Imperceptibility and Robustness of the audio Haar waves using Discrete Wavelet Transformation and Singular Value Decomposition

Dr. Neelu Jhala, Assistant Professor, Department of Computer Science and IT, Faculty of Science, Bhupal Nobles' University, Udaipur

Abstract

The digital multimedia technologies have two main problems the digital multimedia security and intellectual property right. Digital watermarking is proposed to solve these problems. In digital applications, digital watermarking is a process of hiding some information in digital media and extracts it whenever needed to indicate the data owner.

In a watermarking technique, watermark is the information related to the original data or the owner is produced and is embedded in the original data and then the watermarked data distributed through computer networks. Watermarked is then extracted at receiving end to find the authentication of the digital data. In this paper we are presenting two cascading technique for watermarking are discrete wavelet transform and Singular Value Decomposition.

Keywords : multimedia, watermarking, discrete wavelet transform, Singular Value Decomposition

Introduction

Due to technology enhancement, it is easy to reproduce digital data in their exact original form which encourage copyright violation and data misappropriation and become the problems of theft and distribution of intellectual property. Creators and distributors of digital data are investigating the reliable solutions to the problems associated with copyright protection of multimedia data. Digital watermarking in multimedia data in one of them technique to protect the data from piracy[1,4,6].

Digital Watermarking

Digital watermarking is the process that embeds copyright information as 'watermark' into the multimedia object, so that the watermark can be extracted later to make an assertion about the ownership.

The watermark, can be used later to identify the owner of the work, to authenticate the content, and to trace illegal copies of the work. Digital watermarking is the process of embedding copyright and authentication information into digital media forms, such as text, image, audio or video.

Audio Watermarking

Audio watermarking is the application on audio signals, a rather new field. Compared with images and video, inserting imperceptible, robust and secure watermark(s) into digital audio files presents special challenge. Since audio signals have less samples per time interval embedding is difficult[7, 8,10]. As compared to image and video watermarking, audio watermarking methods are not easy due to Human Auditory System (HAS) is more sensitive than Human Visual System (HVS). Due to low sampling frequency, a small amount of noise in HAS system can be detected by ear. In this research we are using Haar Wavelet. A wavelet is a wave with amplitude begins at zero, increases, and then decreases back to zero. We are applying two technique Discrete wavelet transforms and singular value decomposition for embedding and extracting watermark the wavelet.

Discrete Wavelet Transforms

It is impossible to analyze a signal using all wavelet coefficients. The reconstruction of the signal from the corresponding wavelet coefficients is possible by discrete subset of the upper half-plane[5,6, 8]

Decomposition

The Discrete wavelet transforms of a signal x is calculated by passing it through a series of filters. First of all the samples are passed through a low pass filter by the given formula.

$$y[n]=(x\times g)[n]=\sum_{k=-\infty}^{\infty} x[k]g[n-k]$$

Simultaneously, the signal is also decomposed using a high-pass filter h. The detail coefficients are obtained from the high-pass filter and approximation coefficients from the low-pass.

Half the frequencies of the signal have now been removed and half the samples is discarded according to Nyquist's rule. The filter outputs are then subsampled. In the next two formulas, the notation is the opposite: g- denotes high pass and h- low pass as is Mallat's and the common notation:

$$y_{low}[n]=\sum_{k=-\infty}^{\infty} x[k]h[2n-k]$$

$$y_{high}[n]=\sum_{k=-\infty}^{\infty} x[k]h[2n-k]$$

This decomposition has halved the time resolution of the wave and only half of each filter output characterises the signal. However, each output has half the frequency band of the input so the frequency resolution has been doubled [6,9,11].

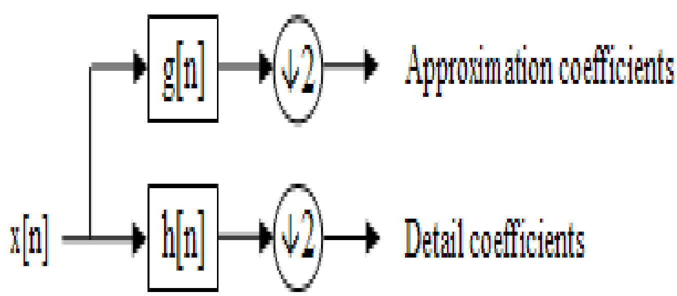


Figure 1: Block diagram of filter analysis

Singular Value Decomposition

The singular value decomposition (SVD) is a factorization of a real or complex matrix. Factorization form of the singular value decomposition of an $m\times n$ real or complex matrix M is a

$M=U\sum V^*$

Experimental Result

The audio file Audio Clip2.WAV as s input signal and Lena.PNG as watermark image. The Clip2.WAV file have a frequency of 44.1 KHz , the 16 bit per second sample. The total duration od aufio file is 25 sec. The total smaple in the audio file is 2005500.

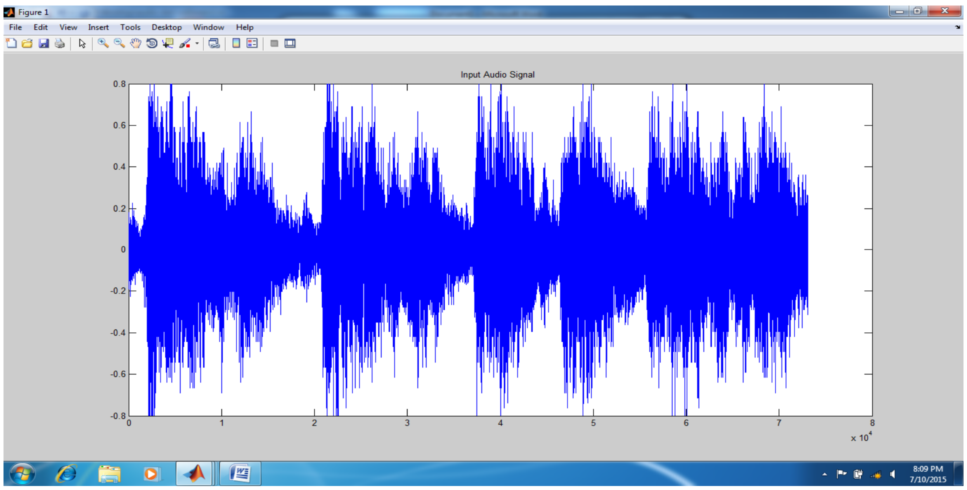


Figure 2: Original input signal (Audio Clip2.WAV)

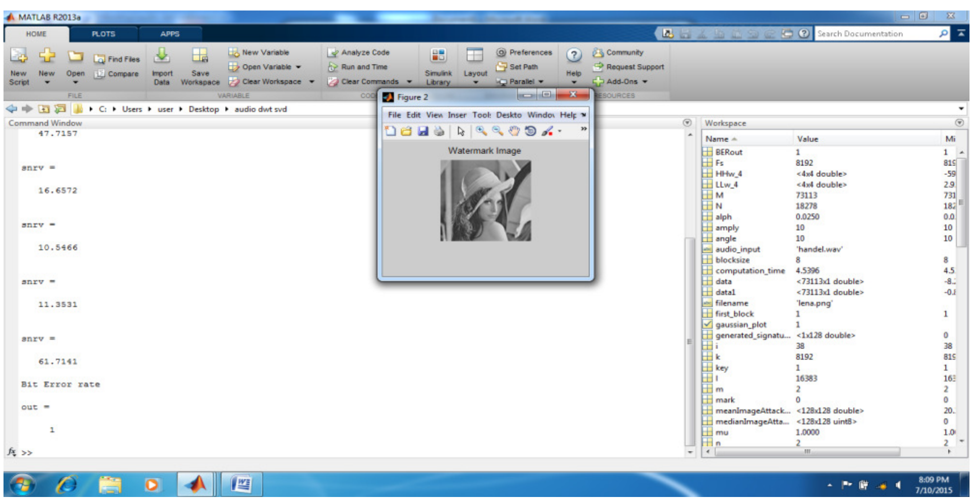


Figure 3: Watermark image (Lena.PNG)

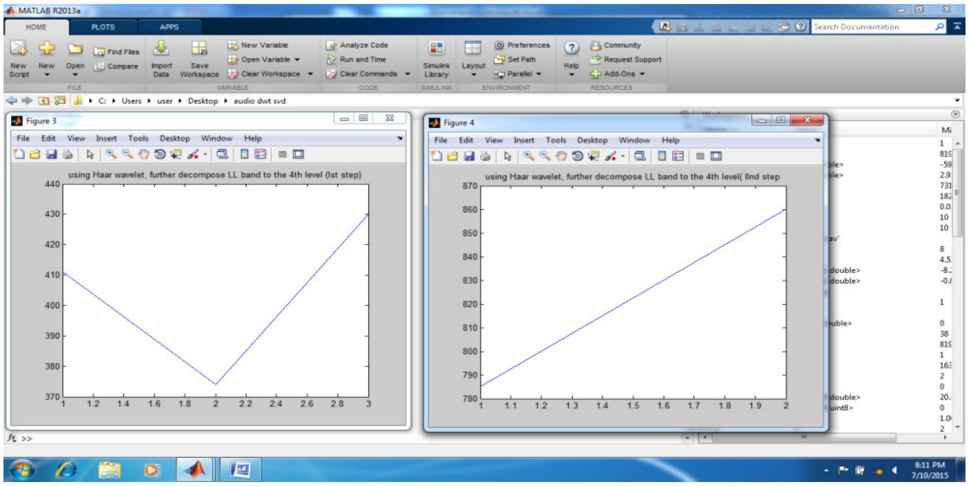


Figure 4: Higher Level of decomposition of LL subband of Audio clip2

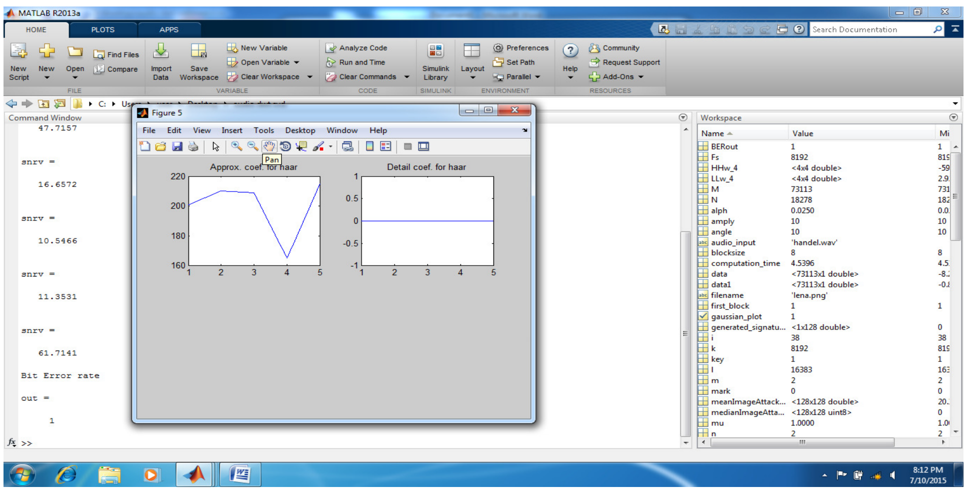


Figure 5: The approximate and detail of the 'HAAR' wavelet

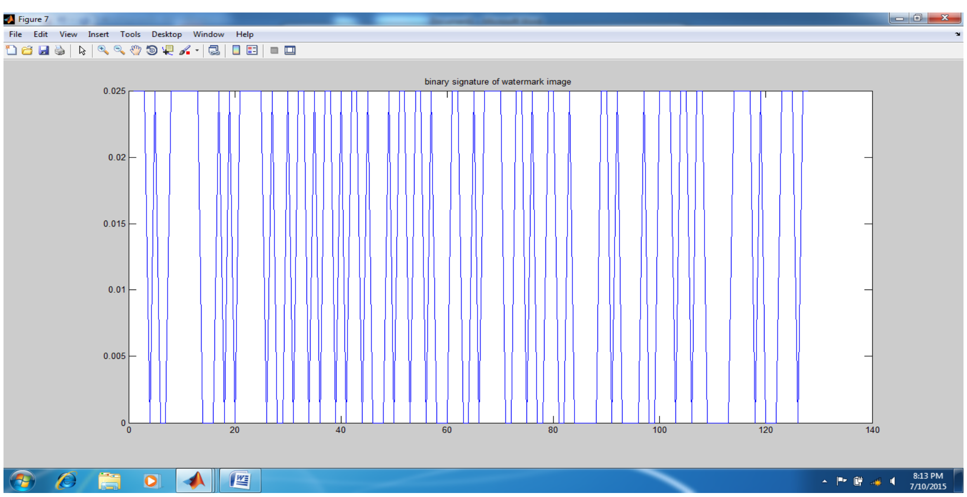
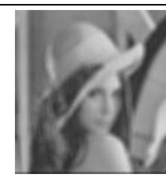
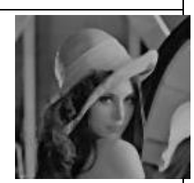


Figure 6: Binary signature of the watermark image Lena.PNG

Attack	Normalized Cross-correlation	BER	PSNR (dB)	SNR (dB)	Water-mark after attack
No Attack	1.000	0	66.59172123	47.56554589	
Cropping	0.9985	0.36	60.15492112	42.96784234	
Down-sampling	1.0000	0	64.89322478	46.35234571	
Up-sampling	1.0000	0	64.71010101	46.22154212	
Low-pass filtering	0.8963	18.56	69.32814152	49.52014575	
Re-quantization	0.9996	0.05	64.30704040	45.93366681	
MP3 Compression (32kpbs)	1.0000	0	65.63494523	46.88211234	
MP3 Compression (64kpbs)	1.0000	0	64.74398187	46.24577451	

MP3 Compression (128kpbs)	1.0000	0	64.44396452	46.03144231	
Denoising	0.9358	0.89	62.62242245	44.73038952	
Echo addition	0.9762	0.67	65.72230425	46.94455231	

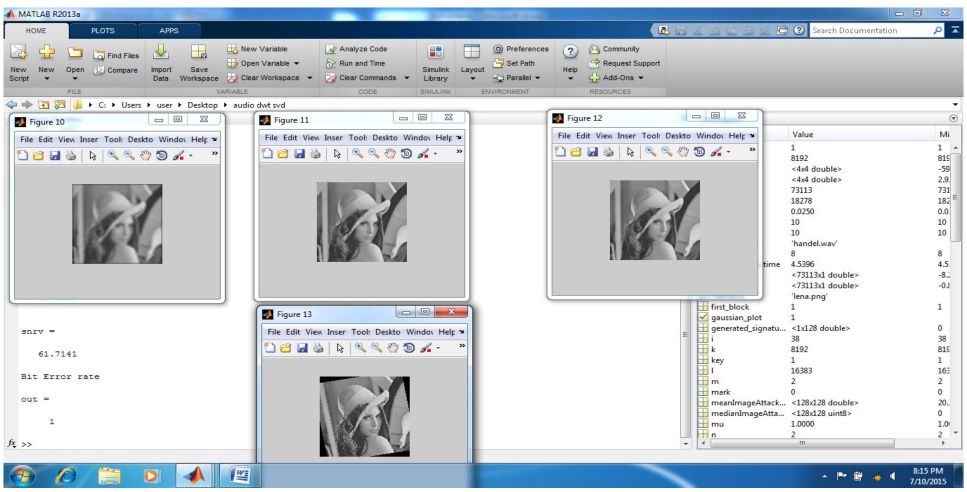


Figure 7: Extracted images (Lena.PNG) after the attacks on the watermarked audio signal

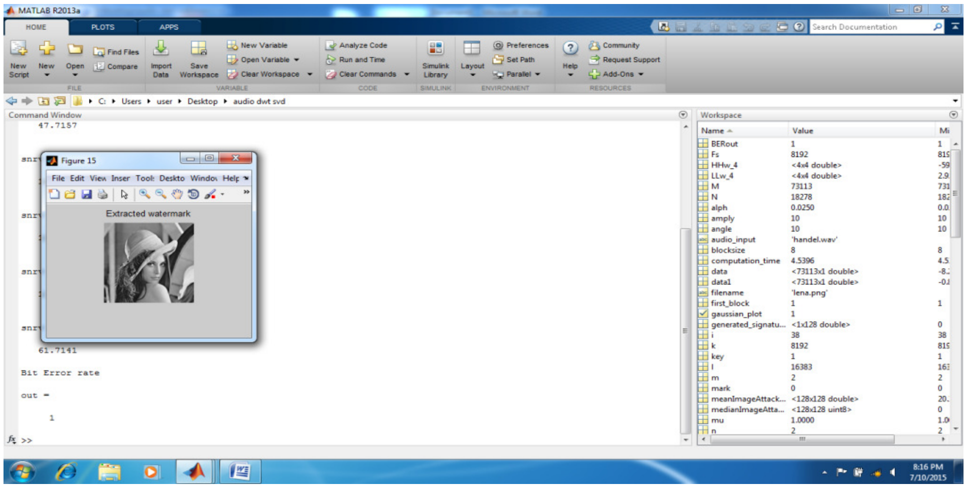


Figure 8: Extracted watermark image Lena.PNG

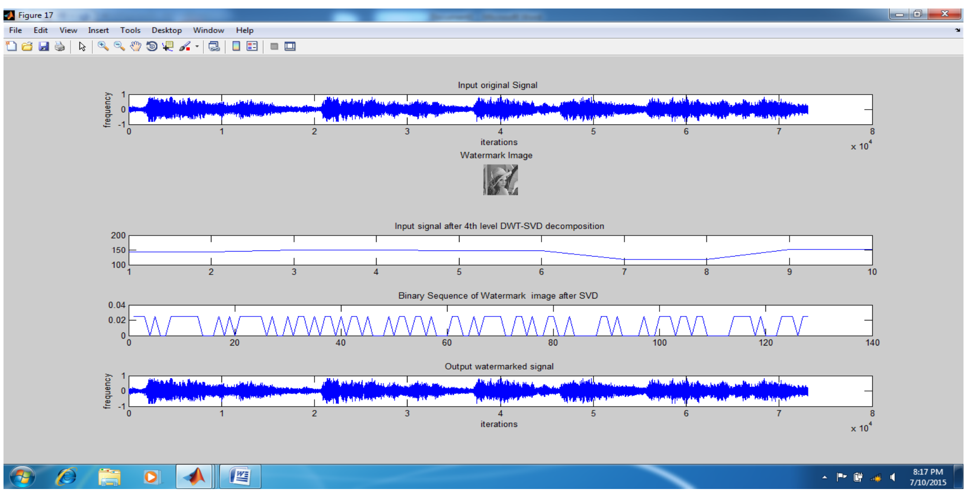


Figure 9: Input Signal (Audio clip2.WAV), watermark image (Lena.PNG), input signal after DWT-SVD decomposition, binary sequence of the image after SVD, watermarked signal

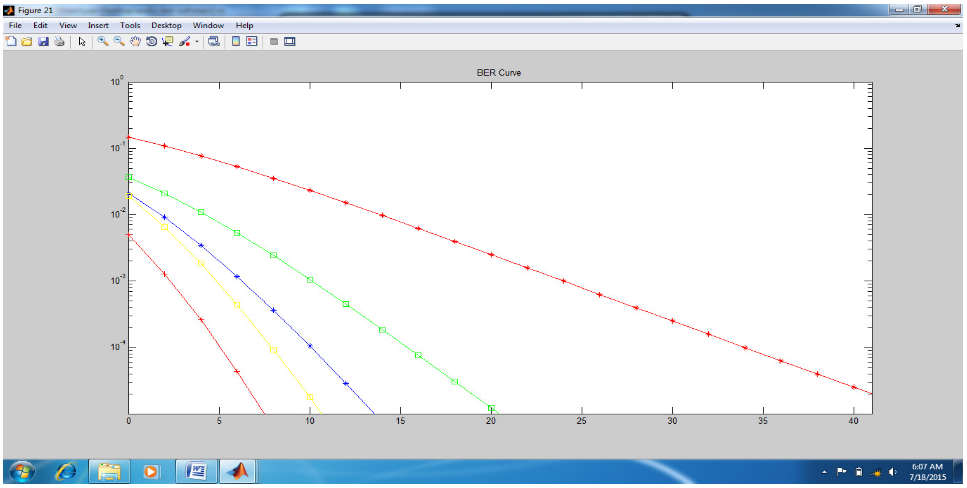


Figure 10: Bit error rate curve

Table1: Normalized cross-correlation, PSNR, SNR, extracted watermark after the no attacked and attacked with the audio signal Audio Clip2.WAV and watermark image Lena.PNG

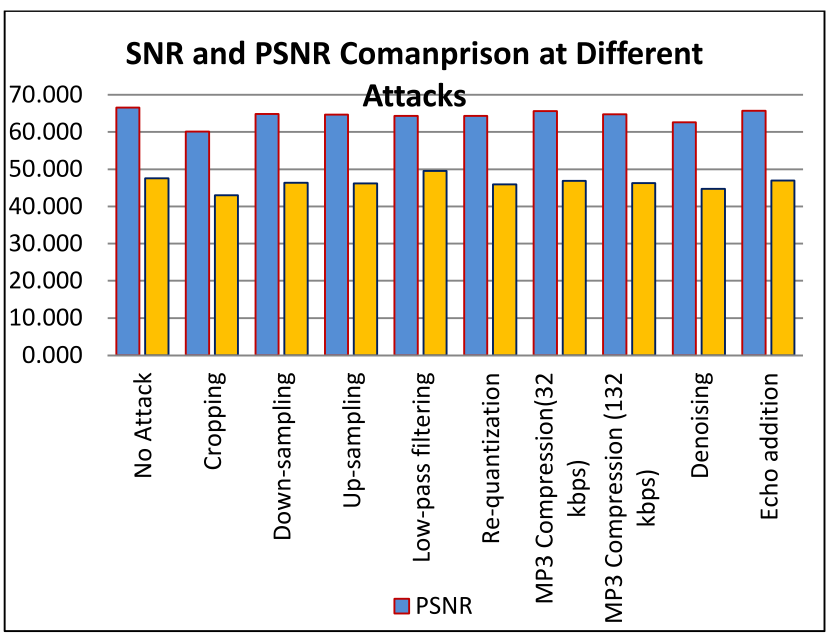


Figure 11: Graph for the SNR and PSNR comparison

at different attacks on Audio Clip2

Table2: Watermark intensity, execution time and SNR, PSNR for Audio clip2

Watermark Intensity	Execution Time	PSNR(dB)	SNR(dB)
Alpha = 10	4.2932	59.5655	35.7393
Alpha = 5	4.2882	63.5221	38.1132
Alpha = 1	3.9843	66.5917	47.5655
Alpha =.05	3.9562	68.2322	47.9393
Alpha = 0.025	4.1256	68.4675	48.0805

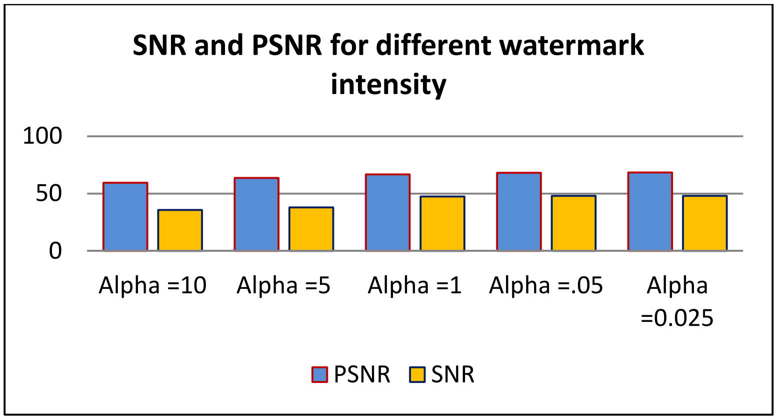


Figure 12: Comparison of the PSNR and SNR on Audio Clip2

Table 3: Normalizes cross relation at different attacks for Audio clip2

Attack	Alpha = 10	Alpha = 5	Alpha = 1	Alpha = 0.05	Alpha = 0 .025
No Attack	1.0000	1.0000	1.0000	1.0000	1.000
Cropping	0.9102	0.9265	0.9856	0.9912	0.9985
Down-sampling	1.0000	1.0000	1.0000	1.0000	1.0000
Up-sampling	1.0000	1.0000	1.0000	1.0000	1.0000

Low-pass filtering	0.8265	0.8466	0.8869	0.8942	0.8963
Re-quantization	0.9256	0.9453	0.9902	0.9985	0.9996
MP3 Compression (32kpbs)	1.0000	1.0000	1.0000	1.0000	1.0000
MP3 Compression (64kpbs)	1.0000	1.0000	1.0000	1.0000	1.0000
MP3 Compression (128kpbs)	1.0000	1.0000	1.0000	1.0000	1.0000
Denoising	0.8945	0.9158	0.9299	0.9305	0.9358
Echo addition	0.9265	0.9575	0.9722	0.9754	0.9762

Discussion and conclusion

The main focus of this study is robustness and imperceptibility of the audio watermark. The embedding process should not introduce any perceptible artifacts means watermark should not affect the quality of the original signal. The PSNR is used to investigate the amount of error which was introduced while embedding the watermark. The robustness of the techniques is also tested by using well known attacks. The correlation coefficient is used to determine the closeness of extracted watermark from the watermarked signal to the original watermark.

In our study, frequently-preferred transform domain technique DWT and decomposition method SVD is combined so that watermarked signal are much more robust against attacks.

The audio signal is decomposed into four suband using DWT which is known as first level of decomposition, further LL subband is decomposed into higher level.

Imperceptibility

The imperceptibility test is done by two methods. First is *subjective* (MOS grade), which is a listening test and was actually performed with 300 listeners to estimate the subjective grade of the watermarked signals. Each listener, listen the original signal and the watermarked signal and was asked to report whether any difference could be detected between the two signals. The 300 people listed to each pair for 10 times, and they gave a grade from 1 to 5 for this pair according to the table 1 in chapter 4 and listener were instructed before listing the audio signals. The average grade for original and watermarked from all listeners is the final grade..

Imperceptibility results (SNR and MOS grade) of above three experiments are summarized in Table 4.

Table 4: SNR and MOS for different input signal and watermark image

Input Signal	Watermark Image	SNR(dB)	MOS Grade
Audio Clip2.WAV	Lena.PNG	48.08050	5

We embedded the watermark with DWT-SVD domain into the audio signal. After embedding watermark, the SNR of all selected audio signals using the proposed method are above 20 dB. This satisfies the IFPI requirement (20 dB). This ensures that imperceptibility of the proposed system. The MOS value is 5 which is the highest value where the watermark is imperceptible. The highest SNR and PSNR is obtained at 0.025 ie at least watermark intensity.

Robustness

We are applying following attacks cropping, sampling, low pass filtering, requatization, MP3 compression, denoising, echo addition etc.

The value of NC of sampling(down and up) attack are 1 and BER is 0 and the value of the is NC of the other attacks cropping, low pass filtering, denoising, echo edition is near by 1, means values are varies from .7 to .9 and BER is very less. Since normalize cross relation is a statistical tool, which define the relation between the two different variable or value. The value of the NC exits between -1 to +1. If NC is 1 or -1 it prove that the variable or data are highly correlated. As value of NC is as much higher the data are highly correlated.

Table 5: NC and BER value for different input signal and watermark image

Input Signal	Watermark Image	BER	NC
Audio Clip2.WAV	Lena.PNG	0	1

Comparison

The performance of the proposed audio watermarking algorithm was compared with other typical audio watermarking algorithms and the DWT SVD domain based algorithm. This comparison is given in Table 6

Table 6: Comparison with other algorithm

Reference	Algorithm	SNR(dB)
Proposed algorithm	*DWT-SVD*	48.0805
Al-Haj *et al.*, 2011	DWT	28.57
B. Y. Lie *et al.* ,2011	SVD	32.53
De Li *et al.*,2013	DWT-DCT	32.57
Dha *et al.*, 2013	DWT-DCT	39.91
Yatin *et al.*, 2013	DCT	26.70
Khalid A. Darabkh *et al.*, 2014	DWT-SVD	37.07

The audio watermarking scheme presented in this study was based on decomposition of audio signal and applying the SVD on the higher frequency. Experiments were carried out and demonstrated that the proposed blind scheme was imperceptible, robust against cropping, sampling, denoising, echo addition MP3 compression and other attacks. Finally, the scheme had a high of about imperceptibility of 48.0805dB.

References:

1. Patil, Mishra, chitode (2013) An Effective Audio Watermarking using DWT-SVD, International Journal of Computer Applications 70(8):6-11, May 2013,70(8):6-11, DOI:10.5120/11980-7851

2. Al-Haj Ali, Ahmad Mohammad and Lama Bata (2011). DWT–Based Audio Water-marking. The International Arab Journal of Information Technology, July 2011 Vol. 8, No. 3.

3. Samah Alshathri, Ezz El-Din Hemdan(2023) An efficient audio watermarking scheme with scrambled medical images for secure medical internet of things systems, Multimed Tools Appl. 2023; 82(13): 20177–20195, 2023 Jan 9. doi: 10.1007/s11042-023-14357-6

4. anya Koohpayeh Araghi, David Megías (2023) Analysis and effectiveness of deeper levels of SVD on performance of hybrid DWT and SVD watermarking, Multimed Tools Appl. Published: 20 May 2023, Volume 83, pages 3895–3916, (2024)

5. Baluja, Yatin, Shray Mishra, Trilok Singh Saini and M.V. Patil (2013). Frequency Domain Based Data Hiding Technique for Audio Signal, *International Journal of Innovative Research in Science, Engineering and Technology*, May 2013, Vol. 2, Issue 5.

6. Erol Duymaz, Aydın Akan (2012). Novel digital audio watermarking approach/IU-JEEE , 2012, Vol. 12(2), 1517-1522.

7. Gupta Akshya Kumar, Raval Akshya Kumar (2012). A robust and secure water-marking scheme based on singular values replacement, *Sadhana* Vol. 37, Part 4, August 2012, pp. 425–440, Indian Academy of Sciences.

8. Schneier, Bruce (2015). The Data Encryption Standard (DES). Crypto-Gram. 15 June 2000, Retrieved 26 March 2015.

9. Xinkai Wang, Pengjun Wang, Peng Zhang, Shuzheng Xu, Huazhong Yang (2013). A norm-space, adaptive, and blind audio watermarking algorithm by discrete wavelet transform, Signal Processing, April 2013, Volume 93, Issue 4, pp. 913–922.

10. Yekta Said Can, Fatih Alagoz, and Melih Evren Burus (2014). A Novel Spread Spectrum Digital Audio Watermarking Technique. *Journal of Advances in Computer Networks*, March 2014, Vol. 2, No. 1.

11. K. Khaldi and I. Abdel - Ouahab Boudraa Member. "Audio watermarking via EMD, Transactions on audio, speech, and language processing", vol. 21, March.