

EFFECTIVE ROBUST PATCHWORK METHOD TO THE VULNERABLE ATTACK FOR DIGITAL AUDIO WATERMARKING

Y.D. Chincholkar and Shalaka Pravin Kude

Department of Electronics and Telecommunications Engineering, Savitribai Phule Pune University, India

Abstract

This paper presents patchwork based digital audio watermarking. The advanced growth in transmission of digital data has resulted in a corresponding elevation in the need for copyright protection of signal. Cryptography and steganography are used for the content protection but do not completely solve the copyright issue. Watermarking is a method to protect and identify the digital data while maintaining the quality of the host media, it permits various types of watermarks to be hidden in audio signal e.g. image, audio and video. This paper limits on image embedding technique using patchwork-based method. In patchwork based method average of all segments of approximate coefficients is calculated for embedding watermark into sound signal. The experimental results shows that proposed method achieves imperceptibility for audio signal as watermarked audio signal is inaudible after embedding watermark and robustness of watermark against different signal processing attacks with higher PSNR. The resulting audio is robust to attacks and exhibits good quality in term of peak signal to noise ratio. The simulation results show the effectiveness of the proposed system.

Keywords:

Audio Watermarking, Stationary Wavelet Transform, Patchwork

1. INTRODUCTION

In recent years, audio watermarking techniques have achieved several good algorithms and significant progress for inserting watermarks into audio data to authenticate copyright. In recent years, there has been augmenting research interest in the protection of digital media concerning intellectual property rights. Digital watermarking is an optimistic technique to solve this problem. From the specialized facet, digital watermarking aims to insert personal data (such as signature, logo, ID number, biometric image etc.) into the host signal. When required, the owners can extract these data to declare their copyright. The digital watermarking can be applied to various multimedia data such as audio [1], image [2], and video [3].

There are two main routes for audio watermarking: Time domain and Transform domain [5], [6]. Watermark Information is directly embedded into original audio in time domain while in frequency domain watermark is encoded by modifying the transform output coefficient. Transform domain techniques found more robust and efficient over time domain. DCT (Discrete Cosine Transform), DFT (Discrete Fourier Transform), DWT (Discrete wavelet Transform), these are the popular transform. DWT is partial to other transform for digital audio watermarking because it gives time and frequency representation of signal. It is facile to implement and provide more information for analysis and synthesis of signal while DCT shows drawback of blockage artifacts in image. This paper limits its attention to audio watermarking. An audio signal contains only one-dimensional data and the human auditory system is too sensitive than the visual

system, it is difficult to hide extra information into host audio signal than into other multimedia data, without reducing the quality of the media signal. Digital audio watermarking by using SWT algorithm shows the moderate robustness and imperceptibility [5].

The main goal of this system is to implement effective and robust audio watermarking algorithm based on patchwork method to prevent copyright offense and intellectual property of original multimedia data. Algorithm possesses properties like robustness, high perceptual quality, security and satisfy IFPI (International Federation of the Phonographic Industry) requirement of PSNR more than 20dB. Encoded watermark reacquired against different signal processing attacks like amplitude, scaling, additive white noise, re-sampling and echo-add.

An effective and practical audio watermarking method should possess three important traits: imperceptibility, robustness, and security. Imperceptibility means that the watermark should be inaudible in the watermarked audio signal. Robustness indicates the ability to reacquire the watermark data from the watermarked signal, both in the situations with and without attacks. Security ensures that an unauthorized user cannot extract or delete the watermark.

The remaining of the paper is categorized as follows. Section 2 shows the literature review. Section 3 shows the embedding and decoding of the subjecting watermarking method, respectively. The simulation results are shown in section 4. Section 5 concludes the proposed system.

2. LITERATURE REVIEW

In recent years, many audio watermarking methods have been developed, which employ various techniques such as spread spectrum [1] echo-hiding [4], support vector regression [8], and patchwork [9], [10]. Among these audio watermarking methods, the patchwork based methods have shown high robustness against many common attacks, including noise, filtering, compression, re-quantization, and re sampling attacks. The concept of patchwork was first introduced in image watermarking [2] and then extended to audio watermarking [1]. The patchwork watermarking method first applies the Fourier transform to each host audio segment to obtain the corresponding Fourier transform coefficients and then divides the Fourier transform coefficients into groups, called patches. Based on an embedding rule, a watermark bit, which is either “0” or “1”, is embedded into the audio segment by modifying the patches. In [5], Yeo and Kim proposed a modified patchwork algorithm in the discrete cosine transform (DCT) domain to improve watermarking performance. To correctly extract the embedded watermarks, the methods in [9] and [11] require the patches to have similar statistical characteristics. Otherwise, they cannot achieve 100% detection

rate (DR) even in the scenario of no attacks. However, the requirement of similar statistical characteristics cannot be guaranteed as the patches are of finite length in practice. In order to overcome the drawback of the methods in [9] and [11], Kalantari et al. proposed a multiplicative patchwork method in [4]. In this method, two patches are formed from each host audio segment by using its discrete wavelet transform coefficients. A watermark bit is embedded into an audio segment if the corresponding two patches have comparable statistical characteristics. Otherwise, no watermark bit will be embedded into the audio segment. Due to this selective watermark embedding scheme, a large number of audio segments will be deemed as unsuitable for watermark embedding. In other words, watermarks are only embedded into some segments of the host audio signal. Thus, at the decoding end, it is vital to know which audio segments of the watermarked signal contain watermarks. If this information is not available, considerable false watermarks will be “extracted” from the un-watermarked audio segments. Whilst correctly identifying the watermarked segments is essential to the method in [4], it does not provide a mechanism to find this information.

In [10], Natgunanathan et al. proposed a patchwork method which can identify the watermarked segments in the watermark decoding process. In this method, the host audio segment is first partitioned into two sub segments and then a number of frame pairs are formed by randomly selecting some DCT coefficients from the two sub segments. The frame pairs suitable for watermark embedding are chosen by a selection criterion and watermarks are embedded into the selected frame pairs accordingly. An important feature of this method is that the selection criterion used at the embedding stage can be applied at the decoding stage to identify the watermarked frame pairs from the watermarked signal. So, unlike the patchwork method in [4], the method in [12] does not need any additional information to find the watermarked frame pairs at the decoding end. However, all of the aforementioned audio watermarking methods require perfect synchronization between the encoder and decoder. As a result, they are vulnerable to de-synchronization attack, which limits their usage in practical applications. In order to deal with de-synchronization attacks, some synchronization schemes have been developed [12], [13]. In [12], Ko et al. proposed to modify the time-spread echo hiding technique by employing a logarithmic operation in the quefrequency axis in the decoding process. Since logarithmic operation can convert multiplication induced by scaling attacks to addition, the embedded pseudo-noise (PN) sequence only differs from the original one (which is available at the decoding end) by a constant delay in the quefrequency axis. Based on this property, the correlation between the PN sequence and the cepstral component of the watermarked signal is computed in the quefrequency axis to find the scaling factor [12], where the scaling factor is equal to one in the absence of scaling attacks. The correct calculation of the scaling factor is vital for establishing synchronization between the encoder and decoder.

In this paper, we propose a patchwork-based audio watermarking method to resist attacks. In the embedding process, we first apply SWT to the host audio signal to obtain the corresponding SWT coefficients and after that divide the audio signal into patches i.e. segments. Then, we remove those SWT coefficients related to low- and high-frequency components, and segment the remaining SWT coefficients. An improved

patchwork embedding scheme is utilized to embed the watermark bits into the non-silent audio segments by modifying the associated SWT coefficients.

3. PROPOSED AUDIO WATERMARKING

General block diagram of audio watermarking is shown in Fig.1(a) shows the embedding process of audio watermark, in which audio signal is divided into patches and simultaneously watermark image is compressed. By using coder patches of the audio signal and watermark bits are embedded together. At the ending these segments are collected and will combined and form a whole watermarked audio signal. The Fig.1(b) shows the decoding process of audio watermark which worked exactly opposite of that embedding process. In decoding, image will be extracted form watermarked audio clip after that image will converted back into grey scale and, watermark will retrieved.

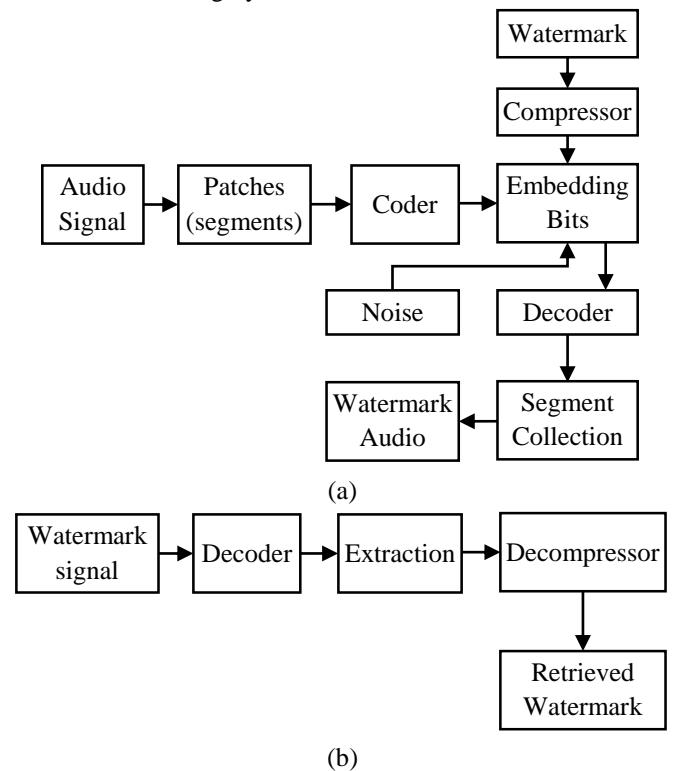


Fig.1. (a) Embedding process (b) Decoding process of System

3.1 EMBEDDING OF WATERMARK

Embedding of watermark to the host audio signal consists of following parts:

3.1.1 Audio Signal:

Audio signal is the signal which is to be protected for the copyright issues which is occurs if that signal isn't protected before uploading. To prevent future problem of ownership that signal is not proven by owner to overcome this issue 'watermark' is the solution. So first block of system is to be implemented is audio signal. Proposed system is worked on basic level so that we have taken audio signal which is in the form of WAV. In proposed system we have worked on a different type of WAV file like jazz song, pop song, normal Bollywood song, classical song, instrumental song. All these signals are taken form database to

prevent copyright authentication and implemented system is applied to these signals.

3.1.2 Watermark Image:

Image is used as a watermark signal. Image can't be directly embedded into audio signal. We are using a grey scale image. This image is a compressed image of original image (32x32) bit size i.e. 1024 bits. This compressed image is grey scale image and is in the form of decimal which is then converted into binary bits.

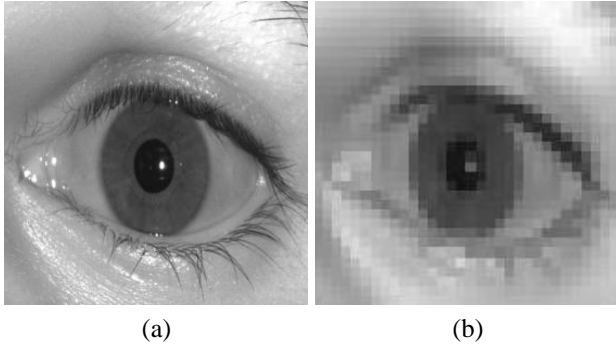


Fig.2. (a) Original image (b) Compressed binary grey scale image

3.1.3 Patchwork (Segmentation):

Audio signal clip is further divided into segments for more security. So that each patch will contain some watermark bits. Before segmentation to enhance the security first we have to go through filtering. For this purpose we are using,

• Stationary Wavelet Transform (SWT):

The Stationary Wavelet Transform (SWT) is a wavelet transform algorithm designed to overcome the lack of translation-invariance of the Discrete Wavelet Transform (DWT). The SWT is an inherently redundant scheme as the output of each level of SWT contains the same number of samples as the input. Embedding of watermark using stationary wavelet transform provides an excellent security. SWT gives a two coefficient detailed and approximated. Among them approximated coefficients are used for embedding platform because they are less vulnerable to attack. SWT is performed on a sound signal with multiple levels. The Fig.3 shows the 3-level SWT.

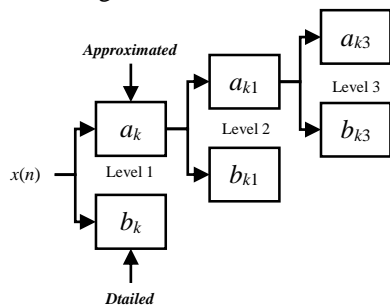


Fig.3. SWT Levels decomposition

Let $\{a_k\}$ and $\{b_k\}$ be the low pass and high pass filter coefficient. Now consider $\{G_{m,n}\}$ is a sequence of approximated coefficients with level m and this sequence is transformed into two sequence i.e. $\{G_{m+1,n}\}$ and $\{H_{m+1,n}\}$ where,

$$\begin{aligned} \{G_{m+1,n}\} &= \frac{1}{\sqrt{2}} \sum_k A_k^m G_{m,n+k} \cdot \{H_{m+1,n}\} \\ &= \frac{1}{\sqrt{2}} \sum_k B_k^m G_{m,n+k} \end{aligned} \tag{1}$$

where $\{H_{m+1,n}\}$ is a detailed coefficient. SWT is redundant, shift invariant and invertible transform.

• Decomposition of Sound Clip:

Original sound clip is decomposed to the 3rd level using Daubechies wavelet in SWT domain. Therefore Ca_3, Cd_3, Cd_2, Cd_1 will get these coefficients after decomposition. Among these coefficients approximate coefficients of 2nd level are selected for encoding the watermark. This coefficient are divided into patches (segment) and average of all segment is calculated as,

$$C_{avg} = \frac{1}{k} \sum_{i \times k}^{(i+1) \times k} C_{a3} \tag{2}$$

where $i = 0$ to 1024 for embedding 32x32 image which is used as a watermark signal and n is length of divided patch then average value plant with image pixel value for embedding. Afterwards inverse SWT is applied to syndicate watermark embedded low frequency coefficients with high frequency coefficients. Output signal is the digital audio watermark signal.

• Patchwork and Embedding Process:

Host sound signal is divided into number of patches (segments). To do framework that means to divide the total length of audio signal into the patches is based upon watermark's bit size. To calculate total number of frame,

$$\begin{aligned} \text{Length of audio signal} &> \text{required signal} \\ \text{Number of frames} &= \text{length of audio} / \text{sample frequency} \\ \text{Required signal} &= \text{required frame} \times \text{sample frequency} \\ \text{Sample frequency} &= \text{total size of watermark bits} \times 100 \end{aligned} \tag{3}$$

The Eq.(3) shows how framing will be done. According to Eq.(3) watermarked bits will embedded in calculated SWT levels of that frame of audio signal. The Fig.4 shows how signal is divided into small segments. Segmentation will be done using SWT. SWT is applied to host signal we will get SWT segments. Watermarks bits are added at the beginning of segments afterwards all segments will be collected by applying inverse of SWT. For embedding the watermarks, we are using image i.e. iris image from database which is in the form of compressed grey scale image and converted into binary image. These binary bits are used as a watermark bits to be embed. The Fig.2 shows the database image.

3.2 DECODING OF WATERMARK

Extraction process is resembled to embedding process and it does not need original sound signal. Knowledge of the pioneer sound clip, the watermark intensity, and the size of the watermark is required in order to yank the watermark. It is semi-blind technique. In watermark extraction process, watermark sound signal is used as an input. It is decomposed into 3rd level SWT. Afterwards average of all low frequency coefficients patches is calculated by using Eq.(2) and each bit of watermark is yank. Acquired watermark is one dimensional; convert it into 2D 32x32 image. Yank image is binary image is converted into decimal format and it is decompressed.

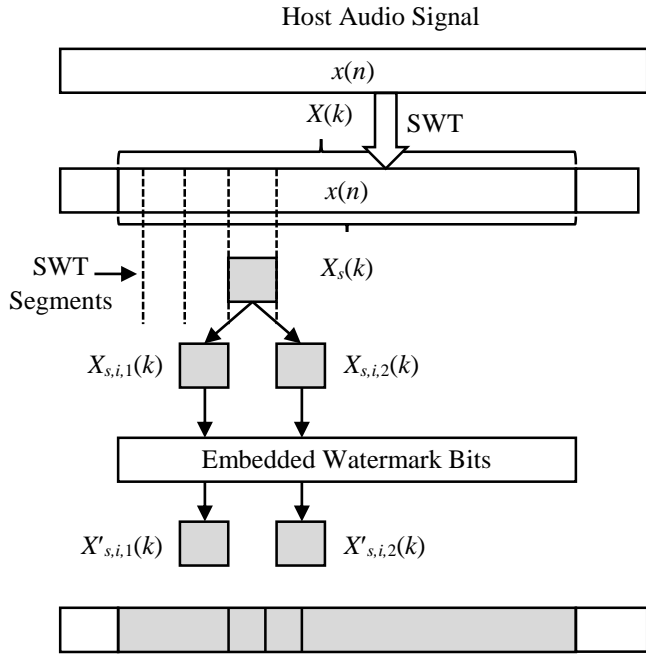


Fig.4. Proposed watermark embedding process

3.3 EVALUATION PARAMETERS

3.3.1 Imperceptibility:

Imperceptibility means that the watermark should be inaudible in the watermarked audio signal. To check this feature parameter SNR is required.

$$SNR = 20 \log_{10} \frac{\text{No. of original bits}}{\text{No. of original bits} - \text{No. of processed bits}} \quad (4)$$

3.3.2 Robustness:

Robustness indicates the ability to reacquire the watermark data from the watermarked signal, both in the situations with and without attacks. To check robustness of the system following parameters are evaluate.

$$PSNR = 10 \log_{10} \frac{32 \times 32}{MSE} \quad (5)$$

where, *MSE* is the mean square error

Common attacks are used to evaluate robustness of proposed system-

- *Noise attack*: Random noise is added to the watermarked signal.
- *Quantization attack*: Each sample is of watermarked signal is quantized from 16 bits to 8 bits.
- *LPF attack*: LPF are applied to watermarked signal, where the cut off frequency is 8 kHz.
- *HP attack*: HPF is applied to the watermarked signal, where the cut off frequency is 50 Hz.
- *Scaling attack*: In scaling pitches of watermark signal are scaled.

4. REPERCUSSION

In this section, the performance of the proposed method against common attacks is illustrated by simulation results. In the simulations, 5 mono-channel audio clips belonging to different genres are used as host audio signals. Each audio clip has duration of 120 seconds, and is sampled at the rate of 44.1kHz. All samples are used to compute the SWT coefficients. Results of proposed system are consisting in this chapter are as follows:

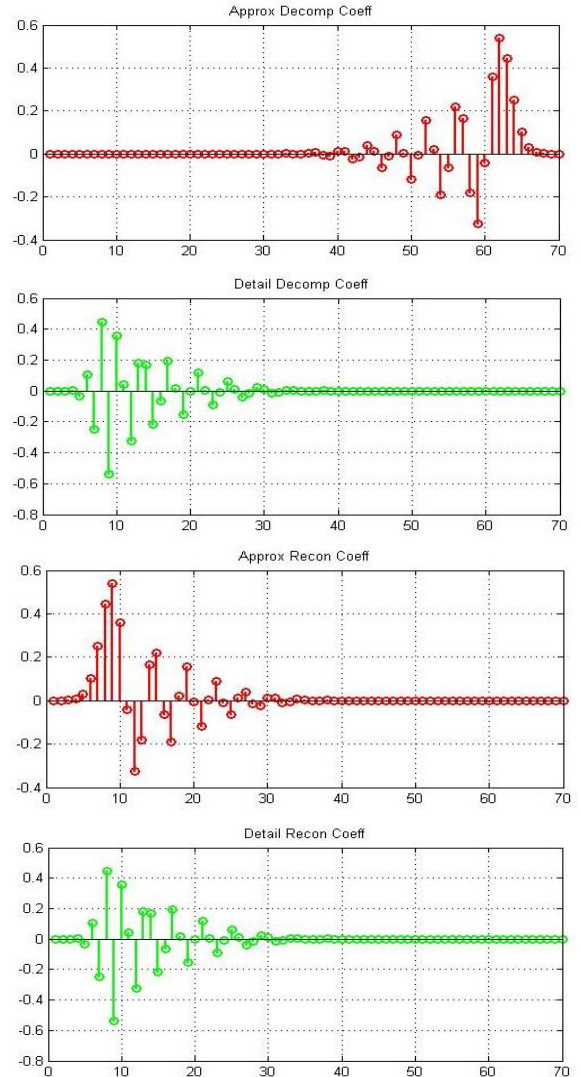


Fig.5. Coefficient estimation of audio signal using SWT

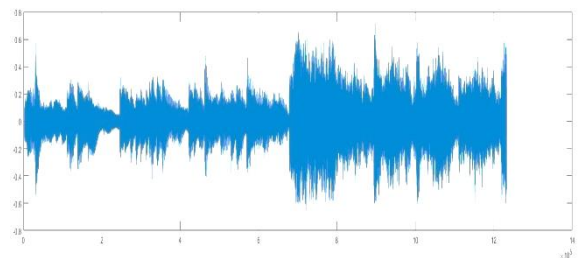


Fig.6. (a) Wave structure of instrumental audio without watermark

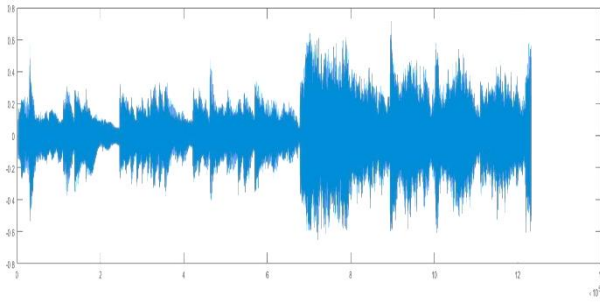


Fig.6. (b) Wave structure of instrumental audio with watermark

Table.1. Performance evaluation parameter

| Signal Types | Parameter | Attack type | | | |
|--------------|-----------|----------------|-------|-----------|--------------|
| | | Without Attack | Noise | Amplitude | Quantization |
| Instrumental | PSNR | 91.72 | 91.71 | 91.70 | 91.71 |
| | SNR | 74.38 | 74.37 | 74.36 | 74.37 |
| | BER | 23.29 | 20.36 | 20.36 | 20.36 |
| Classy | PSNR | 76.86 | 76.88 | 76.85 | 76.88 |
| | SNR | 62.16 | 62.18 | 62.1 | 62.18 |
| | BER | 23.29 | 20.27 | 20.28 | 20.25 |

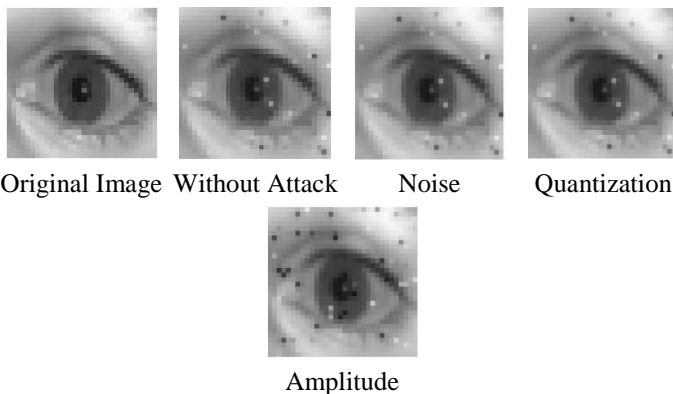


Fig.7. Original image and extracted watermark with and without attack for instrumental audio clip

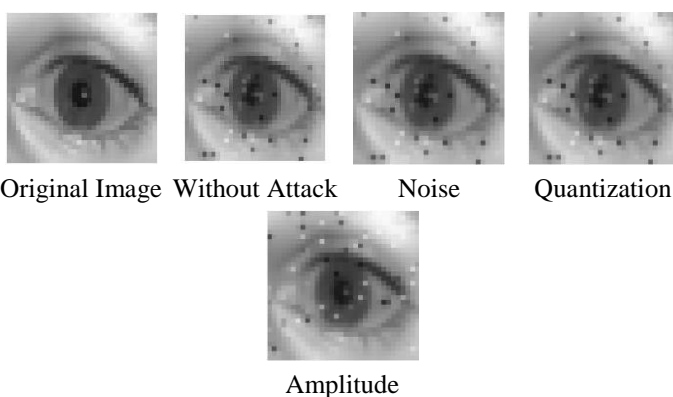


Fig.8. Original image and extracted watermark with and without attack for classy audio clip

5. CONCLUSION

A robust patchwork-based audio watermarking method for digital audio signal is proposed, which is designed to resist assault. It is more robust against common assault and has better perceptual quality than existing DCT based watermarking methods. In the proposed method, watermarks are embedded into the audio signal by modifying its SWT coefficients. Experimental results show that the embedded watermark is perceptually transparent; the performance of proposed method is evaluated by the SNR and PSNR. Imperceptibility and robustness is proved by MSE and PSNR values. This approach is useful in copyright protection application in which the embedded information is related to the owner of the digital audio signal. The proposed technique is robust to conventional attacks such as, noise and filtering (high-pass, low-pass) attacks.

REFERENCES

- [1] Y. Xiang, I. Natgunanathan, S. Guo, W. Zhou and S. Nahavandi, "Patchwork-Based Audio Watermarking Method Robust to Desynchronization Attacks", *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, Vol. 22, No. 9, pp. 1413-1423, 2014.
- [2] Pranab Kumar Dhar and Tetsuya Shimamura. "Blind SVD-based Audio Watermarking using Entropy and Log-Polar Transformation", *Journal of Information Security and Applications*, Vol. 20, pp. 74-83, 2015.
- [3] H. Huang, C. Yang and W. Hsu, "A Video Watermarking Technique based on Pseudo-3-D DCT and Quantization Index Modulation", *IEEE Transactions on Information Forensics and Security*, Vol. 5, No. 4, pp. 625-637, 2010.
- [4] F. Guerrini, M. Okuda, N. Adami and R. Leonardi, "High Dynamic Range Image Watermarking Robust against Tone-Mapping Operators", *IEEE Transactions on Information Forensics and Security*, Vol. 6, No. 2, pp. 283-295, 2011.
- [5] K. Yeo and H.J. Kim, "Modified Patchwork Algorithm: A Novel Audio Watermarking Scheme", *IEEE Transactions on Speech and Audio Processing*, Vol. 11, No. 4, pp. 381-386, 2003.
- [6] W.N. Lie and L.C. Chang, "Robust and High-Quality Time-Domain Audio Watermarking based on Low-Frequency Amplitude Modification", *IEEE Transactions on Multimedia*, Vol. 8, No. 1, pp. 46-59, 2006.
- [7] Y.D. Chincholkar, S.R. Ganorkar and A.S. Sawai, "Implementation Of Audio Watermarking Technique For Copyright Protection using SWT Algorithm", *International Journal of Advanced Research in Computer and Communication Engineering*, Vol. 4, No. 3, pp. 42-48, 2016.
- [8] X. Wang, W. Qi and P. Niu, "A New Adaptive Digital Audio Watermarking based on Support Vector Regression", *IEEE/ACM Transactions on Audio Speech and Language Processing*, Vol. 15, No. 8, pp. 2270-2277, 2007.
- [9] S. Valizadeh and Z.J. Wang, "Correlation-and-Bit-Aware Spread Spectrum Embedding for Data Hiding", *IEEE Transactions on Information Forensics Security*, Vol. 6, No. 2, pp. 267-282, 2011.
- [10] S. Valizadeh and Z. J.Wang, "An Improved Multiplicative Spread Spectrum Embedding Scheme for Data Hiding",

- IEEE Transactions on Information Forensics Security*, Vol. 7, No. 4, pp. 1127-1143, 2012.
- [11] X. Wang, W. Qi, and P. Niu, "A New Adaptive Digital Audio Watermarking based on Support Vector Regression", *IEEE/ACM Transactions on Audio Speech and Language Processing*, Vol. 15, No. 8, pp. 2270-2277, 2007.
- [12] D. Lakshmi, R. Ganesh, R. Marni, R. Prakash and P. Arulmozhiarman, "SVM based Effective Watermarking Scheme for Embedding Binary Logo and Audio Signals in Images", *Proceedings of IEEE Region 10 Conference*, pp. 1-5, 2008.
- [13] M. Arnold, "Audio Watermarking: Features, Applications and Algorithm", *Proceedings of IEEE International Conference on Multimedia Expo*, pp. 1013-1016, 2000.