Performance Evaluation of Digital Audio Watermarking based on Discrete Wavelet Transform for Ownership Protection

Mangal Patil and Janardan S. Chitode

Department of Electronics, Bharati Vidyapeeth University College of Engineering
Dhanakwadi, Pune, Maharashtra, India
mvpatil@bvucoep.edu.in

Abstract: Flexibility of Internet technology gives rise to concern about the protection of digital data. Digital audio watermarking is the art of hiding important data in Digital Audio. This research paper deals with a new methodology which helps to enhance the robustness of Digital Audio Watermarking technique based on Discrete Wavelet Transform (DWT) and singular value decomposition (SVD). In this approach, multiple watermarks (Image -64 X 64 pixels and text) are embedded into the approximate coefficients of Discrete Wavelet Transform of an audio signal. Haar wavelet is used for decomposition while SVD is applied to get the Eigen values of approximate coefficients of audio signal. Experimental work demonstrate that watermarked audio has same perceptual quality as that of original one. The performance of the proposed work is evaluated with the help of objective measures such as Signal to Noise ratio, Peak signal to Noise ratio, Bit error rate, Normalized Cross correlation, etc. Additionally, simulation results on audio signals prove that the capacity of the proposed scheme is considerably better in comparison with the existing watermarking methods. Furthermore, the proposed work is applied for different audio signals such as classical, instrumental, flute, pop & rock to check the robustness against various attacks.

Key Words: Discrete Wavelet Transform, Eigen Value, Imperceptibility, Robustness, Security.

NOMENCLATURE			
DWT	Discrete Wavelet Transform	BER	Bit Error rate
FFT	Fast Fourier Transform	PSNR	Peak Signal to Noise ratio
DCT	Discrete Cosine Transform	dB	Decibel
IFPI	International Federation of the Phonographic Industry	NC	Normalized correlation
SNR	Signal to Noise ratio	MSE	Mean square error

1. Introduction

Now a day the popularity of the Internet technology provides user to store, redistribute, and change the multimedia data easily. Hence the necessity for protection of digital contents for the ownership rights has arisen. Digital watermarking has been marked as an effective way of protection of intellectual property rights over cryptography [1]. The process of hiding sideband information (watermark- information relevant to the owner) into an audio signal is nothing but Digital audio watermarking [2]. However, an implementation of audio watermarking technique is more complex as compared to image watermarking because Human auditory system is more sensitive as compared with the Human visual system. Moreover, the amount of data to be inserted in audio signal is less because audio signal requires fewer samples to represent.

Every audio watermarking algorithm must satisfy certain objectives set by International Federation of the Phonographic Industry (IFPI) [3]. The important aspects of audio watermarking algorithm are imperceptibility, Robustness, Security, Capacity, etc.

Imperceptibility: Imperceptibility ensures that after watermarking process the audio quality should be retained. According to IFPI recommendations, Signal to Noise Ratio between host audio and watermarked audio should be more than 20 db.

Received: April 20th, 2016. Accepted: March 23rd, 2017
DOI: 10.15676/ijeei.2017.9.1.11

Robustness: This term refers to the ability of retaining the watermark in host audio despite different signal processing operations.
Security: Watermark ensures that only authorized person can modify or detect the watermark.
Capacity: Efficient watermarking technique should embed a large amount of data in audio signal.
Payload should be greater than 20 bits per second.
There is always a tradeoff between transparency and robustness as shown in Figure 1.

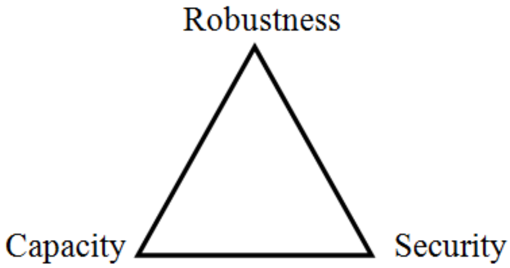


Figure 1. Requirements of watermarking

However, in this proposed work, the multiple watermark i.e. text and image is hidden into the eigen values of approximate coefficients of Discrete wavelet transform [4]. Furthermore, this approach achieves good robustness against attacks. Additionally, good capacity is also achieved with this proposed work.

The rest of the paper includes following sections. Section 2 presents background literature. Mathematical tools are described in section 3. The proposed approach along with the embedding and extraction process are presentated in section 4. Section 5 illustrates objective measures with detail simulation results. Finally, section 6 concludes the proposed work with future scope.

2. BACKGROUND LITERARTURE

To understand the copyright protection of audio signals few exiting techniques are reviewed. Digital Watermarking algorithms for audio signals are classified into two main types namely spatial (time) domain and frequency domain. The very simple method is Least Significant Bit replacement. Spatial domain embedding techniques are very simple and effective. These methods provide high capacity, low computational cost, and good perceptual quality. However, these methods are not suitable for authentication applications as it shows poor robustness than transform domain techniques. Frequency domain masks the watermark data and hides them into unimportant perceptually significant part of audio signal.

Recently various sub band coding based data hiding methods have been introduced for digital audio to improve the robustness. Fast Fourier Transform (FFT) [5], Discrete Cosine Transform (DCT) [6], DWT and singular value decomposition [7-10] are being used to hide the watermark in an audio signal. The implementation of a robust watermarking scheme for audio signals is a big challenge. To understand the existing schemes in frequency domain, several watermarking approaches are reviewed as follows.

Pranab Kr Dhar, et al, introduced sub band decomposition based audio watermarking algorithm along with synchronization code [11]. By changing singular values, an image watermark is hidden in an audio signal. It is based on Quantization index modulation technique. The results show that this method give embedding capacity about 45.9 bps. Moreover, high payload and good results for MP3 compression were obtained along with very low false error probabilities.

B. Lei discussed a method which involves the features of LWT [12]. This approach shows improved robustness, and imperceptibility. Initially, Audio signal is segmented using LWT to obtain approximate coefficients. The watermark is inserted. in the low frequency (approximate) band. SVD is applied separately to every block. This scheme shows good perceptibility and improved robustness against various attacks.

Wavelet based entropy (WBE) adaptive audio watermarking method developed by Pranab Kr. Dhar converts low frequency sub band into the wavelet based entropy form [13]. The mean of each audio is used to embed the watermark. The performance index i.e. SNR obtained is in the range of 13.5dB to 34.4 dB while embedding capacity is about 1000 bits per 11 seconds.

Yang Hong, et al, implemented wavelet based approach along with higher order statistics [14]. Wavelet de-nosing is applied using segmented audio samples. Each audio is decomposed into approximate & detail sub bands and then synchronization code is inserted into the average values of audio samples.

C.M. Juli & C. Janardhan introduced approach based on sub band coding for Indian classical songs [15]. Using Daubechies wavelet watermark is embedded in A8 sub-band wavelet coefficients. This helps to improve results in terms of PSNR, NC & BER against different attacks.

Krishna Rao et al, discussed a technique based on Discrete Wavelet Resonance DWR-SVD [16]. Secret sharing is important feature of this method which results in strong robustness for various cryptographic and compression attacks.

A blind watermarking approach described by Bassia in the time domain modifies the amplitude values [17]. High amplitude watermark is used to improve robustness with little degradation in perceptibility. Moreover, this method shows good robustness to MPEG compression, rescaling, filtering and re-quantization attacks.

Dongmei Wuet. et al, introduced a Haar wavelet based new technique. BCH encoded watermark sequence is hidden in audio signal [18]. This technique provides good robustness and reduces the intensity of watermarking. The watermarked audio has high perceptual quality as it gives SNR value of 42.87dB.

A robust audio watermarking algorithm through use of wavelet transform is proposed in [19]. The sub band decomposition of host signal is obtained to get the particular parts to hide the secrete data, imperceptibly. The high frequency region of audio is selected for watermarking. This method gives SNR about 28.55dB for instrumental audio samples and 25.03 dB for pop audio samples. It fulfills the objectives set by the IFPI for efficient audio watermarking techniques.

In last decade, many audio watermarking approaches have been implemented using a time domain or frequency domain methods. In our previous work a watermarking algorithm was developed using wavelet transform. This method achieved good imperceptibility. Recently, SVD based Image watermarking schemes have been implemented. However, SVD based watermarking algorithms are being used for audio as well.

The literature on transform based audio watermarking is quite good. The contribution of the proposed approach is in terms of improvement in robustness and increase in capacity by embedding multiple watermarks. In this approach, improved robustness is achieved by embedding the multiple watermarks in an audio using wavelet transform along with singular value decomposition technique. Optimizing singular values improve the robustness of this work. Moreover, Imperceptibility & robustness of the proposed work are evaluated using variety of audio samples.

3. Mathematical Tools

To implement proposed work, mathematical tools such as Discrete wavelet transform and singular value Decomposition are used.

A. Discrete Wavelet Transform

DWT divides a signal into approximation coefficients (low frequency) and detail coefficients (high frequency) exploiting multi-resolution for the non-stationary signals analysis [4]. The wavelet transform is generally expressed using a scaling and shifting parameter.

DWT needs a lower computational complexity compared with DCT and DFT. The watermark robustness can be increased to some extent by hiding Watermark signal in the higher level sub bands. The different types of wavelets are being used by researchers to implement watermarking schemes. In this proposed work, Haar function for scaling and wavelet for conversion of the host audio into the DWT domain is used. To enhance the robustness and perceptivity, our proposed work embeds the watermark in Eigen values of approximate coefficients of DWT.

B. Singular Value Decomposition

SVD is used for diagonalzing matrices. Any matrix A can be factored as $U*S*V^T$. The most important feature of SVD is quality does not get affected by changing singular values. Moreover, translation and scaling property of SVD makes it suitable for robust watermarking systems.

By applying SVD, matrix 'A' can be represented as

$$A= U*S*V^T \tag{1}$$

Where,
A– Matrix representation of cover audio signal
U – m x m (unitary matrix)
S – m x n (diagonal matrix) with positive elements
V– n x n unitary matrix
Columns of U represents left horizontal details and columns of V (right singular vectors) which represent vertical details of an audio signal respectively.

4. Proposed Methodology

This section presents the methodology for watermark embedding & extraction. The proposed technique is implemented for hiding image and text watermark into an audio. It consists of two major parts, namely watermark embedding & watermark extraction process.

Embedding process:

The schematic of the proposed watermark embedding process is shown in Figure 2. Here image & text are used as watermarks. After decomposing the cover audio into two sub bands (Approx. and detail), we apply the SVD to approximate sub band. The diagonal singular value coefficients are modified with the watermark(Image) using a scaling factor. Finally, Approximate sub-band coefficients are reconstructed with modified singular values. The Inverse DWT is applied to obtain watermarked audio.

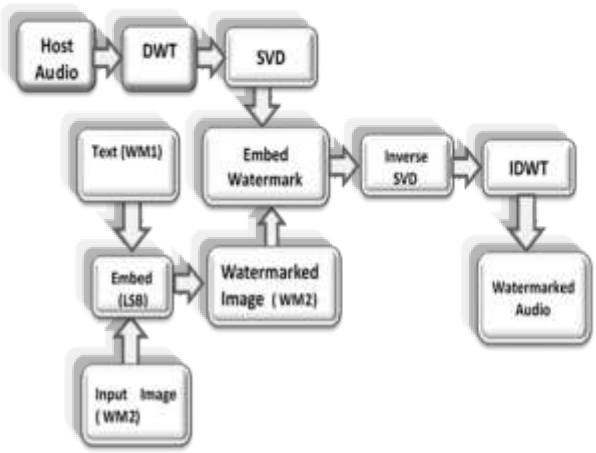


Figure 2. Embedding Process

The watermark embedding steps are described as follows:
S1: Decompose host audio using Haar DWT into detail & approximate sub-bands.
S2: Apply SVD to the approximate sub band (A) of an audio signal.
S3: Read text file to be embedded in an image watermark and encrypt it.
S4. Convert image watermark which is in binary form into a one-dimensional vector.
S5: Hide text (vidyapeeth) i.e. watermarks 1 into an Image (Watermark 2) using LSB technique.
S6: Similarly, embed bits of Image watermark (WM2) into the transformed version of original audio signal using the following equation,

$$S_{emb} = S + \alpha * S_{Wm} \tag{2}$$

Where, S = singular matrix of cover audio
S_{Wm} = singular matrix of watermark2 (Image)
S_{emb} = singular matrix of watermarked signal.
S7: Retrieve stego-audio by taking inverse DWT.

Extraction Process

The scheme of extraction of watermark is illustrated in Figure 3. The detail process of the extraction of watermark is described using following steps.
S1: Apply DWT and SVD to the watermarked signal and the Singular matrix values are computed for all frames.
S2: Obtain the singular matrix of image from transformed watermarked audio signal.
S3: Perform inverse SVD using two unchanged matrices (U and V).
S4: Apply inverse DWT to get all watermark audio frames.
S5: Apply LSB approach to obtain the ASCII values and convert into characters i.e. hidden text message (Vidyapeeth) from the retrieved image.

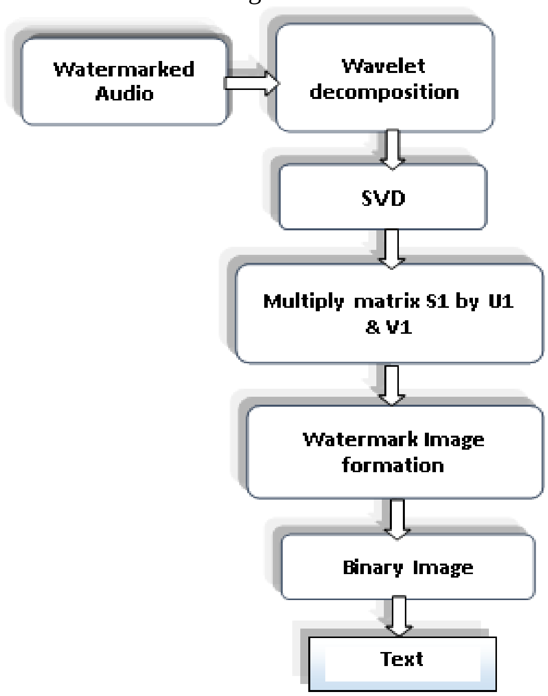


Figure 3. Extraction Process

5. Experimental Results and Analysis

This section demonstrates simulation results using MATLAB. The proposed work is evaluated using subjective as well as objective evaluation criteria [20].

A. Subjective Evaluation

To measure the performance, five different audio signals namely Classical, Pop, Rock, Flute & Instrumental in .wav format were used. Each signal with of 5-10 second duration was sampled at a frequency 44.1 KHz. The listening test is used as a subjective parameter to obtain imperceptibility [21]. The term Imperceptibility is referred as the perceptual quality of the watermark in the host signal. The perceptual quality of watermarked audio is evaluated practically by playing original and watermarked signals randomly to 10 listeners. The listeners were asked to rate each audio using standard mean opinion score (MOS). The MOS values for various samples were measured with the help of average of the rating provided by the listeners.

Table 1 shows the criteria for MOS grades used to evaluate perceptual quality of an audio and Table 2 give the results for MOS for the proposed algorithm. Moreover, MOS results prove that watermarked audio are sounded like the host audio. Thus, the inaudibility of watermark is retained. However, the subjective evaluation is time consuming.

Table 1. MOS criteria

Grade	Description	Quality
5	Imperceptible	Excellent
4	Perceptible	Good
3	Slightly Annoying	Fair
2	Annoying	Poor
1	Very Annoying	Bad

Table 2. MOS for the proposed algorithm

Audio	Average MOS
Classical	4.7
Pop	4.8
Flute	4.7
Rock	4.8
Instrumental	4.8

B. Objective Evaluation

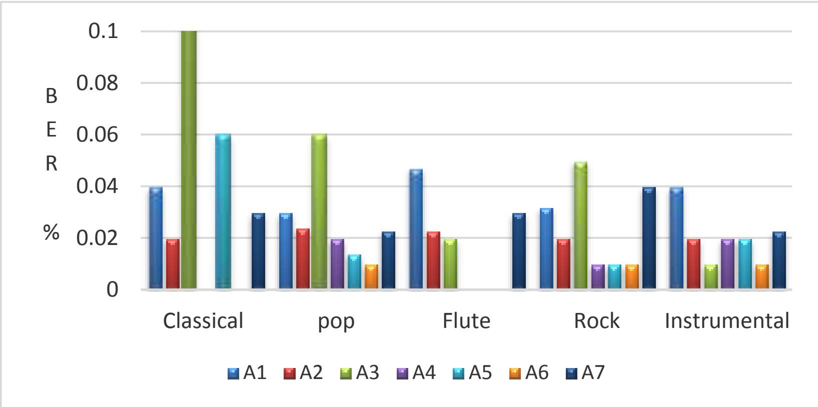
The watermark imperceptibility, robustness and capacity are computed based on an objective measure such as Signal to Noise ratio, the Peak Signal-to-Noise Ratio (PSNR), Normalized Bit error rate and number of bits per second [21].

Different tests have been performed to evaluate the robustness. A volume scaling test was conducted on the watermarked audio with scaling factors 0.8, and 1.1. In resampling attack, the sampling rate of the watermarked signal is changed to 22.05 KHz from original 44.1 kHz, and again resampled at 44.1 kHz. Form the results it is observed that the proposed scheme shows good robustness against resampling attack. In re-quantization, the watermarked audio is re-quantized from 16 bits to 8 bits and then it is quantized from 8 bits to 16 bits. The proposed method can retrieve watermark with minimum BER. In echo addition, an attack is applied on watermarked audio signal with a delay of 0.5 msec.

The performance evaluation of robustness supports the following observations:

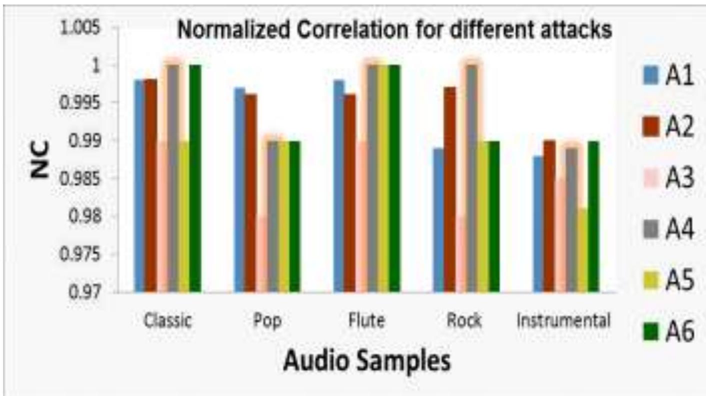
1. Figure 4 demonstrates the bit error rate results obtained for different audio signals. From these observations, it is seen that our method gives the average bit error rate much smaller in comparison with existing schemes. It varies with type of audio samples

2. Graph representing normalized correlation has been depicted in Figure 5 for different audio samples against common signal processing attacks such as Noise addition(A1), Volume Scaling(A2), Resampling(A3), Re-quantization(A4), Echo addition(A5), and Time Scaling(A6). The average NC against several attacks is also improved. The performance is evaluated for Classical (S1), POP(S2), Flute (S3), Rock(S4) and Instrumental(S5) audio signals.
3. The proposed approach has a better performance in terms of payload as it embeds two watermarks (Image 64 x 64 and text). The use of the wavelet and SVD helps to improve the robustness.
4. The perceptual quality evaluation of extracted watermark is done using Peak signal to noise ratio (PSNR) and the results are shown in Figure 6. From the results, it is observed that PSNR is well above 40 dB which satisfies the requirements set for the effective watermarking. Moreover, the normalized correlation between the original watermark and the extracted watermark is measured and found nearly equal to one. Time scaling attack introduces time displacement between hiding and extraction process. As per robustness test requirement of Secure Digital Music Initiative (SDMI), audio watermarking algorithm should resist ±4% Time scaling attack. The proposed method can resist only +1% time scaling attack.



A1-Noise addition A2-Volume scaling A3-Resampling A4-Requantization
A5-Echo addition A6-Time scaling A7-Low pass filtering

Figure 4. Bit Error Rate (%) v/s audio samples



A1-Noise addition A2-Volume scaling A3-Resampling
A4-Requantization A5-Echo addition A6-Time scaling

Figure 5. Results of Normalized correlation

5. The watermark extracted after signal processing operations are depicted in Figure 7. The simulation results show that the proposed approach can retrieve the watermark faithfully. However, these results vary with audio samples. The Figure 8 demonstrates the results for signal to noise ratio against different attacks (A1 to A7). The results show that our approach can achieve better SNR performance (in the range of 30 dB to 70 dB) and satisfies the IFPI criteria for effective audio watermarking.

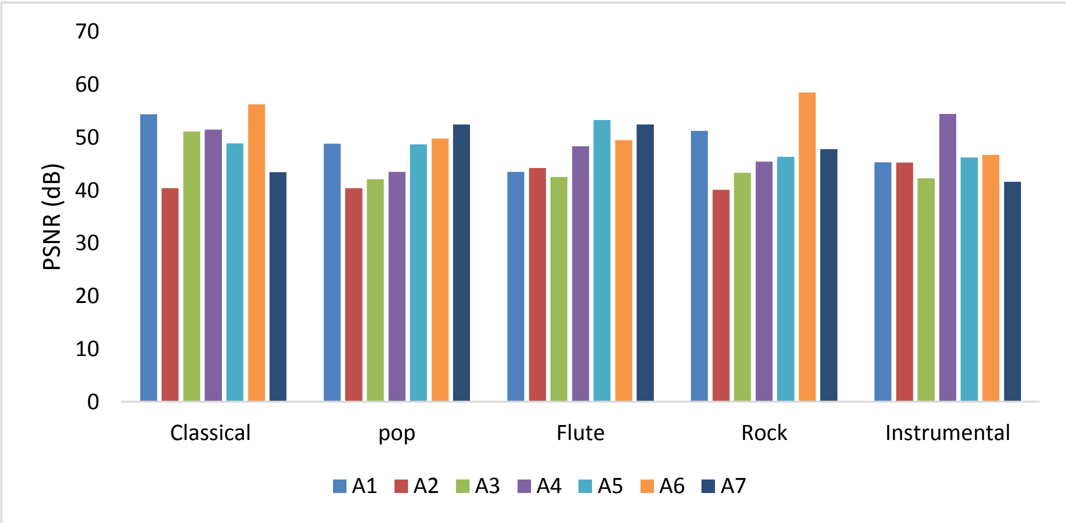


Figure 6. Performance in term of PSNR

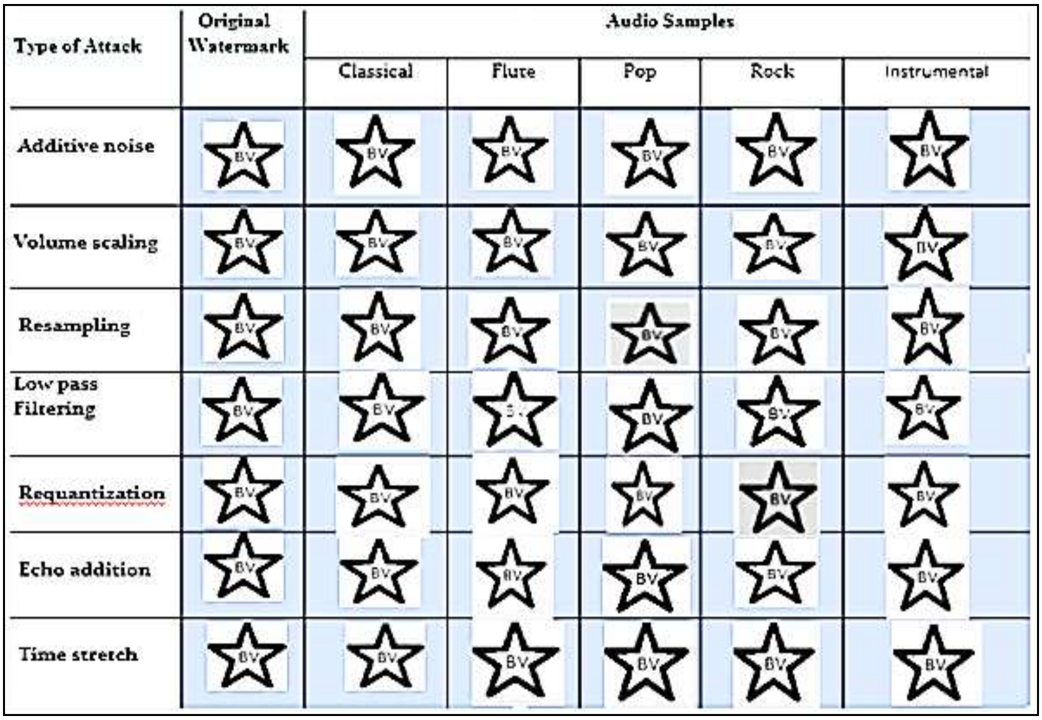
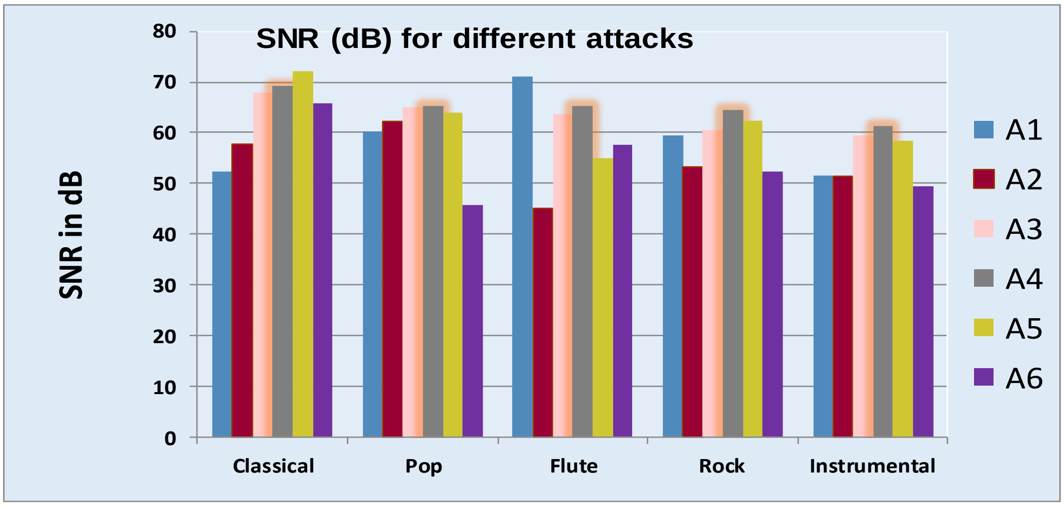


Figure 7. Sample results for Extracted watermark after attacks

6. The proposed algorithm is compared with the existing methods using objective measures such as SNR & BER as shown in Figure 9 and Figure 10 respectively.



A1-Noise addition A2-Volume scaling A3-Resampling
A4-Requantization A5-Echo addition A6-Time scaling

Figure 8. Performance in term of SNR

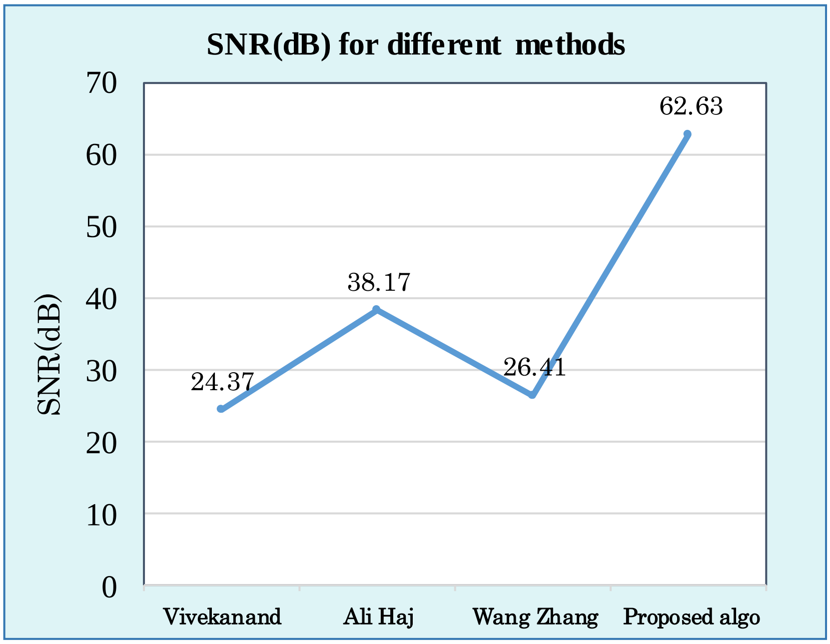


Figure 9. Performance of proposed method in terms of SNR

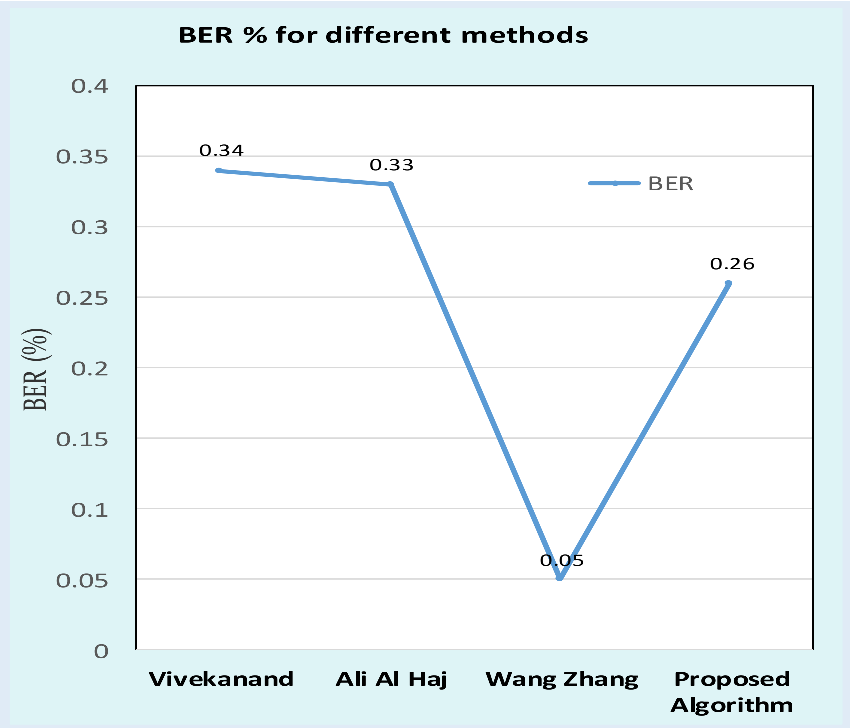


Figure 10. Performance of proposed method in terms of BER

6. Conclusion

This research work demonstrated a data hiding technique for digital audio based on DWT and SVD aimed at the implementation of a methodology which supports protection of the digital audio. In this approach, two watermarks namely image watermark of size 64 X 64 pixels and text watermark (vidyapeeth) are embedded. SVD is applied to get the Eigen values of approximate coefficients of audio. The singular values are modified with the help of watermark scaling factor. An analysis has been carried out to evaluate the robustness of the proposed algorithm using subjective & objective measures. The simulation results demonstrate improved PSNR values before and after attacks. The combination of DWT and SVD gives good robustness under different signal processing attacks. This technique gives low bit error rate and high payload for an audio by embedding an image (wm) of size 64 × 64 pixels & text watermark. This work is novel from the scheme proposed by Ali Haj in terms of objective parameters such as SNR, PSNR, NC, BER & Payload. As per robustness test requirement of Secure Digital Music Initiative (SDMI), audio watermarking algorithm should resist ±4% Time scaling attack. The proposed method can resist only +1% time scaling attack. This work can be extended to improve the robustness against time scaling attack (4%) by inserting synchronization code. However, the quality of audio is quite good as Human Auditory System is less sensitive to this attack.

7. References

[1]. Christine I. Podilchuk, Edward J. Delp “Digital Watermarking: Algorithms & Applications”, *IEEE Signal Processing Magazine*, Vol.18, No.4, pp.33-46, 2001

[2]. C Xu, J Wu, Q Sun, K Xin, “Applications of watermarking technology in audio signals”, *Journal of Audio Engg. Society*, Vol. 47, No.10, pp. 805–812, 1999.

[3]. IFPI (International Federation of the Phonographic Industry), 2009. http://www.ifpi.org.

[4]. Martin Vetterli, Jelena Kovacevic, “Wavelets and Sub-Band Coding”, NJ, Prentice Hall, Englewood Cliffs, 1995.

[5]. Md. Iqbal Hasan Sarker, Mohammad Ibrahim Khan, Kaushik Deb, Md. Faisal Faruque, “FFT-Based Audio Watermarking Method with a Gray Image for Copyright Protection”, *International Journal of Advanced Science and Technology*, Vol. 47, pp. 66-76, 2012.
[6]. X. Wang, H. Zhao, “A novel synchronization invariant audio watermarking scheme based on DWT and DCT”, *IEEE Transaction on Signal Processing*, Vol.54, No.12, pp. 4835–4840, 2006.
[7]. Ali Al-Haj, “A dual transform audio watermarking algorithm”, *Multimedia Tools and Applications*, Vol. 73, No.3, pp. 1897-1912, 2014.
[8]. Baiying Lei, Ing Yann Soon, Feng Zhou, Zhen Li, Haijun Lei, “A robust audio watermarking scheme based on lifting wavelet transform and singular value decomposition”, *Signal processing*, Vol. 92, No.9, pp. 1985-2001, 2012.
[9]. Vivekananda Bhat, Abhijit Das, Indranil Sengupta, “An adaptive audio watermarking based on the singular value decomposition in wavelet domain”, *Digital Signal Processing*, Vol.3, pp. 1547-1558, 2010.
[10]. R. Liu, T. Tan, “An SVD-based watermarking scheme for protecting rightful ownership”, *IEEE Transactions on Multimedia*, Vol. 4, No.1, pp. 121–128, 2002.
[11]. Pranab Kr Dhar, Saif Ahmad, “A new DCT based watermarking method for copyright protection of digital audio”, *International Journal of Computer Science & Information Technology*, Vol.2, No.5, pp. 91-101, 2010.
[12]. Y. Lei, I. Y. Soon, Zhen Li, “Blind and robust audio watermarking scheme based on SVD–DCT”, *Signal Processing*, Vol. 91, No.8, pp. 1973–1984, 2011.
[13]. Pranab Kumar Dhār, Tetsuya Shimamura, “Audio watermarking in transform domain based on singular value decomposition and Cartesian-polar transformation”, *International Journal of Speech Technology*, Vol.17, No.2, pp. 133-144, 2014.
[14]. Yang Hong-ying, Wang Xiang-yang, MaTian Xiao, “A robust digital audio watermarking using higher-order statistics”, *International Journal of Electronics and Communications*, Vol.65, No.6, pp. 560–568, 2011.
[15]. M. Juli Janardhanan, C. Satish Kumar, “Performance Analysis of Discrete Wavelet Transform based Audio Watermarking on Indian Classical Songs”, *International Journal of Computer Application*, Vol.73, No. 6, pp. 975-988., 2013.
[16]. Krishna Rao Kakkirala, Srinivasa Rao Chalamala, “Digital audio watermarking using DWR-SVD and secret sharing”, *International Journal of signal processing system*, Vol. 1, No.1, pp. 59-62, 2013.
[17]. P. Bassia, Pitas, “Robust audio watermarking in the time domain”, *IEEE Transaction on Multimedia*, Vol. 3, No. 2, pp. 232-241, 2001.
[18]. Dongmei Wu, Shangguo Yang, Zuochen Wang, Wenhua Qin, “BCH Code-Based Robust Audio Watermarking Algorithm in the DWT Domain”, *International Conference on Intelligent Computation Technology and Automation*, Vol.3, Changsha, May 11-12, pp. 878-880, 2010.
[19]. W. Bender, D. Grul, N. Morimoto, A. Lu, “Techniques for data hiding”, *IBM System Journal*, Vol. 35, No.4, pp. 313-316, 1996.
[20]. A. Fabien, P. Petitcolas, “Watermarking Schemes Evaluation”, *IEEE Signal Processing*, Vol.17, No. 5, pp. 58-64, 2000.
[21]. Yiqing Lin, W. H. Abdulla, “Perceptual Evaluation of audio watermarking using objective quality measures”, *IEEE International Conference on Acoustics, Speech and Signal Processing*, Las Vegas, NV, pp. 1745 – 1748, Mar-31-April 4 2008.

Mangal Patil, has obtained the B.E degree in Electronics Engineering from Shivaji University, Kolhapur, India in 1999. She received M.E. degree from the Bharati Vidyapeeth Deemed University College of Engineering at Bharati Vidyapeeth University, India in 2006. Currently pursuing Ph.D. (Electronics Engineering) from Bharati Vidyapeeth University College of Engineering, Pune. She has 15 years of teaching experience.

Janardan S. Chitode is Currently working as professor in the Bharati Vidyapeeth University College of Engineering, Pune (India). received B.E. degree in Industrial Electronics Engineering from Bharati Vidyapeeth University, Pune, Maharashtra in 1991. He received M.E. degree from College of Engineering (COEP), Pune in 1995 and Ph.D. degree in Electronics from Bharati Vidyapeeth Deemed University, Pune in 2009. His research interest includes Signal processing, Speech Synthesis, etc. He has published more than 55 research papers in various National and International Journals and conferences