IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 3, Issue 11, November 2016

Comprehensive Study & Implementation of Audio Watermarking using DWT-SVD-Firefly Algorithm

Er. Harmanpreet Singh[1], Er. Shagun Sharma[1]

Electronics & Communication Engineering Dept, Modern Institute of Engineering & Technology, Mohri, Kurukshetra[1]

Abstract: This paper investigates the development of digital audio watermarking in addressing issues such as copyright protection. Over the past two decades, many digital watermarking algorithms have been developed, each with its own advantages and disadvantages. The main aim of this thesis was to develop a new watermarking algorithm within an existing discrete wavelet Transform (DWT) and singular value decomposition (SVD) framework. This resulted in the development of a combination of DWT-SVD-Firefly (Firefly algorithm) watermarking algorithm. In this new implementation, the embedding depth was generated dynamically thereby rendering it more difficult for an attacker to remove, and watermark information was embedded by manipulation of the spectral components in the spatial domain thereby reducing any audible distortion. Further improvements were attained when the embedding criteria was based on bin location comparison instead of magnitude, thereby rendering it more robust against those attacks that interfere with the spectral magnitudes. The further aim of this thesis is to analyze the algorithm from a different perspective. Improvements were investigated using DWT-SVD and DWT-SVD-firefly algorithms. A comparison of these two algorithms for different watermark images on two types of audio input signal is presented here. The whole thesis work is divided into five main chapters with 4^{th} 5^{th} chapter describing proposed algorithm and results with discussions.

Keywords: Discrete wavelet transform (DWT), singular value decomposition (SVD), wavelet transform, survey papers.

I. INTRODUCTION

Technological advances in computing, communications, consumer electronics and their convergence have resulted in phenomenal increases in the amount of digital content that is being generated, stored, distributed, and consumed. The term “content” broadly refers to any digital information, such as digital audio, video, graphics, animation, images, text, or any combinations of these types. This digital content can be easily accessed, perfectly copied, rapidly disseminated and massively shared without it losing quality, as opposed to the situation with earlier analogue media, such as audio cassettes and Video Home System (VHS) tapes. However, these advantages of digital media formats over analogue transform into disadvantages with respect to copyright management, because the possibility of unlimited copying without a loss of fidelity has led to a considerable financial loss for copyright holders.

In terms of a solution to the financial losses incurred from unauthorised copying, content owners predominantly turn to cryptography, which is one of the most commonly used methods of protecting digital content. In the cryptography process, the content is encrypted prior to the delivery to the consumer, and then a decryption key is provided only to those who have purchased legitimate copies of the content. However, cryptography does not offer a robust solution to content piracy. For example, a pirate could purchase the encrypted content legitimately and then use the decryption key to produce and distribute copies of the content illegally. In other words, once decrypted, the content has no further protection. Thus, there is a strong need for an alternative or complement to cryptography. In terms of the solution to the problems encountered with cryptography, watermarking has been proposed as it has potential to offer more robustness. It can protect the digital content during its normal usage because the copyright information is placed within the digital content in such a way that it cannot be removed. This unique feature of watermarking makes it one of the most promising techniques for digital content protection, which has been the motivating factor behind much of the research in the last two decades. This chapter is organized as follows:

First, a brief history of watermarking is given followed by an overall illustration of how watermarking algorithms work [1]. Then, typical applications of watermarking are introduced. The motivation and contributions of this thesis are explained subsequently. Finally, an overview of this thesis is provided.

II. AUDIO WATERMARKIMG USING DWT-SVD AND FIRE FLY ALGORITHM

- **Illustration of watermarking algorithm**

In this section, the definition and characteristics of a watermarking algorithm will be illustrated.

- **Watermarking definition**

Digital watermarking is the process by which a discrete data stream called a watermark is embedded within a

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 3, Issue 11, November 2016

digital content. It is a special form of steganography, which is concerned with developing methods of writing hidden messages in such a way that no one, apart from the intended recipient, knows of the existence of the message. Generally, the process of watermarking can be divided into two parts: embedding and detection, the embedding process is depicted in Figure 1.1. As seen from Figure 1.1, the process of watermark embedding is very straightforward, that is, the watermark is generated and then is embedded into the original content by some means. A variety of embedding algorithms have been developed which rely on the manipulation of some properties of the original content.

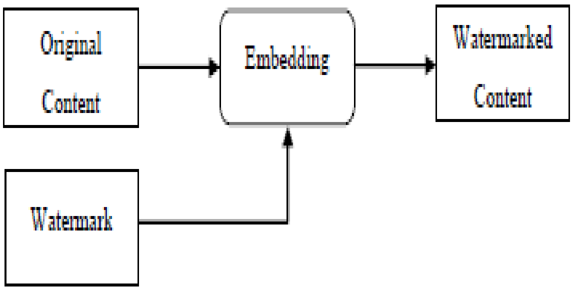


Figure 1.1 the flowchart of the embedding process

The detection process, as depicted in Figure 1.2, illustrates the embedded watermark information being recovered only by the use of the key.

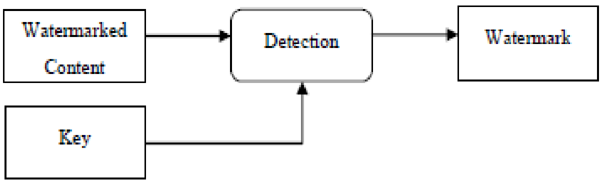


Figure 1.2 One type of detection processes

A more complex detection process, as depicted in Figure 1.3, illustrates the detection of the presence of watermark information, which requires not only the key, but also the original watermark and/or the original content. Compared with the detection process described in Figure 1.2, this needs extra storage capability for the original watermark information or original host information. However, the advantage of using the detection process described in Figure 1.3 is that the detection performance can be greatly improved when the original host or original watermark bit sequence is available.

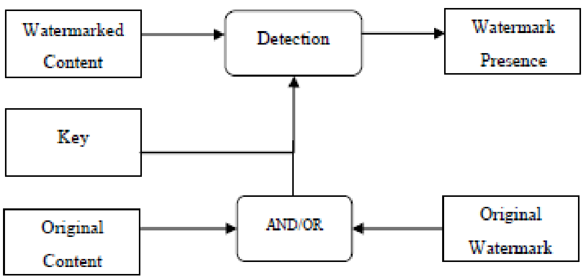


Figure 1.3 the other type of detection process

• **Watermarking characteristics**

A watermarking algorithm has the following characteristics [NC06, Cve07]:

• **Imperceptibility**: in general, the embedded watermark should not affect the human perception of the content. Namely, the watermark should be "invisible" in an image/video or "inaudible" in audio. However, in some special application scenarios, the watermark should be obtrusive to serve as a statement of ownership. Therefore, watermarking algorithms can be classified as „perceptible" where the embedded watermark can be perceived and„ imperceptible" where the embedded watermark cannot be perceived.

• **Robustness**: any manipulation of the watermarked content is defined as an attack. The embedded watermark should be robust against attacks. That is, the watermark recovery accuracy should not be decreased significantly after the watermarked content is attacked. Attacks on watermarks can be accidental or intentional. Accidental attacks are the result of standard signal processing that the signal might undergo, such as Analogue to Digital (A/D) conversion, Digital to Analogue (D/A) conversion and lossy compression. Intentional attacks refer to those that deliberately distort or remove the embedded watermarks. As far as audio is concerned, the main attacks, both intentional and accidental, can be divided into the following groups:

I. Filtering: such as highpass filtering, lowpass filtering and equalization. An equalizer only increases or decreases specified spectral regions.
II. Lossy compression: such as MP3 or Advance Audio Coding (AAC), which are used to reduce the amount of audio data.
III. Noise: such as noise adding or removal.
IV. Conversion: such as A/D, D/A or conversion of the sampling frequency (for example, from 32 kHz to 48 kHz).
V. Time stretch: increasing or decreasing the duration of an audio signal without changing its pitch.
VI. Pitch shift: change the pitch without changing the speed of the audio.

The task of designing a robust watermarking algorithm, which is able to withstand all or even a subset of possible attacks, appears to be quite difficult. Each algorithm currently proposed has its own weakness against certain types of attack. However, not every attack is possible with particular applications. Thus, identifying potential attacks that are associated with a specified application is essential. The most powerful attacks are those that can remove or distort the watermark information without severely degrading the content quality. With these attacks, the watermark information cannot be recovered but the audio can be used normally. Watermarking algorithms can be classified as „robust", „fragile" and „semi-fragile" according to their robustness against attacks.

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 3, Issue 11, November 2016

• **Capacity**: the watermarking algorithm should be capable of embedding a large amount of watermark information into the digital content.

• **Blindness**: in general, the watermarking algorithm should be blind, that is, the embedded watermark can be detected without requiring access to the original content or original watermark. However, some watermarking algorithms can only detect the presence of the original watermark. Therefore, watermarking algorithms can be classified as „blind" and „informed" in terms of how much information is required at the detection stage. The „blind" watermarking algorithm only requires a „key" to detect the watermark. While the „informed" watermarking algorithm needs the original watermark or the original content to detect the presence of the watermark information.

• **Computational efficiency**: the efficiency of the watermarking algorithm will determine if it can be applied in time-critical applications.

• **Security**: Kerchhoff"s Principle states that a cryptosystem should be secure even if everything about the system, except the key, is publicly known. This principle was reformulated by Claude Shannon as "the enemy knows the system", which is embraced by cryptographers worldwide. As far as watermarking systems are concerned, the algorithm might be published or made public. However, an unauthorized user, who may even know the exact watermarking algorithm, should not be able to detect the embedded watermark unless the secret keys are disclosed.

• **Adjustability**: The algorithm should be tuneable to various degrees of robustness, imperceptibility and capacity to facilitate diverse applications.

• **Watermarking trade-off**
The imperceptibility, robustness and capacity are the three most important characteristics of a watermarking algorithm. However, there is a trade-off between these three characteristics. This trade-off can be represented as shown in Figure 1.4.

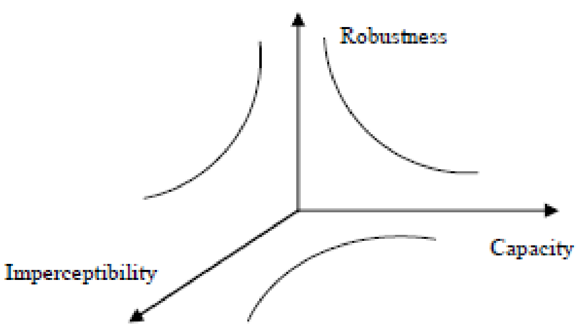


Figure 1.4 Trade-off presents in watermarking system

As seen from Figure 1.4, imperceptibility, robustness and capacity are conflicting characteristics of a watermarking system. For example, a specific application may determine what the capacity is needed. After it is determined, there exists a trade-off between imperceptibility and robustness. If one then wants to make the watermark more robust against attacks, a larger modification of the signal"s properties to embed the watermark will be necessary. However, this will worsen the imperceptibility [4]. Another scenario may be that with a predefined requirement for the imperceptibility, there will exist a trade-off between the capacity and robustness. For instance, the fewer the message bits that are embedded, the more redundant the watermark can be. Therefore, the watermarking will have a better error correction capability against attacks, that is, it is more robust.

III. PROPOSED SYSTEM

From literature study it has been observed that hybrid transformation algorithm performs well than individual DWT, DCT and SVD transformation techniques for watermarking. It is always a tradeoff between robustness and exact retrieval of watermark from the audio signal along with the data packets watermarked. DWT-SVD seems to be a promising than others. It is also resistive to attacks. The watermarking is done by the generalize formula [5].

$$Sw=S+a*W$$

Where Sw is the watermarked matrix, S is the original matrix to be watermarked and W is the watermark. 'a' is the watermark intensity which should be chosen to tune the trade-off between robustness and imperceptibility. So in our work this tuning will be done by optimization and a value of 'a' will be reached which will give tradeoff results. Following will be our key objectives for our proposed work:

- Firefly algorithm will be used for tuning purpose aling with DWT-SVD watermarking techniques.
- A binary image will be used as watermark and it will always remain size constraint.
- Results will be analyzed on basis of PSNR (peak signal to noise ratio), NCC (normalized cross correlation), MSE (mean square error).

Tool Used: MATLAB's signal processing toolbox. The audio watermarking is necessary to avoid the original sound form theft and tempering. To protect and watermark the audio signal the algorithm should be such that it audio signal shouldn't be changed in any aspect and it should be robust and message hidden in the audio signal should not be retrieved very easily by others. To fulfill these we proposed the combination of Discrete Wavelet transform (DWT) and Singular Value Decomposition (SVD) is used which is further modified by firefly algorithm (FA). The embedding formula used in our case is

$$S_w = S + \alpha . W$$

Where S_w is the watermarked audio signal, S is the original audio signal, α is the gain factor which is the

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 3, Issue 11, November 2016

measure of robustness and imperceptibility of message and W is the watermark message bit.

The gain factor is the deciding factor for imperceptibility and robustness of watermarked audio signal and it should be optimum so that retrieval of watermark message is easier as high gain factor can leads to distorted extraction of message but robustness increases with high gain value. So we have used FA optimization algorithm (discussed in previous chapter) to tune the gain factor value for trade off between PSNR (peak signal to noise ratio) and MSE (mean square error). The objective function chosen for this purpose considers the three evaluation parameters: PSNR, NCC (normalized cross correlation), MSE, which is discussed in next sections.

Till now many researchers have used optimization algorithms to tune the gain value but signal value for whole audio signal was used[2]. This increases the chance of perception for gain value by third user. In our work, we have divided the audio signal into chunks and gain value is calculated by FIREFLY as per the number of chunks. The number of chunks depends upon the watermark level used. We have developed the dynamic MATLAB script which can work for desired DWT levels[3].

Message chunks = (numel (watermark) / (dwtlevel^2-dwtlevel))

- **Algorithm for watermark embedding**

STEP1. Convert the binary image watermark into a one-dimensional vector b of length M×N. A watermark bit bi may take one of two values: 0 or 1.

$$b_i = \{0,1\} \qquad 1 \le i \le M \times N$$

STEP2. Sample the original audio signal at a sampling rate of 44,100 samples per second and partition the sampled file into N frames. The optimal frame length will be determined experimentally in such a way to increase data payload.

STEP3. Perform a four-level DWT transformation on each frame. This operation produces five multi-resolution sub-bands: D1, D2, D3, D4, and A4. The D sub-bands are called 'detail sub-bands' and the A4 sub-band is called 'approximation sub-band'. The five sub-bands are arranged in the vector.

STEP4. Arrange the four detail sub-bands D1, D2, D3, and D4 in a matrix D as shown in Figure 4.2. The matrix formation is done this way to distribute the watermark bits throughout the multi-resolution sub-bands D1, D2, D3, and D4. Forming the matrix with the Ds, rather than using A alone, is done to allow for matrix formation and subsequent application of the matrix-based SVD operator. The size of matrix D is 4 × (L/2), where L refers to the length of the frame.

STEP5. Decompose matrix D using the SVD operator. This operation produces the three orthonormal matrices Σ, U, and V^T as follows:

$$D = U * \Sigma * V^T$$

where the diagonal matrix Σ has the same size of the D matrix. The diagonal σii entries correspond to the singular values of the D matrix. However, for embedding purposes, only a 4 × 4 subset of matrix Σ, assigned the name S hereafter, is used as shown below. This is a trade-off between imperceptibility (inaudibility) and payload (embedding capacity). That is, using the whole Σ matrix for embedding will increase embedding capacity but will lead to severe distortion in imperceptibility (inaudibility) of the watermarked audio signal.

$$S = \begin{bmatrix} S_{11} & 0 & 0 \\ 0 & S_{22} & 0 \\ . & . & . \\ . & . & S_{nn} \end{bmatrix}$$

STEP6. Arrange 12 bits of the original watermark bit vector b into a scaled 4 × 4 watermark matrix W. The watermark bits must be located in the non-diagonal positions within the matrix, as shown below.

$$W = \begin{bmatrix} 0 & \text{bit1} & \text{bit2} & \text{bit3} \\ \text{bit4} & 0 & \text{bit5} & \text{bit6} \\ \text{bit7} & \text{bit8} & 0 & \text{bit9} \\ \text{bit10} & \text{bit11} & \text{bit12} & 0 \end{bmatrix}$$

STEP7. FA algorithm starts from this step. The searching space dimension of firefly is equal to the number of chunks as it will be final tuned gain value. The table 4.1 correlates the bio term in FA algorithm with our technical counterpart.

Table 4.1: Technical counterpart of bio inspired variables

	Variable in Bio Inspired Algorithm	Terms in our technical concept
1	Position of firefly	Gain factor values
2	Number of dimension of searching space	Number of gain factors to be tuned for embedding
3	Update in positions of firleflies	Change in the gain factor's value

The fitness function used for our proposed scheme in FA takes PSNR, MSE and NCC in consideration which is:

$$\text{fitness value} = \left(\frac{1}{\text{PSNR} + \text{NCC}}\right) + \text{MSE}$$

Since fitness value should be minimized so inverse of PSNR and NCC is considered here. For each collection of gain values (number of gain values is equal to number of audio signal chunks) the embedding algorithm is executed and fitness value is calculated by above equation. Once all iterations are finished, the gain factor values for minimum fitness function is picked as the final gain values, which are used further for embedding of watermark message.

STEP8. Embed watermark matrix W bits into matrix S according to the additive-embedding formula of equation 4.1.

STEP9. Decompose the new watermarked matrix Sw using the SVD operator. This operation produces three new orthonormal matrices as follows:

$$S_w = U_1 * S_1 * V_1^T$$

The matrices U_1 and V_1^T are stored for later use in the extraction process. This makes the proposed watermarking algorithm semi-blind, as the whole original audio frame is not required in the extraction process.

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 3, Issue 11, November 2016

STEP10. Apply the inverse SVD operation using the U and VT matrices, which were unchanged, and the S1 matrix, which has been modified according to Equation below. The Dw matrix given below is the watermarked D matrix.

$$D_w = U * \Sigma' * V^T$$

where matrix Σ′ is the original Σ matrix with the S sub-matrix replaced by the S1 sub-matrix.

STEP11. Apply the inverse DWT operation on the Dw matrix to obtain the watermarked audio frame.

STEP12. Repeat all previous steps on each frame. The overall watermarked audio signal is obtained by concatenating the watermarked frames obtained in the previous steps.

IV. RESULT AND DISCUSSIONS

As discussed in previous chapter this thesis work is to suggest a new method for audio watermarking and analysis is done in MATLAB. We have used MATLAB 2013a's signal processing toolbox to test our proposed algorithm and a comparison is done with the already existing DWT-SVD algorithm. Results of reference paper are not quoted here, as test conditions are different along with sample watermark images and input audio samples. MATLAB's signal processing toolbox provided may functions ready to use which reduces our hassle to write script for those and we were able to concentrate on our proposed work's implementation. The results has been tested for a recorded signal at 44100 Hz frequency as well as at live recording of audio signal at the same frequency. Different types of watermark messages with varying sizes are used for analysis purpose. As described in 4th chapter PSNR, NCC and MSE will be defining parameters in our work. The recorded input signal is shown in subplot 1 in figure 5.1.

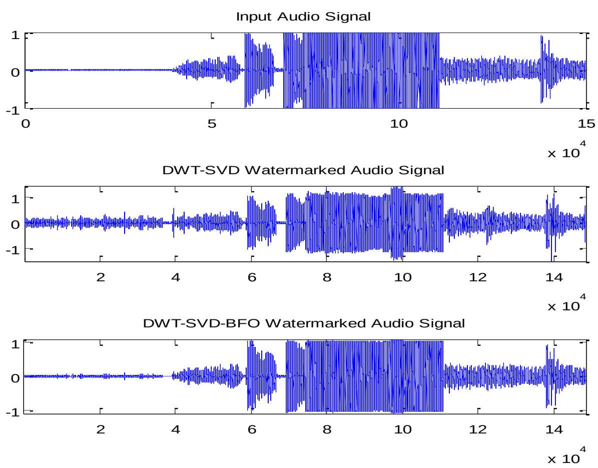


Figure: (a) Recorded input signal (b) DWT-SVD watermarked signal (c) DWT-SVD-FIREFLY watermark signal

The watermark message embedded is show in figure 5.2. the message is embedded into audio signal and depth of embedding is based on the gain factor used. The gain factor should be optimal so that there is a tradeoff between PSNR and MSE as discussed in previous chapter. For this purpose a bio optimized algorithm named firefly algorithm (FA) is used which gives the tuned gain factor value which results in high PSNR and low MSE. The FA parameters used for the purpose are tabulated in table 5.1. The variables alpha, beta and gamma, for the movement of firefly and searching space dimension are important for a fine tuning and it can be observed by plotting the objective function value plot for whole iterations. With increase in iteration number the value of objective function should decrease and if it settles to a minimum value after certain iterations and doesn't change further then that would be a case of best optimization of FA. Figure 5.3 shows the objective function values plot for iterations in our case for the embedding message in figure 5.2.

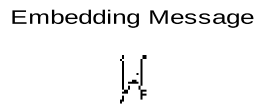


Figure : Embedding message with dimension 27*22

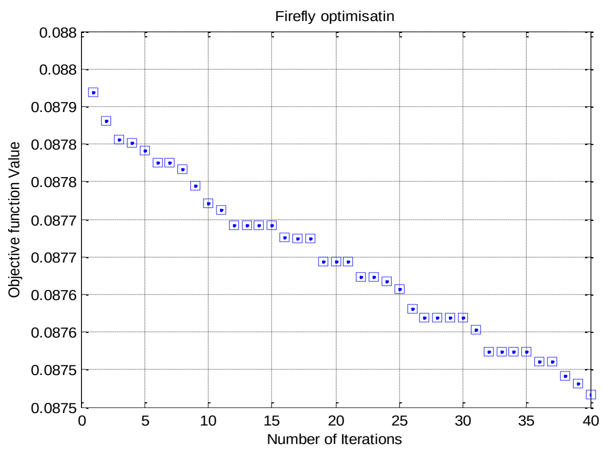


Figure: Objective function values plot

Table: FA initialization parameters

No of Fireflies	4
No of iterations	10*4
alpha	0.3
beta	0.2
gamma	0.09

The searching space dimension of firefly in FA depends upon the number of tuning variables in the application used. In our case the input audio signal is divided into 50 chunks and for each chunk a different gain factor is used which gives optimal results for embedding. Thus we have 50 gain values for each chunk instead of single common value as in the case of DWT-SVD audio signal embedding process. An output comparison of NCC, PSNR and MSE is shown in figure (a), (b) and (c) respectively. For DWT-SVD embedding process the gain factor has been fixed at 10 and in proposed it is 50 in numbers having different

IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

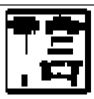
ISO 3297:2007 Certified

Vol. 3, Issue 11, November 2016

values with mean of 2.0637. Results have been checked for different watermark message of varying size with same recorded audio signal. Results are shown in table. The size of watermark message used for iteration isn't chosen so large as FA is an iterative process and large message size will increase the execution time of algorithm.

Figure (a): NCC comparison of DWT-SVD and DWT-SVD-FA (b) PSNR Comparison (c) MSE Comparison

Table: Evaluation of proposed scheme for different watermark message

	Input Audio	Watermark Message	DWT-SVD			DWT-SVD-FA		
			PSNR	NCC	MSE	PSNR	NCC	MSE
1	Recorded Input signal of frequency 44100 Hz	Size=27*22	7.2973	0.17689	0.18631	7.3959	0.18631	0.18213
2		Copyright Size=20*50	20.0076	0.9219	0.0099818	34.1841	0.99844	0.00038156
3		 Size=36*39	5.2703	0.25985	0.29713	5.4447	0.29179	0.28543

V. CONCLUSION

This thesis proposes a new algorithm considering the tuning of gain factor for embedding of watermark message in the audio signal. Proposed algorithm is based on quantization in DWT domains with SVD while considering the more active components of the signal. The performance of the algorithm is provided by evaluating the performance parameters such as peak signal to noise ratio, normalized correlation, and mean square error. From the results it is inferred that proposed algorithm is more robust than the DWT-SVD. The performance of the algorithm is improved by using the tuning of gain factor depending upon the number of chuncks of audio signal in the embedding process. Choosing proper gain value and wavelet filters have considerable effect on the performance of the algorithm.

Wavelet filters' decomposition level are considered to evaluate the algorithm performance completely. From studies level 4 produces better results bringing a tradeoff between PSNR and MSE.

VI. FUTURE SCOPE

- The audio watermarking is relatively new and has wide scope for research. This thesis is limited to binary image embedding and can be continued to gray scale images.
- The technique can be implemented on live signals rather than a fixed signal as considered in this thesis. Some of the real time audio signals include speech and conversation of pilot with ground controllers.
- Further research can be carried on embedding watermark in video sequences i.e. movies or surveillance. Applying watermarking technique on the surveillance system will decrease the security issues by keeping track of the voice communication.
- One other application that can be targeted is the watermarking of the live objects such as a person taking his tone and image.
- The research can be extended by developing watermarking technique using neural networks.

REFERENCES

[1]. Komal V. Goenka, Pallavi K. Patil," Overview of Audio Watermarking Techniques" International Journal of Emerging Technology and Advanced Engineering, Volume 2, Issue 2, February 2012

[2]. Ali Al-Haj," An imperceptible and robust audio watermarking algorithm" EURASIP Journal on Audio, Speech, and Music Processing 2014.

[3]. Darabkh, K.,"Imperceptible and Robust DWT-SVD-Based Digital Audio Watermarking Algorithm", Journal of Software Engineering and Applications, 2014.

[4]. Yekta Said Can, Fatih Alagoz, Melih Evren Burus," A Novel Spread Spectrum Digital Audio Watermarking Technique" Journal of Advances in Computer Networks, Vol. 2, No. 1, March 2014

[5]. Hwai-Tsu Hu, Hsien-Hsin Chou, Chu Yu and Ling-Yuan Hsu," Incorporation of perceptually adaptive QIM with singular value decomposition for blind audio watermarking" EURASIP Journal on Advances in Signal Processing 2014