion of watermark into the host signal inevitably introduces some perceptible quality degradation. Another problem is the inherent conflict between imperceptibility and robustness. To overcome these limitations, zero-watermarking technique has been introduced. In this technique, instead of embedding watermark, it extracts some essential characteristics from the host signal and uses them for detecting watermark. Zero-watermarking algorithm was first proposed by Quan et al. for image [13]. Li and Guangjun introduced an audio zero-watermarking algorithm based on Bayesian information criterion (BIC) and double-direction wavelet coefficients mapping (DWCM) matrix [14]. A blind zero-watermarking algorithm based on DWT has been presented in [15]-[16]. Here, the authors generated secret keys using the characteristics of DWT coefficients. In [17]-[18], authors introduced audio zero-watermarking methods based on DWT and DCT where watermark sequence is generated by using the characteristics of transformed coefficients. Ciptasari [19] proposed a modified version of the method presented in [18] by generating a single secret key rather than three secret keys. An audio zero-watermarking algorithm combined with DCT and Zernike moment was presented by Xiong et al. [20]. Tsai [21] proposed a zero-watermarking method based on energy for audio signals. Most of the methods

Manuscript received August 5, 2017
Manuscript revised August 20, 2017

explained earlier don not provide good robustness against various attacks specially echo, reverse, and MP3 compression. To overcome this limitation, in this paper, we present a zero watermarking method based on discrete Hartley transform (DHT) and nonnegative matrix factorization (NMF) for audio signal. The main features of the proposed method include (i) it utilizes the DHT and NMF jointly in zero watermarking for the first time, (ii) it generates a watermark key from the fractional part of the summation of each weight matrix obtained from the DHT coefficients of each frame, (iii) it provides superior performance than the recent methods [11], [13] in terms of robustness, while keeping comparable robustness with the method reported in [17].

The rest of this paper is organized as follows. Section 2 presents background information including DHT and NMF. Section 3 proposes our watermarking method including watermark embedding and detection processes. Section 4 shows the performance of the proposed method with some recent methods in terms of robustness. Section 5 provides error analysis of the proposed method. Lastly, the conclusion of this paper is provided in Section 6.

2. Background Information

2.1 Discrete Hartley Transform

Discrete Fourier transform (DFT) plays a vital role in signal processing. Despite its tremendous application, the DFT has an unattractive feature. It transforms a real-valued sequence into a complex-valued sequence. R. N. Bracewell [22] proposed an inherently real-valued transform called the discrete Hartley transform (DHT). The new transform has the advantage that a real-valued signal always generates a real-valued transformed signal. Also, unlike the DFT, the DHT is symmetric, i.e., both the forward and inverse transforms are identical. Moreover, for computing real sequence, fast Hartley transform (FHT) is faster than fast Fourier transform (FFT) [19].

Mathematically, the DHT can be written as:

$$H(k)=\sum_{n=0}^{N-1}x(n)\left[\cos\left(\frac{2\pi nk}{N}\right)+\sin\left(\frac{2\pi nk}{N}\right)\right],\quad k=0,1,......,N-1 \qquad (1)$$

where $x(n)$ is the original signal in time domain and $H(k)$ is transformed signal in DHT domain with length N.

2.2 Non-Negative Matrix Factorization

NMF was designed to find a representative basis vector with nonnegative element [13], which overcomes the limitation of SVD. In many applications, negative elements may contradict physical realities. Let A be an arbitrary matrix of size $m\times m$ with NMF of the form $A=BH$, where B is an $m\times r$ nonnegative matrix, containing the NMF basis vector and H is an $r\times m$ nonnegative weight matrix, containing the associate coefficients.

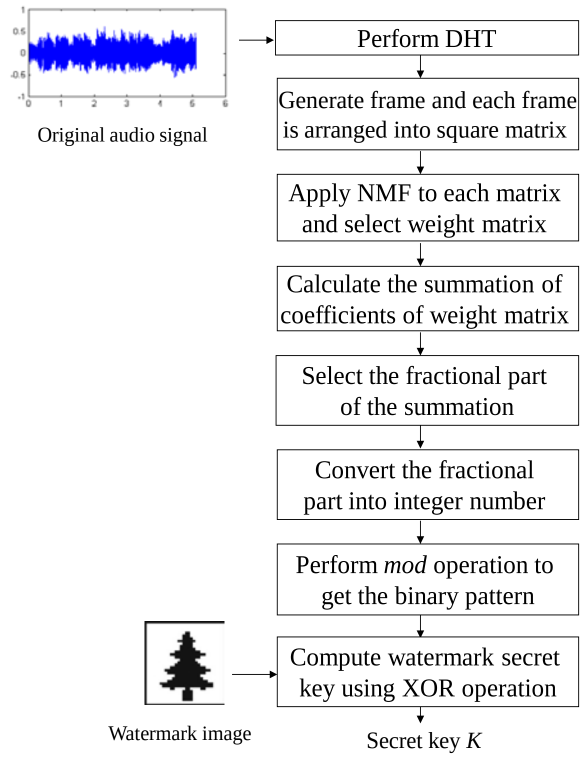


Fig 1. Watermark embedding process

3. Proposed Watermarking Method

Let $A=\{a(i),\ 1\le i\le L\}$ be a host audio signal with L samples, $W=\{w(i,j),\ 1\le i\le M,\ 1\le j\le M\}$ be a binary image to be embedded into the host signal, and $w(i,j)\in\{0,1\}$ be the pixel value at point (i, j).

3.1 Watermark Embedding Process

The watermark embedding process is shown in Fig. 1. This process is described as follows:

1) The original audio signal A is transformed into DHT domain to calculate the DHT coefficients $C=\{c(i),\ 1\le i\le L\}$, where i is the coefficient number and L is the total number of coefficients.

2) The DHT coefficients C is segmented into m non-overlapping frames $F=\{F_j,\ 1\le j\le m\}$, where j indicates the frame number. Each frame F_j contains r numbers of coefficients denoted by $R=\{R_k,\ 1\le k\le r\}$.

3) The DHT coefficients of each frame F_j are rearranged into a $P \times P$ square matrix D_j. This is done by dividing the coefficient set into P segments with P coefficients.

4) NMF is applied to each block D_j of the each frame F_j. NMF is represented as follows:

$$D_j = B_j H_j \tag{2}$$

where B_j is a non negative matrix, containing the NMF basis vectors and H_j is a weight matrix containing the associate coefficients.

5) The matrix H_j is arranged into a one dimensional sequence Y_j. Each sequence Y_j contains r numbers of coefficients denoted by $U=\{U_k, 1 \le k \le r\}$.

6) The summation $S=\{S_j, 1 \le j \le m\}$ of the absolute value of the coefficients of each sequence Y_j is calculated using the following equation:

$$S_j = \sum_{k=1}^{r} |U_k| \tag{3}$$

where $|R_k|$ represents the absolute value of R_k.

7) Select the fractional part of the summation S_j and convert it into an integer number T_j using the following equation:

$$T_j = convert\left(S_j - floor\left(S_j\right)\right) \tag{4}$$

8) A binary pattern $X=\{X_j, 1 \le j \le m\}$ is generated from T_i using the following equation:

$$X_j = \mathrm{mod}\left(\left(T_j\right)^P, 2\right) \tag{5}$$

where P represents the T_j-th prime number.

9) Convert the binary pattern X into an $M \times M$ matrix represented by $Y = \{y(i,j),\ 1 \le i \le M,\ 1 \le j \le M\}$.

10) Compute the watermark secret key $K = \{k(i,j),\ 1 \le i \le M,\ 1 \le j \le M\}$ using the following equation:

$$k(i,j) = w(i,j) \oplus y(i,j) \tag{6}$$

where $\oplus$ is the exclusive-or (XOR) operation. The key K is used in watermark extraction process.

3.2 Watermark Detection Process

The proposed watermark detection process is shown in Fig. 2. This process does not need the original audio signal to extract the watermark. The detection process is described as follows:

1) The DHT is applied to the attacked watermarked audio A^* to get the DHT coefficients $C^* = \{c^*(i),\ 1 \le i \le L\}$.

2) The DHT coefficients C^* is segmented into m non-overlapping frames $F^* = \{F_j^*,\ 1 \le j \le m\}$ and each frame F_j^* contains r numbers of coefficients represented by $R^* = \{R_k^*,\ 1 \le k \le r\}$.

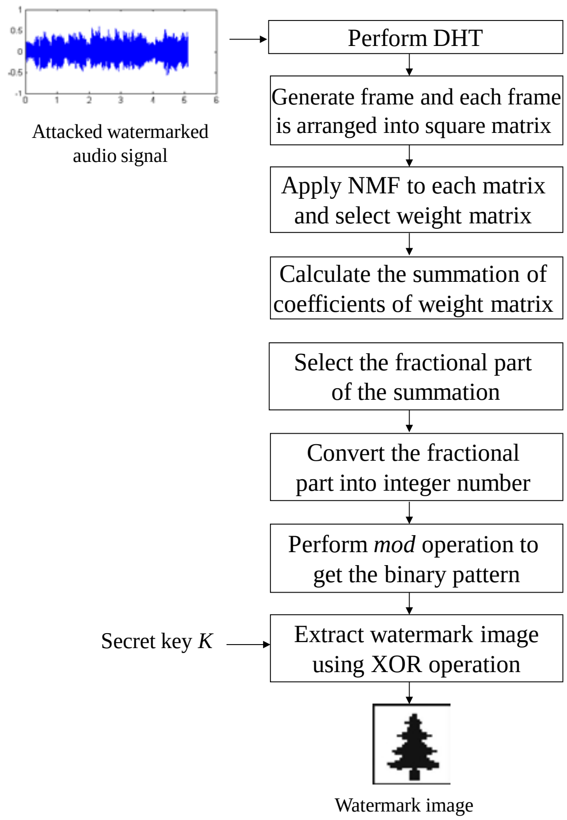


Fig 2. Watermark detection process

3) The DHT is applied to the attacked watermarked audio A^* to get the DHT coefficients $C^* = \{c^*(i),\ 1 \le i \le L\}$.

4) The DHT coefficients C^* of each frame F_j^* are rearranged into a $P \times P$ square matrix D_j^*.

5) NMF is applied to each block D_j^* of the each frame F_j^*.

6) The matrix H_j^* is arranged into a one dimensional sequence Y_j^*. Each sequence Y_j^* contains r numbers of coefficients denoted by $U=\{U_k^*, 1 \le k \le r\}$.

7) The summation $S^* = \{S_j^*,\ 1 \le j \le m\}$ of the coefficients $|U_k^*|$ of each sequence Y_i^* is calculated using the following equation:

$$S_j^* = \sum_{k=1}^{r} |U_k^*| \tag{7}$$

8) Calculate T_j^* using the following equation:

$$T_j^* = convert\left(S_j^* - floor\left(S_j^*\right)\right) \tag{8}$$

9) A binary pattern $X^* = \{X_j^*,\ 1 \le j \le m\}$ is generated from T_i using the following equation:

$$X_j^* = \mathrm{mod}\left(\left(T_j^*\right)^P, 2\right) \tag{9}$$

where P represents the T_j^*-th prime number.

10) Convert the binary pattern X^* into an $M\times M$ matrix represented by $Y^* = \{y^*(i,j),\ 1 \le i \le M,\ 1 \le j \le M\}$.

11) Watermark image $W^* = \{w^*(i,j),\ 1 \le i \le M,\ 1 \le j \le M\}$ is extracted using the following equation:

$$w^*(i,j) = k(i,j) \oplus y^*(i,j) \quad (10)$$

where $\oplus$ is the exclusive-or (XOR) operation.

4. Experimental Results and Analysis

In this section, several experiments were carried out to demonstrate the performance of the proposed watermarking method. The performance of the proposed method was evaluated in terms of robustness and imperceptibility. In this study, we selected four different sound files from the album Rust [23]. These are: (a) Citizen, Go Back to Sleep, (b) Beginning of the End, (c) Breathing on Another Planet, and (d) Thousand Yard Stare. All audio files contain 4194304 samples (duration 95 seconds) with sampling rate 44.1 KHz having 16 bits per sample as shown in Fig. 3. In our experiment, frame length was fixed with 4096 samples. A binary image with size 32×32 is used as a watermark shown in Fig. 4.

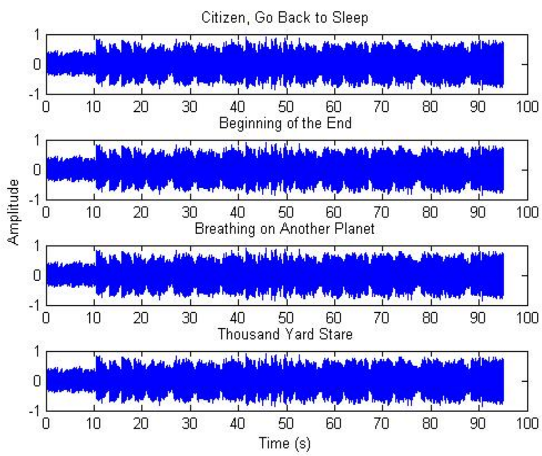


Fig 3. The original audio signals used in this study

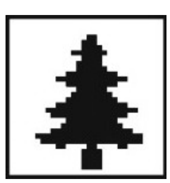


Fig 4. Watermark image

4.1 Imperceptibility Analysis

The watermarked audio signal is identical to the original one because the watermark is actually, embedded into the secret key, not into the audio signal. Therefore, without processing various imperceptibility analysis techniques, we can say that the proposed watermarking method provides excellent watermarked audio signals.

4.2 Robustness Analysis

Normalized correlation (NC) coefficient is used to measure the similarities between the original watermark W and the extracted watermark W^*, which is calculated by the following equation:

$$NC\ (W,W^*) = \frac{\sum_{i=1}^{M}\sum_{j=1}^{M} w(i,j)\cdot w^*(i,j)}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{M} w(i,j)\cdot w(i,j)}\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{M} w^*(i,j)\cdot w^*(i,j)}} \quad (10)$$

where i and j are the indices of the binary watermark image. The correlation between W and W^* is very high if NC (W, W^*) is near to 1. On the other hand, the correlation between W and W^* is very low if NC (W, W^*) is near to 0.

The robustness of a watermarking method can be measured using BER which is calculated by the following equation:

$$BER\ (W,W^*) = \frac{\sum_{i=1}^{M}\sum_{j=1}^{M} w(i,j) \oplus w^*(i,j)}{M \times M} \times 100\% \quad (11)$$

For evaluating the robustness of the proposed method, following nine different types of attacks were applied to the audio signals:

1) Noise addition: Additive white Gaussian noise (AWGN) is added to the original audio signal until the resulting signal has an SNR of 20 dB.
2) Re-sampling: The audio signal originally sampled at 44.1 kHz, is re-sampled at 22.050 kHz, and then restored by sampling again at 44.1 kHz.
3) Low-pass filtering: A low-pass filter with cut-off frequency 11.025 kHz is used.
4) Re-quantization: The 16 bit audio signal is quantized down to 8 bits/sample and again re-quantized back to 16 bits/sample.
5) Echo: An echo signal with a delay of 0.5s is added to the audio signal.
6) Reverse: The audio is reversed to its original.
7) MP3 compression: MPEG-1 layer 3 compression with 128 kbps, 64 kbps, and 32 kbps is applied to the original signal.

Gold Wave [24] was used in this experiment for re-sampling, low-pass filtering, echo, reverse, and MP3 compression. MATLAB [25] was used for implementing the noise addition and re-quantization attacks.

Table 1 shows the NC and BER result of the proposed watermarking method in terms of robustness against several attacks for the audio signal "Citizen, Go Back to

Sleep". The extracted watermark images are also shown in Table 1. We observed that the extracted watermark images are almost similar compared to the original image. It is seen that all NC values range from 0.90 to 0.92, whereas all BER values range from 9% to 10%. This clearly indicates a good performance of the proposed method against different attacks. Table 2 shows similar results for the audio signal "Beginning of the End", "Breathing on Another Planet", and "Thousand Yard Stare", respectively. The NC values range from 0.90 to 0.92 and BER values range from 9% to 12%, demonstrating the high robustness of our proposed method against various attacks. This is because it generates a watermark key from the fractional part of the summation of each weight matrix obtained from the DHT coefficients of each frame.

Table 3 provides a comparative analysis between the proposed and some recent watermarking methods for audio signal 'Citizen, Go Back to Sleep' in terms of NC and BER, respectively. For the echo and reverse attacks, the proposed method clearly shows better NC and BER values compared with the other methods. For the noise addition, low-pass filtering and MP3 compression attacks, the proposed method provides better NC and BER values than the methods reported in [11] and [13] while keeping comparable robustness with the method proposed in [17]. Fig. 5 and Fig. 6 show a comparative graphical representation between the proposed method and other related algorithms in terms of NC and BER, respectively. Here, each attack is placed on the horizontal axis and the values of NC and BER for each attack of different methods are placed in the vertical axis against respective attack. From these two figures, we observed that, the NC and BER values of the proposed method range from 0.90 to 0.92 and 9% to 12%, respectively whereas the NC and BER values of the methods proposed in [11] and [13] range from 0.40 to 0.96 and 25% to 60%, respectively. This clearly indicates the good performance of the proposed method compared with other methods. Moreover, the proposed method shows comparable robustness with the method proposed in [17].

Table 1. NC and BER of Extracted Watermark Image for the Audio Signal 'Citizen, Go Back to Sleep'

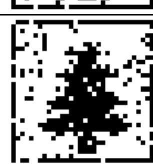
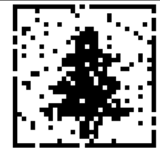
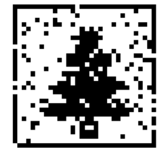
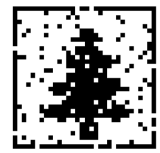
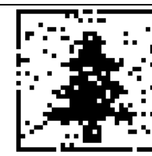
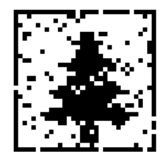
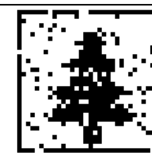
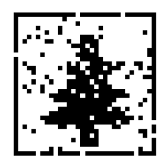
Attack Type	NC	BER	Extracted watermark
Noise addition	0.9206	10.06	
Re-sampling	0.9158	10.64	
Low-pass filtering	0.9145	10.84	
Re-quantization	0.9172	10.45	
Echo	0.9268	9.28	
Reverse	0.9199	10.16	
mp3 compression (32 kbps)	0.9225	9.77	
mp3 compression (64 kbps)	0.9279	9.18	
mp3 compression (128 kbps)	0.9271	9.28	

Table 2. NC and BER of the Extracted Watermark for Different Audio Signals

Audio Signal	Attack type	NC	BER
Beginning of the End	Noise addition	0.9141	10.84
	Re-sampling	0.9172	10.45
	Low-pass filtering	0.9232	9.77
	Re-quantization	0.9268	9.38
	Echo	0.9108	11.33
	Reverse	0.9233	9.77
	MP3 compression (32 kbps)	0.9270	9.28
	MP3 compression (64 kbps)	0.9162	10.64
	MP3 compression (128 kbps)	0.9154	10.65
Breathing On Another Planet	Noise addition	0.9264	9.57
	Re-sampling	0.9129	11.04
	Low-pass filtering	0.9135	10.94
	Re-quantization	0.9228	9.77
	Echo	0.9213	9.96
	Reverse	0.9170	10.45
	MP3 compression (32 kbps)	0.9194	10.16
	MP3 compression (64 kbps)	0.9156	10.64
	MP3 compression (128 kbps)	0.9108	11.23

Thousand Yard Stare	Noise addition	0.9187	10.25
	Re-sampling	0.9194	10.26
	Low-pass filtering	0.9132	11.04
	Re-quantization	0.9214	9.96
	Echo	0.9162	10.64
	Reverse	0.9041	12.11
	MP3 compression (32 kbps)	0.9108	11.33
	MP3 compression (64 kbps)	0.9136	10.94
	MP3 compression (128 kbps)	0.9194	10.25

Table 3. Comparison of BER Between The Proposed And Other Recent Methods for the Signal ‘Citizen, Go Back to Sleep’

Attack	BER [Proposed]	BER [17]	BER [11]	BER [13]
Noise addition	10.06	1.17	47.85	37.50
Re-sampling	10.64	0.59	0.98	25.00
Low-pass filtering	10.84	0.68	43.65	50.00
Re-quantization	10.45	0.00	4.69	49.00
Echo	9.28	27.83	49.02	25.00
Reverse	10.16	56.64	48.63	43.75
MP3 compression (32 kbps)	9.77	4.79	49.51	62.50
MP3 compression (64 kbps)	9.18	4.10	49.71	56.25
MP3 compression (128 kbps)	9.28	4.10	52.25	31.25

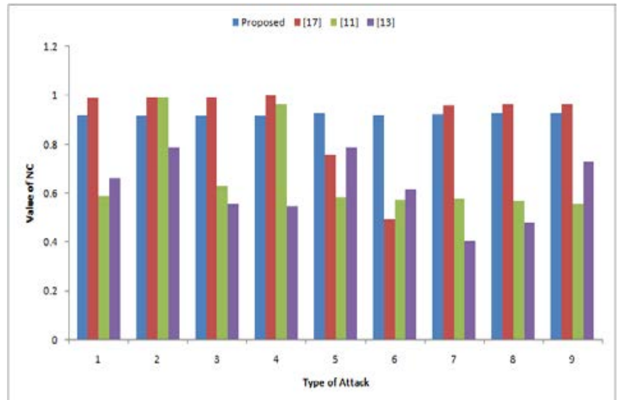


Fig 5. NC comparison between the proposed and several recent methods in terms of graphical representation

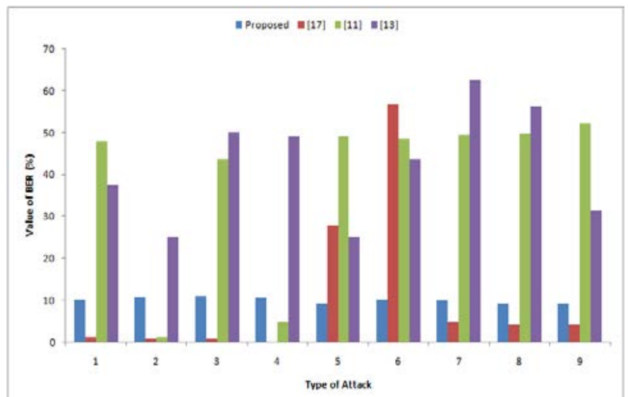


Fig 6. BER comparison between the proposed and several recent methods in terms of graphical representation

5. Error Analysis

In watermarking, generally two types of error are analyzed: (1) false positive error (FPE) (2) false negative error (FNE). These errors are very harmful because they impair the credibility of watermarking method. It is difficult to give an exact probabilistic model for an FPE and an FNE. The probability of FPE and FNE can be measured using a binomial probability distribution similar to that in [10].

5.1 False Positive Error:

The FPE is the probability that an unwatermarked audio signal is declared as watermarked signal by the detector. Then, FPE probability P_{f_p} can be calculated as:

$$P_{f_p} = 2^{-k} \sum_{m=\lceil 0.8k \rceil}^{k} \binom{k}{m} \qquad (12)$$

where $\binom{k}{m}$ is the binomial coefficient, k is the total number of watermark bits and m is the total number of matching bits. Fig. 7 shows the FPE probability for $k \in (0,70]$. It is noted that P_{f_p} approaches 0 when k is larger than 15. In our proposed method, k = 1024, therefore, the FPE probability is approximately equal to 0.

5.2 False Negative Error:

The FNE is the probability that a watermarked audio signal is declared as unwatermarked signal by the detector. Then FNE probability P_{f_n} can be calculated as:

$$P_{f_n} = \sum_{m=0}^{\lceil 0.8k \rceil - 1} \left[\binom{k}{m} (P)^m (1-P)^{k-m} \right] \qquad (13)$$

where P is the bit error rate probability of the extracted watermark. The approximate value of P can be obtained from BER under different attacks. From the Tables I and II we observed that all BER values are less than 0.13. Thus, P is taken as 0.87. Fig. 8 plots the FNE probability for $k \in (0,70]$. It is noted that P_{f_n} approaches 0 when k is larger than 10.

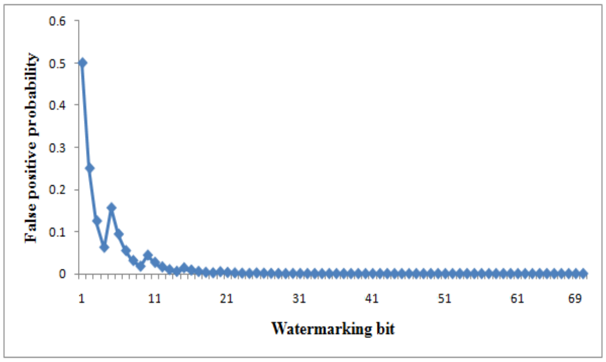


Fig 7. False positive probabilities under various k

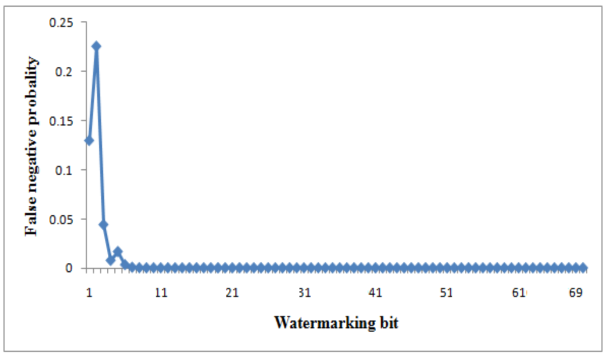


Fig 8. False negative probabilities under various k

5. Conclusion

In this paper, a zero-watermarking method based on DHT was introduced for audio signal. The proposed method does not modify the original audio signal, which ensures the imperceptibility of the signal. Moreover, it provides good robustness against various attacks such as such as noise addition, re-sampling, re-quantization, low-pass filtering, echo, MP3 compression and so on. This is because it generates a watermark key from the fractional part of the summation of the weight matrix obtained from the DHT coefficients of each frame. The NC and BER values of the proposed method range from 0.90 to 0.92 and 9% to 12%, respectively. This is in contrast to other methods whose NC and BER values range from 0.40 to 0.99 and 4% to 60%, respectively. These results verify that the proposed audio zero watermarking method can be a good competent for copyright protection.

References

[1] N. Cvejic and T. Seppanen, Digital audio watermarking techniques and technologies: applications and benchmarks, IGI Global, 2007.

[2] P. K. Dhar, and T. Shimamura, Advances in Audio Watermarking Based on Singular Value Decomposition, Springer, USA, 2015.

[3] P. Bassia, I. Pitas, and N. Nikolaidis, “Robust audio watermarking in the time domain,” IEEE Trans. Multimedia, vol. 3, no. 2, pp. 232-241, 2001

[4] A. Lemma, J. Aprea, and W. Oomen, “A temporal domain audio watermarking technique,” IEEE Trans. Signal Processing, vol. 51, no. 4, pp. 1088–1097, 2003.

[5] J. Singh, P. Garg, and A. De, “Multiplicative watermarking of audio in DFT magnitude,” Multimedia Tools and Applications, vol. 71, no. 3, pp. 1431-1453, 2014.

[6] G. Zhao, “A new digital audio watermarking algorithm based on DWT,” in Proc. International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC), pp. 6513-6516, 2011.

[7] S. T. Chen, G. D. Wu, and H. N. Huang, “Wavelet domain audio watermarking scheme using optimization based quantisation,” IET Signal Processing, vol. 4, no. 6, pp. 720–727, 2010.

[8] Y. Yang, R. Huang, and X. Mintao, “A novel audio watermarking algorithm for copyright protection based on DCT domain,” in Proc. International Symposium on Electronic Commerce and Security, pp. 184-188, 2009.

[9] Q. Guo, Y. Zhao, P. Cheng, and F. Wang, “An audio digital watermarking algorithm against A/D and D/A conversions based on DCT domain,” in Proc. International Conference on Consumer Electronics, Communications and Networks, pp. 871–876, 2012.

[10] P. K. Dhar and T. Shimamura, “Blind SVD-based audio watermarking using entropy and log-polar transformation”, Journal of Information Security and Application, vol. 20, pp. 74-83, Elsevier, 2015.

[11] P. K. Dhar and T. Shimamura, “Audio watermarking in transform domain based on singular value decomposition and Cartesian-polar transform,” International Journal of Speech Technology, vol. 17, no. 3, pp. 133-144, Springer, 2014.

[12] P. K. Dhar and T. Shimamura, “A DWT-DCT based audio watermarking method using singular value decomposition and quantization,” Journal of Signal Processing, vol. 17, no. 3, pp. 69-79, 2013.

[13] W. Quan, T. Sun, and S. Wang, “Zero-watermark watermarking for image authentication,” in Proc. International Conference Signal and Image Processing, pp. 503–508, 2002.

[14] X. Li and H. Guangjun, “Efficient audio Zero-watermarking algorithm for copyright protection based on BIC and DWCM matrix,” International Journal of Advancements in Computing Technology, vol. 4, no. 6, pp. 109-117, 2012.

[15] Y. Yang, M. Lei, H. Liu, Y. Zhou, and Q. Luo, “A novel robust zero-watermarking scheme based on discrete wavelet transform,” Journal of Multimedia, vol. 7, no. 4, pp. 303–308, 2012.

[16] G. H. Wu, Q. H. Wu, and X. D. Zhou, “An audio zero-watermarking algorithm based on DWT,” Mechanical and Electrical Engineering Magazine, vol. 9, no. 6, 2009.

[17] H. L. Dai and D. He, “An efficient and robust zero-watermarking scheme for audio based on DWT and DCT,”

in Proc. Asia Pacific Conference on Postgraduate Research in Microelectronics and Electronics, pp. 233-236, 2009.
[18] N. Chen and J. Zhu, “A robust zero-watermarking algorithm for audio,” EURASIP Journal on Advances in Signal Processing, vol. 2008, no. 103, pp. 1-7, 2008.
[19] R. W. Ciptasari, “An efficient key generation method in audio zero-watermarking,” in Proc. International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIHMSP), pp. 336-339, 2011.
[20] Y. Xiong and R. Wang, “An audio zero-watermark algorithm combined DCT with zernike moments,” in Proc. International Conference on Cyberworlds, pp. 11–15, 2008.
[21] S. M. Tsai, “A robust zero-watermarking scheme for digital audio,” International Journal of Information and Electronics Engineering, vol. 5, no. 2, pp. 117–121, 2015.
[22] R. N. Bracewell, “The discrete hartley transform,” Journal of Optical Society of America, vol. 73, no. 12, pp. 1832–1835, 1983.
[23] http://www.jamendo.com/en/album/7365 (Accessed on: June 2, 2016)
[24] https://www.goldwave.com (Accessed on: October 28, 2016)
[25] https://www.mathworks.com/products/matlab

Pranab Kumar Dhar received the Bachelor of Science (B.Sc.) Degree in Computer Science and Engineering from Chittagong University of Engineering and Technology (CUET), Chittagong, Bangladesh in 2004. He received the Master of Science (M.Sc.) Degree from the School of Computer Engineering and Information Technology, University of Ulsan, South Korea in 2010. He received his Ph.D Degree from Graduate School of Science and Engineering, Saitama University, Saitama, Japan in 2014. In 2005, he joined as a Lecturer in the Department of Computer Science and Engineering, Chittagong University of Engineering and Technology (CUET), Chittagong, Bangladesh where he is currently serving as an Associate Professor. He is an author of one book. His research interest includes Multimedia Security, Digital Watermarking, Multimedia Data Compression, Sound Synthesis, Digital Image Processing, and Digital Signal Processing.

Emtiaj Hasan received the Bachelor of Science (B.Sc.) Degree in Computer Science and Engineering from Chittagong University of Engineering and Technology (CUET), Chittagong, Bangladesh in 2016. His research interest includes Digital Watermarking, Multimedia Systems, and Digital System Design.

Tetsuya Shimamura received the B.E, M.E., and Ph.D Degrees in Electrical Engineering from Keio University, Yokohama, Japan, in 1986, 1988, and 1991, respectively. In 1991, he joined Saitama University, Saitama, Japan, where he is currently a Professor. He was a visiting researcher at Longhborough University, U.K. in 1995 and at Queen’s University of Belfast, U.K. in 1996, respectively. Prof. Shimamura is an author and coauthor of 6 books. He serves as an editorial member of several international journals and is a member of the organizing and program committees of various international conferences. His research interests are in digital signal processing and its application to speech, image, and communication systems.