

A Novel Spread Spectrum Digital Audio Watermarking Technique

Yekta Said Can, Fatih Alagoz, and Melih Evren Burus

Abstract—Watermarking has become important in the last decade because of the copyright protection applications. Embedding information into an audio file is more difficult as compared to images because human auditory system is more sensitive than human visual system. Therefore, the proposed watermarking algorithms for digital audio have been less than those for digital image and video. This paper presents a digital audio watermarking scheme based on spread spectrum technique to embed the watermark. This method does not need the original audio carrier signal when extracting watermark using the blind extraction audio watermark. The experimental results demonstrate that the embedding technique is not only less audible but also more robust against the common signal processing attacks like low-pass filter, adding white Gaussian noise, shearing, and compression.

Index Terms—Audio signals, spread-spectrum, audio watermarking, blind extraction.

I. INTRODUCTION

With the widespread usage of the internet, illegal copying and distribution have become very common. Multimedia documents are transmitted through the internet but ownership of these documents is generally unknown. Digital watermarking has been used in recent years to embed imperceptible signatures or messages into multimedia contents, for the purpose of copyright protection, distribution control, integrity verification, broadcast monitoring, data hiding, and other aspects of multimedia security [1]. Digital watermarking has been proposed in recent years mainly to protect ownership information in multimedia documents without sacrificing their usability [2].

Three requirements [3] which are necessary for any audio watermarking scheme are: inaudibility, robustness and data rate. Inaudibility means that embedding should be performed in such a way that there is a little difference between original signal and watermarked signal. Therefore, inaudibility is more important metric than the others for any audio watermarking scheme. Other two requirements, that is, robustness and data rate are also important but these two cannot be achieved at the same time [4]. So, in this paper, particular importance is attached to robustness by making a little sacrifice from the data rate.

Digital watermarking techniques can be mainly examined into two categories: blind and non blind approach. A watermarking scheme which does not require original signal

during extraction process is known as blind watermarking scheme [4]. Blind approach is used in this paper.

These two types of watermarking techniques can be used either in time domain or in transform domain. In time domain approach [5], information is embedded directly into the amplitudes of the audio signal [4]. The second approach is the transform domain approach. Transform domain watermarking can be classified into three types: Frequency domain, Discrete Cosine Transform (DCT) domain, Discrete Wavelet Transform (DWT) domain. In transform domain approaches [6], [7], the first step is the transformation of host signal into the specified domain and after that embedding information is carried out into the transformed media.

Spread spectrum technique is one of the most important techniques for hiding secret information into an audio signal. Spreading nature of this technique makes the detection of embedded information is almost impossible for malicious user. There are two implementations of this technique: either using direct sequence spread spectrum approach or frequency hopping spread spectrum. In the direct sequence spread spectrum approach, Exclusive-OR operation is used for spreading the embedded information whereas in the frequency hopping spread spectrum technique, spreading is performed by using the hopping sequence [4].

In this paper frequency hopping spread spectrum technique is applied in time domain. Length five hopping sequence is used and watermark is embedded to the bit of audio that hopping sequence specifies. The results illustrated that our method increases robustness against common signal processing attacks as compared to transform domain approaches.

Organization of this paper is like the following. In the next section, the details of the watermarking embedding and extraction procedures of the proposed algorithm will be explained. In section three, applied common signal processing attacks are explained. In section four, performance of the algorithm is evaluated and experimental results of the inaudibility and robustness requirements are presented. In the last section, there will be some important remarks about the proposed algorithm.

II. THE AUDIO WATERMARKING METHOD

In this paper, a frame based audio watermarking scheme using spread spectrum technique in time domain is proposed. The algorithm consists of two procedures; watermarking embedding procedure and watermarking extraction procedure respectively.

The watermarked image is $M \times N$ pixels. The audio that watermark is embedded is a wav file. In order to embed the image to the audio file, the image and audio should be

Manuscript received June 15, 2013; revised October 2, 2013.

Yekta Said Can is with Boğaziçi University, Computer Eng. Dept., Bebek 34342 Istanbul, Turkey. He is also with Tübitak-BİLGEM Kocaeli, Turkey (e-mail: yektasaid.can@gmail.com).

Fatih Alagoz and Melih Evren Burus are with Boğaziçi University, Computer Eng. Dept., Bebek 34342 Istanbul, Turkey.

preprocessed.

A. Watermark Preprocessing

1) Represent the grey-scale image watermark as a two dimensional $M \times N$ matrix. Each value of this matrix represent the pixel value which is between 0-255. `Imread()` method at MATLAB is used.

2) Convert this matrix into a one dimensional array. The rows will be concatenated after each other. This created array has $M \times N$ length.

3) Zeros and ones are needed in order to apply spread spectrum to embed different parts of audio bytes. First convert the pixel value into 8 bit binary number. Then, a new array to hold this zeros and ones is created. This new watermark array has a length that 8 times the old one. This new array's elements will be embedded to audio.

4) Frequency hopping spread spectrum technique is used. Hopping sequence is necessary for this purpose. Length five hopping sequence is created where every element of this sequence is an integer between 1-8. This sequence will tell us which bit to embed the watermark. 'n' means embed the watermark to the nth least significant bit.

5) Convert an audio signal into byte array.

B. Watermark Embedding

1) Determine the position that the watermark bit will be embedded by using hopping sequence. This hopping sequence will be repeated. For example if sixth bit will be embedded and the length of the hopping sequence is five, this bit will be embedded to the position that first element of hopping sequence specifies. $(6 \bmod 5)$

2) Obtain the embedded watermark bit from binary array.

3) Embed every one byte in four bytes of the audio file.

4) If the element is '0' use bitwise 'AND'. Otherwise use bitwise 'OR' with the audio byte arrays specified bit. This means replacing the specified bit in the audio byte with the watermark bit. For instance if the embedded bit is '1', embedding will be done as follows:

```
audioBytes(i)=bitor(audioBytes(i),2n-1)
If the bit is zero:
audioBytes(i)=bitand(audioBytes(i),255-2n-1)
```

where `audioBytes` is the audio array, `n` is the bit's embedding position.

5) Obey the hopping sequence periodically until the watermark bit array finishes.

C. Watermark Extraction

1) In the watermarked audio byte array, embedded bits location is obtained by using hopping sequence.

2) Retrieve the embedded bit by applying bitwise AND operation with the binary number which has all zeros but a one in the embedding position.

```
embeddedBit=bitand(audioBytes(i), 2n-1) / 2n-1
```

3) Put the retrieved bits into bit array.

4) Convert these binary representations into decimal representations. This operation will result in the pixel value array ($M \times N$).

5) Convert pixel value array as an $M \times N$ two - dimensional matrix.

6) Represent this resulting 2D matrix as a gray scale image. `Mat2gray()` method in MATLAB is used.

III. SIGNAL PROCESSING ATTACKS

To evaluate the robustness of the watermarking algorithm common signal processing attacks are applied. [8] The watermarked material should be similar to the original embedded one in order for our algorithm to be successful.

A. Low Pass Filter

Watermark signal is filtered with Butterworth low-pass filter which is second order and has cutoff frequency is $= 0.5$. This filter is implemented by using MATLAB built-in function.

`[b, a] = butter (n, Wn)` designs an order `n` low pass digital Butterworth filter with normalized cutoff frequency `Wn`.

B. Shearing

Shearing is also another common signal processing attack. During transmissions, the part of transmitted audio might be lost due to channel conditions, jamming and so on. So, this attack is implemented on MATLAB, by assigning some part of watermarked audio to zeros. However, since spread spectrum is spreading the signal into more time, when applying this attack, shearing time is hold longer on spread spectrum than the DWT method.

C. Lossy Compression

Lossy compression is another most commonly used signal processing attack. It compresses data by discarding (losing) some part of it. The metric that determines how much compression affects data is the compression ratio. In other words, it means how much of the data is discarded during compression. Lossy compression with compression ratios 12.5, 25, 37.5, 50% is applied.

This attack is implemented by cutting the least significant bits then replacing them with random bits. For example, if compression ratio is 25%, then the last two bits of a byte are removed then randomly replaced.

D. Additive White Gaussian Noise

The last common signal processing attack is adding AWGN to watermarked audio signal. The metric that determines how strong this attack is the signal to noise ratio. This attack is implemented on MATLAB by using the built-in function.

$$Y = \text{AWGN}(X, \text{SNR})$$

IV. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed audio watermarking algorithm, 90 seconds mono speech file was used. The clip was sampled at a sampling rate of 88000 sample per second. Each sample was represented by 8 bits.

128x128 gray-scale Lena image was used as a watermark for embedding in the audio clip. The size of the image watermark was constant. The audio clip size was adjusted so that watermarking has finished after the image bytes and audio has finished. The metrics that have been used to

evaluate the performance of proposed audio watermarking algorithm are inaudibility and robustness.

A. Inaudibility

Inaudibility(imperceptibility) is determined by the perceptual difference of the embedded watermark data from the original audio signal. The goal of inaudibility is that the quality of host signal is not corrupted much, so that the listener can not perceive the difference. To measure imperceptibility, Signal-to-Noise Ration (SNR) is used as an objective measure. Signal to Noise Ratio (SNR) is a difference metric that is used to calculate the similarity between the original audio signal and the watermarked audio signal. The SNR computation is carried out according to equation (1), where I_n is the original audio signal, and E_n corresponds to the watermarked audio signal [5].

$$SNR(dB) = 10 \log \frac{\sum_n I_n^2}{\sum_n (E_n - I_n)^2} \quad (1)$$

Audio SNR results of Spread Spectrum method with the Al-Haj *et al.* 2009 DWT is compared and also implemented on MATLAB. As it can be observed from Table I, SNR of host audio file is higher at Spread Spectrum technique. This means that Spread Spectrum technique is more inaudible.

TABLE I: SNR OF WATERMARKED AUDIO

	SNR of Watermarked Audio
Al-Haj <i>et al.</i> 2009 DWT	56.26 dB
Our Spread Spectrum Tech.	65.46 dB

B. Robustness

Common signal processing attacks such as linear filtering, lossy compression, shearing and AWGN among many others are implemented on watermarked audio digital signal. Even though effects of these attacks might not be perceived from the quality of the host signal, they corrupt the embedded watermark image more.

We measured the similarity between the original watermark and the watermark extracted from the attacked watermarked images using the correlation factor ρ , which is computed as shown in equation below [5]:

$$\rho(w, e) = \frac{\sum_n (w_n * e_n)}{\sqrt{\sum_n w_n^2} * \sqrt{\sum_n e_n^2}} \quad (2)$$

where e 's are extracted image pixels, w 's are watermarked image pixels.

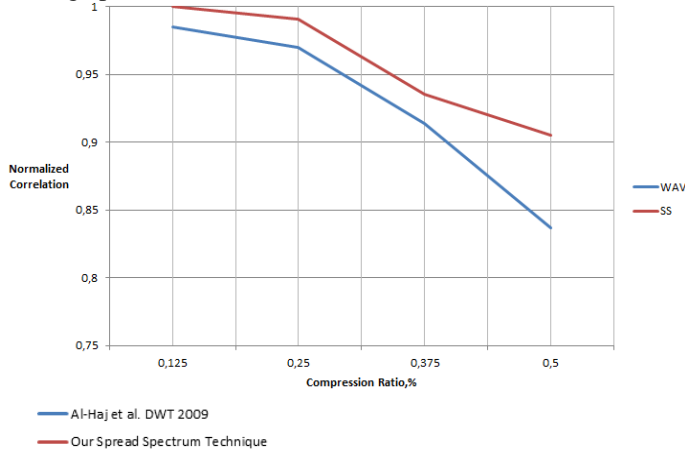


Fig. 1. Correlation value versus compression ratio (%).

C. Compression Attack

Compression attack is applied on both our Spread Spectrum algorithm and Al-Haj *et al.* DWT. Vertical axis shows the correlation factor. Horizontal axis shows compression ratio. The correlation factor is measured as in the Fig. 1.

Spread Spectrum method's correlation values are always higher as the compression ratio increases. In addition, as it can be observed from the Fig. 1, the decreasing rate of DWT method is accelerating as the compression ratio increases. On the contrary the slope of SS method is decreasing because the Spread Spectrum algorithm works better at more noisy environments than its competitor.

D. AWGN Attack

Additive White Gaussian attack is applied on both our Spread Spectrum algorithm and Al-Haj *et al.* DWT. Vertical axis shows Normalized Correlation, whereas horizontal axis shows the SNR value of AWGN.

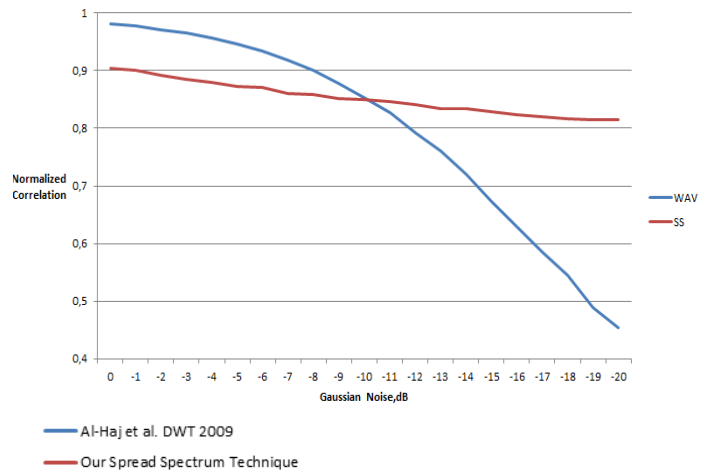


Fig. 2. Normalized correlation versus gaussian noise (dB).

The correlation factor of DWT method starts higher than Spread Spectrum method (near 1) when AWGN is 0dB. When the AWGN SNR approaches -10dB, correlation factors of the two methods are starting to become closer. They become equal at -10dB. After that, our Spread Spectrum method's correlation factors are higher than the DWT's.

It can be seen from Fig. 2 that Spread Spectrum method's correlation factors are smaller at low noises. The decrease in correlation factors as noise increases is approximately linear. On the contrary, the DWT method's correlation factors start with higher values at low noise. The decrease in correlation factors as noise increases is approximately exponential. It can be concluded that our method is more robust against AWGN attack more than -10dB. Our method works better at noisier environments than DWT.

E. Shearing and Low Pass Filter

Shearing and Butterworth attacks are also applied. Since the Spread Spectrum spreads the information 8 times more than the other technique, the shearing (cut) time is hold 8 times more than DWT technique.

Second order and 0.5 cutoff frequency Butterworth low pass filter is used. The correlation factor results are given in Table II.

TABLE II: NORMALIZED CORRELATION WITH SHEARING AND LPF ATTACKS

	Shearing	Butterworth LPF
Al-Haj <i>et al.</i> 2009 DWT	0.2542	0.1203
Our Spread Spectrum Tech.	0.9917	0.7642

From Table II, it can be observed that for shearing and low pass filter attacks, Spread Spectrum technique outperforms the DWT technique with regarding robustness.

V. CONCLUSION

A spread spectrum based audio watermarking technique on time domain is presented in this paper. Some of the existing methods are implemented and compared with proposed one. Experiments have shown that the inaudibility and robustness performance goals can be achieved. When we evaluate the robustness performance against signal distortions, the algorithm definitely outperformed its competitor.

Another advantage of the proposed method to its competitor is blindness. This method does not need the original audio file to extract watermark.

Finally, as it can be observed at compression and AWGN attacks, the difference between the proposed algorithm and its competitor becomes more clear when attack's power is increased. Our algorithms works much better than its competitor at noisy environments.

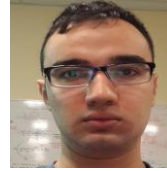
ACKNOWLEDGMENT

This project was supported in part by Argela and Turk Telekom Research Collaboration Award Program.

REFERENCES

- [1] P. Zhang, S. Xu, and H. Yang, "Robust and transparent audio watermarking based on improved spread spectrum and psychoacoustic masking," in *Proc. 2012 International Conference on Information Science and Technology (ICIST)*, 23-25 March 2012, China, pp. 640-643.
- [2] I. Cox, M. Miller, and J. Bloom, *Digital Watermarking*, Academic Press, USA, 2002.
- [3] N. Cvejic, "Algorithms for audio watermarking and steganography," Ph.D. dissertation, Department of Electrical and Information Engineering, Information Processing Laboratory, University of Oulu, 2004.
- [4] N. Baranwal and K. Datta, "Comparative study of spread spectrum based audio watermarking techniques," in *Proc. 2011 International Conference on Recent Trends in Information Technology (ICRTIT)*, pp. 896-900, 3-5 June 2011.

- [5] 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.
- [6] A. Al-Haj, L. Bata, and A. Mohammad, "Audio watermarking using wavelets," in *Proc. First International Conference on Networked Digital Technologies, 2009. NDT '09*, 28-31 July 2009, pp. 398-403.
- [7] M. Yiping and H. Jiqin, "Audio watermark in dct domain Strategy and algorithm," *Chinese Journal Electronics*, pp. 1260-1264, 2003.
- [8] B. Chen, Y. Hao, R. Liu, G. Hu, "The audio watermarking system based on wavelet transform algorithm," in *Proc. Cross Strait Quad-Regional Radio Science and Wireless Technology Conference (CSQRWC)*, 26-30 July 2011, vol. 2, pp. 1274, 1277.



Yekta Said Can was born in Bursa, Turkey on June 18, 1990. He studied Computer Engineering in Faculty of Engineering, Boğaziçi University, İstanbul, Turkey where he obtained his B.Sc Degree in 2012. He is pursuing an M.Sc degree right now in Computer Engineering at Boğaziçi University, İstanbul, Turkey. He also works as a researcher in TUBITAK His research interest includes watermarking, target tracking, speech and speaker recognition.



Fatih Alagöz received the BSc degree in electrical engineering from Middle East Technical University, Turkey, in 1992, and the MSc and DSc degrees in electrical engineering from George Washington University in 1995 and 2000, respectively. He is currently a professor in the Department of Computer Engineering, Bogazici University, İstanbul, Turkey. From 2001-2003, he was with the Department of Electrical Engineering, United Arab Emirates University. In 1993, he was a research engineer in a missile manufacturing company, Muhimmatsan AS, Turkey. His current research interests include satellite networks, wireless networks, sensor networks, and UWB communications. He has contributed to 10 research projects for organizations including the US Army of Intelligence Center, US Naval Research Laboratory, UAE Research Fund, Turkish Scientific Research Council, State Planning Organization of Turkey, and BAP. He has edited five books and published more than 100 scholarly papers. He is the satellite systems advisor to the Kandilli Earthquake Research Institute, İstanbul, Turkey. He has served on many technical committees and organized many technical sessions at international conferences. He has numerous professional awards. He is a member of the IEEE Satellite and Space Communications Technical Committee.



Melih Evren Buruş studied Chemical Engineering in Faculty of Engineering, Boğaziçi University, İstanbul, Turkey where he obtained his B.Sc Degree in 2004. He has got his M.Sc degree in Software Engineering at Boğaziçi University, İstanbul, Turkey in 2012. He is a research assistant now at Boğaziçi University. His research interest includes digital signal processing, information hiding, object recognition.