

A Zero-watermarking Scheme Based on Discrete Hartley Transform for Audio Signal

First A. Author, *Fellow, IEEE*, Second B. Author, and Third C. Author, Jr., *Member, IEEE*

Abstract— This paper proposes a zero-watermarking scheme based on Discrete Hartley Transform (DHT) for audio signal. This algorithm does not embed any information into original audio, therefore, is imperceptible. So it actually proposes a solution in which there is no thinking about the tradeoffs between imperceptibility and robustness. In this scheme, DHT is performed on the audio. By doing this, some coefficients have been found. Then, the coefficients are divided into an arbitrary number of frames. After that, the binary pattern is generated. In this scheme, a binary-valued image is used as a watermark. XOR operation is performed between watermark and generated binary pattern. The output is a matrix of key, which can be used for watermark extraction. The experimental results show that this algorithm can resist various attacks and it shows much better result than other existing schemes.

Keywords—Audio watermarking, Discrete Hartley Transform, Zero-watermark.

I. INTRODUCTION

Internet is the fastest medium of transferring data to any place in a world and a popular digital media is audio. It contains a huge database of digital audio. Moreover, for the cheaper price of different storage device e.g., flash drive, make it easier for people to share, distribution of audio files easily and sometimes illegally. For these reasons, the chances of copyright infringement are higher than ever and measures against piracy were never so demanded. In this view, an audio watermarking field is very important.

Audio watermarking techniques that proposed so far can be categorized into time domain, transform domain, zero-watermarking technique, and also into other techniques.

Among these methods, time domain technique is the simplest and easy to implement. But it shows worse robustness and imperceptibility comparing with other techniques. Various techniques have been proposed so far based on time domain [1, 2].

In transform domain watermarking techniques, an audio is processed by means of a specific transform. The frequency domains such as Discrete Fourier Transform (DFT) [3], Discrete Wavelet Transform (DWT) [4, 5], Discrete Cosine Transform (DCT) [6, 7] are generally used in audio watermarking. The combination of various transformations is also used in audio watermarking, e.g., a combination of DWT and DCT with Singular Value Composition (SVD) has been

proposed [8].

So various watermarking schemes have been proposed but they may slightly distort the original signal and cannot provide a perfect balance between robustness and imperceptibility.

In such a case, a zero-watermark algorithm was applied to the image by Quan [9] for the first time. Zero watermarking technique does not modify any contents of the original signal and produces significant features from a signal. So it is imperceptible.

Based on BIC and DWCM matrix, a zero-watermarking scheme has been proposed [10]. A blind zero-watermarking algorithm based on the DWT has been used to construct secret keys [11, 12]. DWT and DCT have been combined to generate the watermarking sequences [13, 14]. Ciptasari [15] proposed a modified version of [14] by producing a secret key rather than three secret key. An audio zero-watermarking algorithm combined with DCT and Zernike moment proposed by Xiong [16]. The concept of the zero-watermark based on energy was also proposed to protect audio signals [17].

Though various zero-watermarking schemes have been proposed, the problem is, for some signal processing manipulation i.e., compression, adding echo, re-sampling, etc, existing schemes do not perform satisfactory result. To overcome the limitation, here, we propose a new zero-watermarking scheme based on DHT.

II. BACKGROUND INFORMATION

Discrete Fourier Transform (DFT) plays a vital role in signal processing. Despite its tremendous application, the DFT has an unattractive feature, that, it transforms a real-valued sequence also into a complex-valued sequence. R. N. Bracewell [18] 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 transform 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].

Considering a sequence of N real numbers x_n for $n = 0, 1, \dots, N-1$, the DHT of this sequence is defined by equation (1).

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

III. WATERMARK EMBEDDING SCHEME

Let $A = \{a(i); i = 1, 2, \dots, L\}$ be the host audio signal and $W = \{w(i,j); i=1,2, \dots, M_1, j=1,2, \dots, M_2\}$ be the binary-valued image watermark to be embedded. The watermark embedding procedure shown in Fig. 1 can be described as follows:

Step 1: The discrete hartley transform is applied to the audio A , and can get the coefficients $C = \{c(i); i = 1, 2, \dots, C_A\}$, which has C_A samples.

Step 2: C is segmented into M frames, each frame includes Q samples.

Step 3: for $i=1:M$

3.1) Compute $s_i = \sum_{n=1}^Q |C_n|$; where $|C_n|$ means absolute value of C_n .

3.2) Compute $t_i = \text{convert}(s_i - \text{floor}(s_i))$; where $\text{floor}(s_i)$ is a function which gives the largest integer less than or equal to s_i , in a word, a function which extract integer part from a real value, and $\text{convert}()$ is a function which converts a fractional value into an integer value.

3.3) Compute $x_i = \text{mod}(t_i^p, 2)$; where p is the t_i th prime.

Step 4: Convert x into $M_1 \times M_2$ matrix.

Step 5: Compute the watermarking extraction secret key $k(i,j)$ by doing the exclusive or operation between $w(i,j)$ and $x(i,j)$. Therefore, $k(i,j) = w(i,j) \oplus x(i,j)$.

Here, in step 3.2, the output value is the fractional part of the value found from previous step i.e., 3.1. Suppose after calculation of step 3.1, the summation value is 13.24. In that case, the output of step 3.2 is 24. And about step 3.3, here, considering previous example, as 24th prime is 89, the two parameter of modulus operation are 24^{89} and 2 respectively.

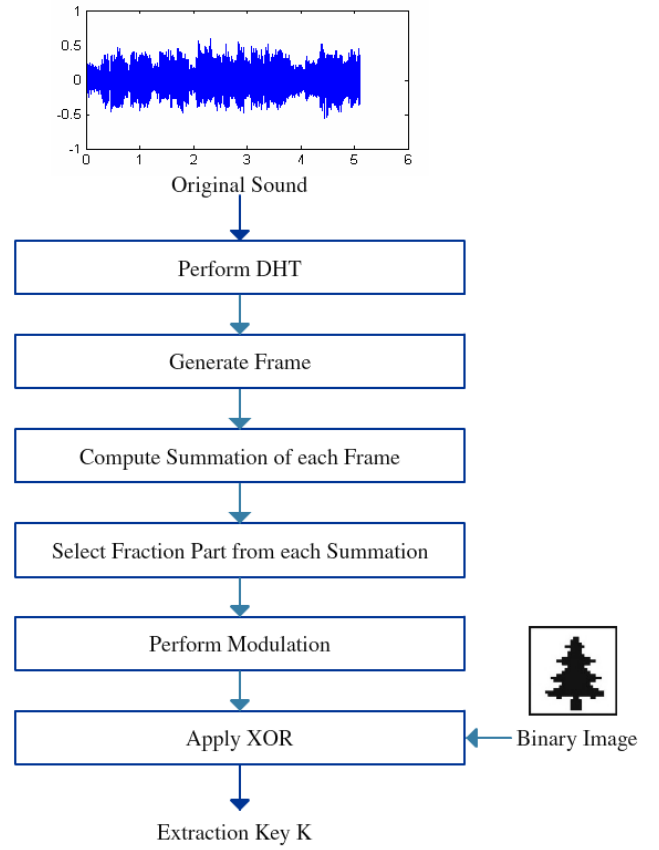


Fig. 1. Watermark embedding process

IV. WATERMARK EXTRACTION SCHEME

The watermark recovery procedure shown in Fig. 2 can be carried out as follows:

Step 1: The discrete hartley transform is applied to the attacked audio A^* , and can get the coefficients $C^* = \{c^*(i); i = 1, 2, \dots, C_A^*\}$, which has C_A^* samples.

Step 2: C^* is segmented into M frames, each frame includes Q samples.

Step 3: for $i=1:M$

3.1) Compute $s_i^* = \sum_{n=1}^Q |C_n^*|$

3.2) Compute $t_i^* = \text{convert}(s_i^* - \text{floor}(s_i^*))$;

3.3) Compute $x_i^* = \text{mod}(t_i^{*p}, 2)$; where p is the t_i^* th prime.

Step 4: Convert x^* into $M_1 \times M_2$ matrix.

Step 5: Compute the watermarking information $W^*(i,j) = k(i,j) \oplus x^*(i,j)$ through computing the exclusive or operation between $k(i,j)$ and $x^*(i,j)$.

V. EXPERIMENTAL RESULTS AND ANALYSIS

In this experiment, a binary image taken from [20] with size 32X32 is used as the watermark shown in Fig. 3.

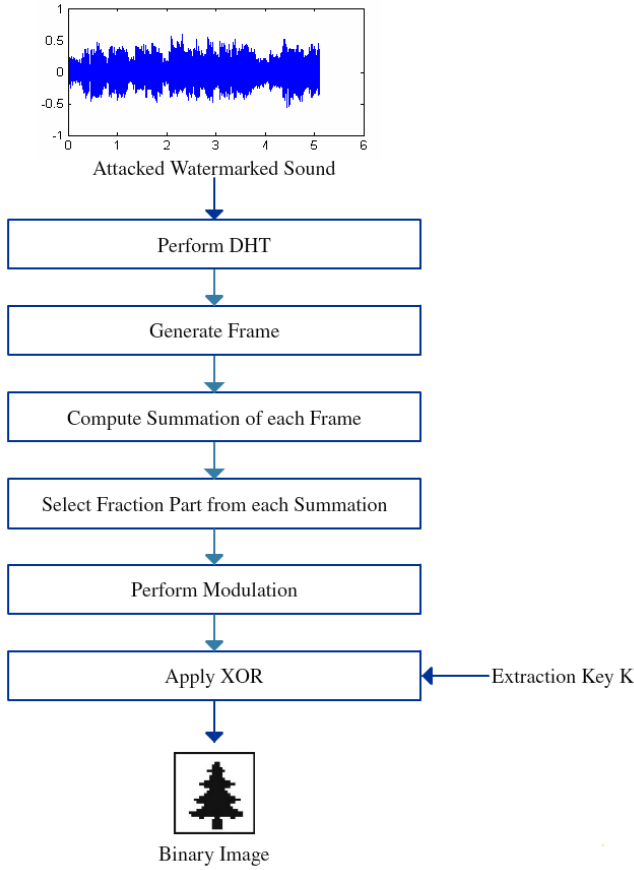


Fig. 2. Watermark extraction process

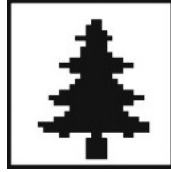


Fig. 3. Watermark image with size 32X32

For evaluation the performance of the proposed watermarking method different types audio signals are used. The sound files are: (a) Citizen, Go Back to Sleep, (b) Beginning of the End, (c) Breathing On Another Planet, and (d) Thousand Yard Stare, included in the album *Rust* [21]. All audio files contain 4194304 samples, 95 seconds long, are sampled at 44.1 KHz with 16 bits per sample as shown in Fig. 4. For this experiment, frame length was fixed at 4096 samples.

To test the robustness of the watermarked algorithm, the following normalized coefficient (NC) and bit error rate (BER) are employed respectively:

$$NC(W, W^*) = \frac{\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} W(i, j) W^*(i, j)}{\sqrt{\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} W(i, j)^2} \sqrt{\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} W^*(i, j)^2}}$$

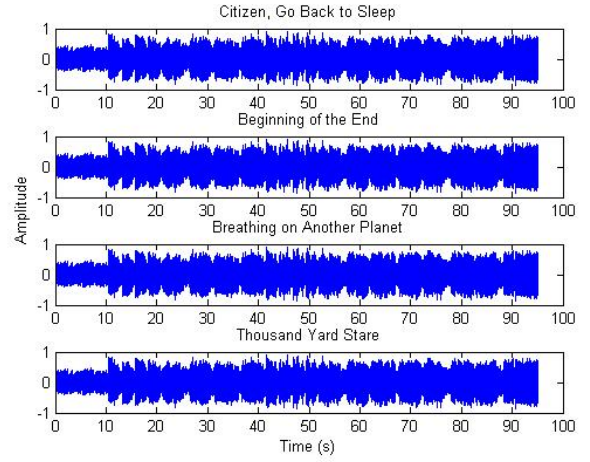


Fig. 4. The signal of original audio

$$BER(W, W^*) = \frac{\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} |W(i, j) - W^*(i, j)|}{M_1 X M_2} \times 100\%$$

where $W(i, j)$ and $W^*(i, j)$ are the original watermark and extracted watermark respectively, and $M_1 X M_2$ is the size of watermark image. In general, BER is expressed in terms of percentage.

To determine the probability of declaring a non-watermarked audio as a watermarked audio and declaring a watermarked audio as non-watermarked audio, false positive error and false negative error are calculated respectively. If these values tend to zero, then it indicates that the system is better. The false positive error probability and false negative error probability are employed respectively:

$$P_{fp} = 2^{-k} \sum_{t=\text{ceil}(0.8k)}^k C(k, t)$$

$$P_{fn} = \sum_{t=0}^{\text{ceil}(0.8k)-1} [C(k, t) p^t (1-p)^{k-t}]$$

Where k is the total number of bits, t is the number of matching bits, $C(k, t)$ is the binomial coefficient of k and t , ceiling is a function which returns smallest integer greater than or equal to a number, and p is the bit error probability of extracted watermark.

A. Imperceptibility Analysis

The watermarked audio is identical to original one because the watermark is, actually, embedded into the secret key, not into the audio. So without processing with various imperceptibility analysis methods, it is said to be an imperceptible scheme.

B. Robustness Test and Analysis


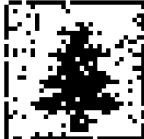


For evaluating the robustness of the proposed method, nine different types of attacks are employed to original audio as following:

- 1) Noise addition: 20 dB additive white Gaussian noise (AWGN) is added to the original audio signal.
- 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: Cut-off frequency is 11.025 kHz.
- 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 32 kbps is applied to the original signal.
- 8) MP3 compression: MPEG-1 layer 3 compression with 64 kbps is applied to the original signal.
- 9) MP3 compression: MPEG-1 layer 3 compression with 128 kbps is applied to the original signal.

GoldWave [22] was used in this experiment for above mentioned attacks except attack 1 and 4. Those were implemented using MATLAB [23].

Table I and II show the values of NC and BER for the proposed watermarking method in terms of robustness against several kinds of attacks applied to four different types of audio signal. In table I, extracted watermark images are also shown. We observe that extracted images almost look like the original image. It is seen that all NC values are ranges from 0.90 to 0.92. Also, all BER values are among 9% to 12%. By considering this performance, we can say that this scheme has good robustness against various kind of attack.

Table I
The result of attacks

Attack	Citizen, Go Back to Sleep		
	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	

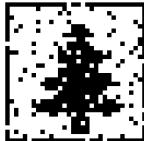

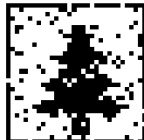
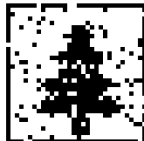
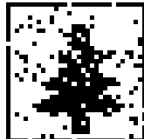
Echo	0.9268	9.28	
Reverse	0.9199	10.16	
mp3 compression (32 kbps)	0.9225	9.77	
mp3 compression (64 kbps)	0.9279	9.18	
mp3 compression (128 kbps)	0.9271	9.28	

Table II
The result of attacks

Audio Signal	Attack	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 III and IV give the comparative results of proposed scheme with other watermarking schemes for audio *Citizen*, *Go Back to Sleep* in terms of NC and BER respectively.

Table III
Comparison of NC among proposed and other algorithm

Attack	NC[Proposed]	NC[17]	NC[11]	NC[13]
1	0.9206	0.9901	0.5870	0.6619
2	0.9158	0.9950	0.9924	0.7847
3	0.9145	0.9942	0.6272	0.5541
4	0.9172	1.0000	0.9634	0.5454
5	0.9268	0.7559	0.5820	0.7867
6	0.9199	0.4904	0.5737	0.6150
7	0.9225	0.9591	0.5774	0.4034
8	0.9279	0.9650	0.5673	0.4806
9	0.9271	0.9650	0.5524	0.7270

For attack 5 and 6, i.e., adding echo and doing the reverse, the proposed scheme clearly shows the best result comparing with other schemes. For attacks of adding noise, low-pass filtering and three different kinds of mp3 compression, this scheme produces better NC values than [11] and [13]. Except attack 5 and 6, [17] clearly indicates that it is better than proposed scheme. But it should be mentioned that NC values greater than 90% indicate its robustness against several attack.

Table IV
Comparison of BER among proposed and other algorithm

Attack	BER[Proposed]	BER[17]	BER[11]	BER[13]
1	10.06	1.17	47.85	37.50
2	10.64	0.59	0.98	25.00
3	10.84	0.68	43.65	50.00
4	10.45	0.00	4.69	49.00
5	9.28	27.83	49.02	25.00
6	10.16	56.64	48.63	43.75
7	9.77	4.79	49.51	62.50
8	9.18	4.10	49.71	56.25
9	9.28	4.10	52.25	31.25

BER is used to evaluate the watermark detection accuracy after different kinds of signal processing operations and the lesser the value of BER, the better the scheme. A value of 0 or very close to zero is required for good robustness.

The proposed scheme clearly shows the best result comparing with other schemes for attacks of adding echo and doing reverse operation. For attacks of adding noise, low-pass

filtering and mp3 compression, this scheme produces better NC values than [11] and [13] but shows comparatively worse result than [17]. And table IV shows that values of [17] clearly superior to the proposed scheme in term of BER except attack 5 and 6.

Fig. 5 and Fig. 6 show the comparative graphical representation of proposed and other algorithms in terms of NC and BER for table III and IV respectively.

Here, each attack is placed on the horizontal axis and value of NC and BER for each attack for different schemes is placed in the vertically against respective attack. From these two graphical charts, we can easily observe which scheme is better or worse against different attacks. For example, from Fig. 6, we can see that, for attack number 6, i.e., doing the reverse of audio, proposed scheme produces below 15% of BER, whereas other three schemes produce above 40% of BER, which indicates the outperformance of the proposed scheme.

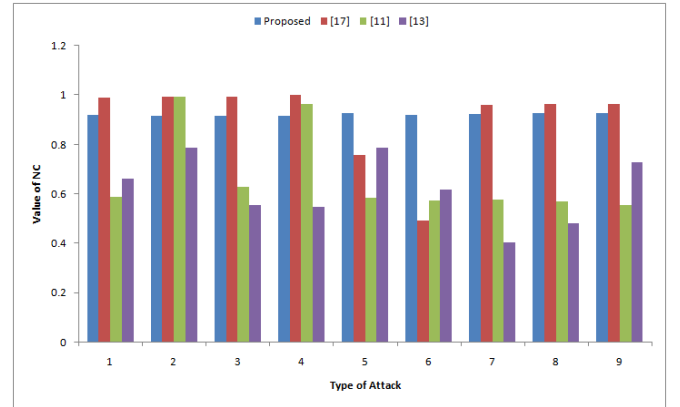


Fig. 5. Graphical representation of NC comparison

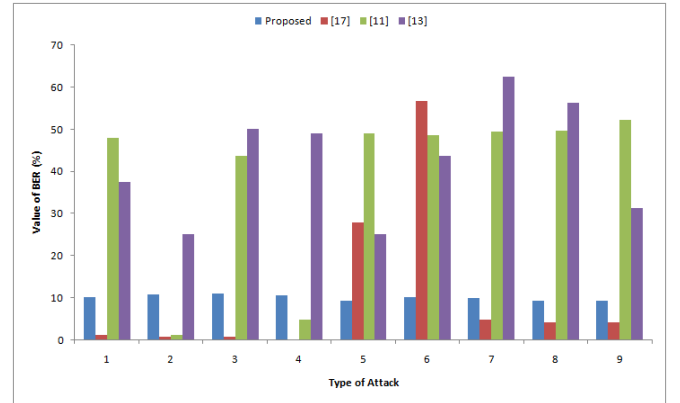


Fig. 6. Graphical representation of BER comparison

Fig. 7 gives the false positive probabilities when watermarking bit (q) belongs to (0, 70], and it tells us that the false positive probability trends to 0 when q is bigger than 15. In this proposed scheme, $q = 1024$, therefore, the false positive probability is approximately equal to 0.

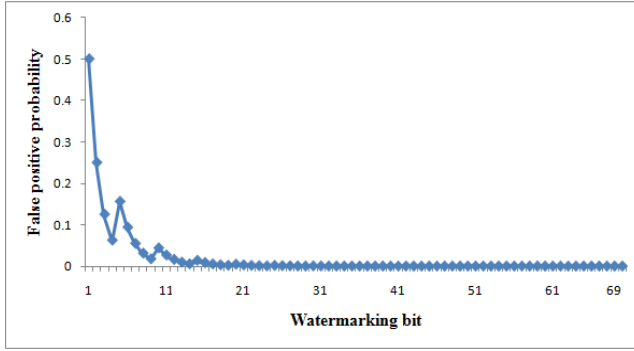


Fig. 7: False positive probabilities under various watermarking bit

Fig. 8 gives the false negative probabilities when q belongs to $(0, 70]$, and it tells us that the false negative probability tends to 0 when q is bigger than 10. In this proposed scheme, $q = 1024$, so, the false negative probability of the proposed scheme is also approximately equal to 0. From tables I, and II, it is easily seen that the BERs are all less than 13%, so the value of p is assumed to be 0.87 in this proposed scheme and put this value in the false negative equation.

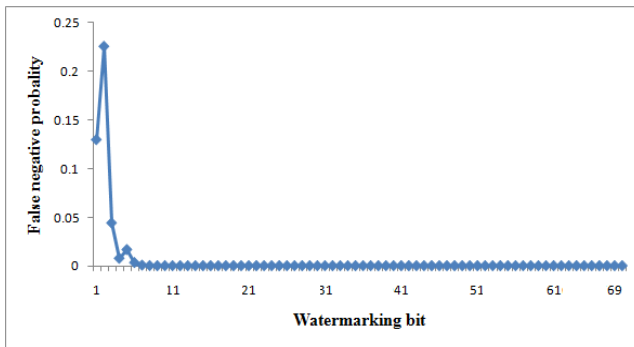


Fig. 8: False negative probabilities under various watermarking bit

VI. CONCLUSION

This paper presents a new zero-watermarking algorithm based on DHT. It does not modify the original audio signal, ensuring the imperceptibility of the watermarking. The proposed method provides normalized coefficient (NC) values ranging from 0.90 to 0.92 and also provides bit error rate (BER) values ranging from 9% to 12%. Therefore, the proposed scheme provides better robustness against various attacks such as re-sampling, adding echo, compression, etc. Comparison of results with other schemes shows that the proposed scheme can be a good competent in the field of audio watermarking.

REFERENCES

- [1] P. Bassia, I. Pitas, and N. Nikolaidis, "Robust audio watermarking in the time domain," *IEEE Transactions on Multimedia*, vol. 3, no. 2, pp. 232–241, 2001.
- [2] A. Lemma, J. Aprea, and W. Oomen, "A temporal domain audio watermarking technique," *IEEE Transaction on Signal Processing*, vol. 51, no. 4, pp. 1088–1097, 2003.

- [3] J. Singh, P. Garg, and A. De, "Multiplicative watermarking of audio in DFT magnitude," *International Journal on Multimedia Tools and Applications*, vol. 71, no. 3, pp. 1431–1453, 2014.
- [4] G. Zhao, "A new digital audio watermarking algorithm based on DWT," in *Proceedings of the International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC)*, pp. 6513–6516, 2011.
- [5] S. T. Chen, G. D. Wu, and H. N. Huang, "Wavelet domain audio watermarking scheme using optimization based quantisation," *Journal of IET Signal Processing*, vol. 4, no. 6, pp. 720–727, 2010.
- [6] Y. Yang, R. Huang, and X. Mintao, "A novel audio watermarking algorithm for copyright protection based on DCT domain," in *Proceedings of the International Symposium on Electronic Commerce and Security*, pp. 184–188, 2009.
- [7] 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 *Proceedings of the International Conference on Consumer Electronics, Communications and Networks*, pp. 871–876, 2012.
- [8] 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.
- [9] W. Quan, T. Sun, and S. Wang, "Zero-watermark watermarking for image authentication," in *Proceedings of the International Conference Signal and Image Processing*, pp. 503–508, 2002.
- [10] 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, 2012.
- [11] 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.
- [12] 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.
- [13] H. L. Dai and D. He, "An efficient and robust zero-watermarking scheme for audio based on DWT and DCT," in *Proceedings of the Asia Pacific Conference on Postgraduate Research in Microelectronics and Electronics*, pp. 233–236, 2009.
- [14] 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.
- [15] R. W. Ciptasari, "An efficient key generation method in audio zero-Watermarking," in *Proceedings of the International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIHMSP)*, pp. 336–339, 2011.
- [16] Y. Xiong and R. Wang, "An audio zero-watermark algorithm combined DCT with zernike moments," in *Proceedings of the International Conference on Cyberworlds*, pp. 11–15, 2008.
- [17] 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.
- [18] R. N. Bracewell, "The discrete hartley transform," *Journal of Optical Society of America*, vol. 73, no. 12, pp. 1832–1835, 1983.
- [19] G. E. J. Bold, "A comparison of the time involved in computing fast hartley and fast fourier transforms," *Proceedings of the IEEE*, vol. 73, no. 12, pp. 1863–1864, 1985.
- [20] V. Bhat, I. Sengupta, and A. Das, "An audio watermarking scheme using singular value decomposition and dither-modulation quantization," *Multimedia Tools and Applications*, vol. 52, no. 2, pp. 369–383, 2011.
- [21] [Online]. Available: <http://www.jamendo.com/en/album/7365>
- [22] [Online]. Available: <https://www.goldwave.com/> (Accessed on: Oct. 28, 2015)
- [23] [Online]. Available: <https://www.mathworks.com/products/matlab/>