Secure Audio Watermarking based on Haar Wavelet and Discrete Cosine Transform

Varun Mehta
M.Tech (ECE)
BUEST, Baddi
Himachal Pradesh, India

Neha Sharma
Assistant Professor
BUEST, Baddi
Himachal Pradesh, India

ABSTRACT
All sorts of information today is getting digitized and internet among others is the easiest and a very effective source and host for sharing data. So, it becomes very important to track, protect and monitor the data and maintain its security, confidentiality, data integrity, authentication, robustness and ownership. And, techniques like steganography, watermarking and cryptography help us to do so. In steganography, information is hidden inside the data (like image, audio, video, text) and the confidentiality is maintained as only the sender and the receiver know how to retrace the information. In this paper, a secure way of Audio- Watermarking is implemented using Discrete Wavelet Transform (Haar Wavelet) and Discrete Cosine Transform.

General Terms
Watermarking, audio signal, bitmap image, copyrighting.

Keywords
Discrete Wavelet Transform, Discrete Cosine Transform, Arnold Transform.

1. INTRODUCTION
Information, today, is mostly transmitted in a digital format, and the growth in this trend will not fade in the foreseeable future. Digital data is susceptible to duplicity with the same quality as the original. And security of data is an issue. Though it can be achieved by cryptography, steganography or watermarking. Cryptography deals with encryption which is defined as protecting the information by encrypting it into unrecognizable format. It doesn't hide the existence of the message from the attacker instead it renders the content of the message garbled to unauthorized people. Steganography and watermarking help in data hiding. Steganography hides the information in such a way that the existence of information is undetectable. It requires a cover file that can hide the message that is to be transmitted. Cover may be an image, audio, text file or even video. Hidden information may be a text file, a cipher text, audio or image. Stego key is required for the embedding and extracting the secret message. Embedding a message in a cover file is carried out in a way that the quality of the cover file is not compromised. But watermarking hides only the legal information inside the carrier for copyright protection. It embeds information into carrier in a way that its removal becomes impossible. It can also be used for authentication. Watermarks, originally, were faint imprints on paper that could only be seen at a specific angle. This process was designed to prevent counterfeiting and is still in use. Printed watermarks though are intended to be proportionally visible but digital watermarks are designed to be completely invisible, and in the case of audio files, totally inaudible.

This paper is further organized as follows: section II details the audio watermarking and its methodologies. Section III deals in the techniques used in audio watermarking, section IV deals with the proposed framework for audio watermarking and the transforms used for the audio watermarking, section V and VI explain the algorithm used for encoding and decoding respectively and finally, section VII shows the simulation results and VIII concludes the paper.

2. AUDIO WATERWARKING
Dynamic supremacy of human auditory system (HAS) over human visual system (HVS) makes audio watermarking quite challenging as compared to image watermarking. The audio watermark should be inaudible, statistically unnoticeable to prevent unauthorized removal and robust to intentional signal processing attacks such as compression, filtering, re-sampling, noise adding, digital-to-analog/analog-to-digital conversion, etc, and self-clocking for the ease of detection in the presence of time scale modification attacks. Various watermarking methodologies are discussed below:

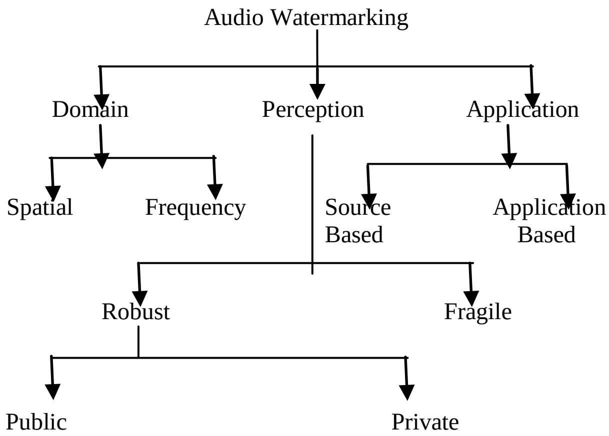


Fig. 1. Audio Watermark Methodologies

3. WATERWARKING TECHNIQUES
The audio watermarking can be classified into temporal watermarking and spectral watermarking, based on spatial and transform domain [1] where watermarks are inserted. Temporal watermarking helps in hiding watermarks directly into digital audio signals in the time domain, and spectral watermarking methods first transform the given audio signal, where FFT (Fast Fourier Transform), DCT (Discrete Cosine Transform) [2], and DWT (Discrete Wavelet Transform) [3], etc. are commonly used as the underlying transform, and hides watermarks in the transform domain.

3.1 Spatial Domain Methods

Spatial domain methods hide information on the basis of geometric characteristics of audio signal. Majority of these methods employ LSB techniques. Low Bit Encoding and Echo Hiding are among popular techniques used under spatial domain. Least Significant Bit techniques are an easy way to hide information. Tolerance to noise at low levels has been achieved through different LSB methods but with a low hiding capacity. But spatial domain methods lack in maintaining characteristics like robustness and security of the watermarked signal.

3.2 Transform Domain Methods

Transform domain methods help in embedding information along the frequency distribution of the carrier signal. Human auditory system has several characteristics which can be exploited by various methods of transform domain to hide data. Spread spectrum technique produces redundant copies of the data signal using M-sequence and embedded in the audio carrier.

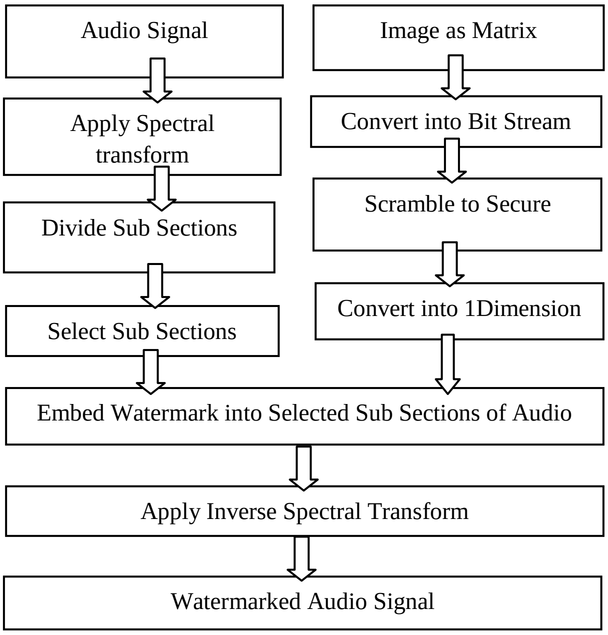


Fig. 2. Watermark Embedding

DWT (Discrete Wavelet Transform) technique is used to hide data in transform coefficients of the audio signal. DWT was developed as an alternative to short time Fourier Transform. DWT actually decomposes the audio signal into many multi resolution sub-bands, which in turn helps to locate the most appropriate sub-bands for embedding bits of watermark gray image. Tone insertion is the indirect exploitation of the psychoacoustic masking phenomenon. Psycho-acoustical or auditory masking is actually the characteristic of human auditory system (HAS) where the presence of stronger tone renders the weaker tone in its spectral domain [4].

The masked sound becomes inaudible in the presence of another louder sound. And, phase coding exploits the fact that Human Auditory System can't recognize the change in phase as easily as it can recognize the noise in the signal. Research shows that hiding data in transform domain gives much better results than hiding data in spatial domain.

4. TRANSFORMS USED

A significant number of watermarking techniques have been reported in recent years in order to create robust and imperceptible audio watermarks. In [2], authors explain a blind watermarking method which inserts watermarks into Discrete Cosine Transform (DCT) coefficients using quantization index modulation method. In [4], authors propose a blind audio watermarking method by using adaptive quantization against synchronization attack. In addition the multi resolution characteristics of discrete wavelet transform (DWT) and the energy compression characteristics of discrete cosine transform (DCT) are combined in this method. It also improves the transparency of digital watermark. Watermark is then inserted into low frequency components using adaptive quantization according to human auditory system (HAS).

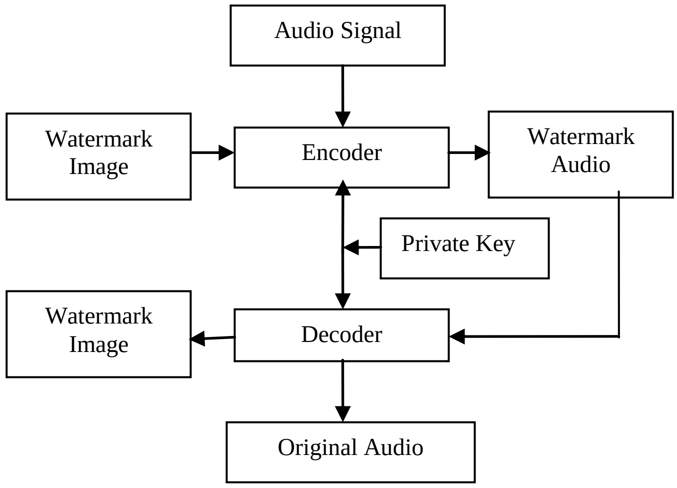


Fig. 3. Watermark Encoding and Decoding

In [5], the authors introduce a new method for inserting watermarks into audio signals by using components which have low frequency. This proposed method uses lifting-based wavelet transform (LBWT). The novel technology of inserting image data into the audio signal and additive audio watermarking algorithm based on DCT domain is proposed in [6]. In [7], the authors propose a watermarking method where DCT coefficients are used.

In [8] and [9], the authors propose a new watermarking system using Discrete Fourier Transform (DFT) and First Fourier Transform (FFT) of digital contents for copyright protection respectively. In [10], the authors propose a novel method to generate a unique personalized watermark which can act authenticating information for any watermarking algorithm.

In [11], the authors describe the watermark construction process where scrambled version of watermark is obtained with the help of Arnold transform. This transform generates a private key which helps in making the watermark image non-detectable even if it is extracted from the audio signal.

4.1 Discrete Wavelet Transform

The wavelets were introduced in 1982 by a French geophysicist named Jean Morlet [12]. The two types of wavelets that are commonly used are Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT). A Continuous Wavelet Transform (CWT) is used to divide a continuous-time function into wavelets. Discrete wavelet transform (DWT) are those wavelets which discretely samples the wavelets. Wavelets are special functions which, in a way analogous to sine and cosines in Fourier analysis, and are used as basal functions for representing signals. They provide powerful multi-resolution tool for the analysis of non-stationary signals with good time localization information. Due to its excellent spatio-frequency localization properties, the DWT is very suitable to identify areas in an audio signal where a watermark can be embedded effectively. The principle objective of the wavelet transform is to hierarchically decompose an input signal into a series of successively lower frequency approximation sub band and their associated detail sub bands. For DWT, the link between the spatial/temporal domain signals, f (t), and the DWT of f(t), d (k, l), is Where ψ (•) denotes the mother wavelet

$$f(t)=\sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} d(k,l)2^{-k/2}\psi(2^{-k}t-l).$$

4.1.1 Haar Wavelet

Haar wavelet filter is one of the most basic and important filter of orthogonal type. In this, the filters used for decomposition and reconstruction of image are orthogonal to each other. The mathematical formulae for the mother wavelet function and scaling function are as:

$$\Phi(t) = \begin{cases} 1 \text{ for } 0 \le t \le 1, \\ 0 \text{ Otherwise} \end{cases}$$

$$\psi(t) = \begin{cases} 1\ for\ 0 \le t \le \frac{1}{2} \\ -1\ for\ \frac{1}{2} \le t \le 1 \\ 0\ otherwise \end{cases}$$

Where ϕ (t) and Ψ (t) is the scaling function and mother wavelet function respectively.

4.2 Discrete Cosine Transform

The Discrete Cosine Transform is a technique for converting a signal into elementary frequency components. The most common DCT definition of a 1-D sequence of length N is

$$C(u)=\alpha(u)\sum_{x=0}^{N-1} f(x)\cos\left[\frac{\pi(2x+1)u}{2N}\right]$$

For u=0,1,2,….,N-1. Similarly, the inverse transform is defined as,

$$f(x)=\sum_{u=0}^{N-1}\alpha(u)C(u)\cos\left[\frac{\pi(2x+1)u}{2N}\right]$$

For x=0, 1, 2,….N-1. In both equations (1) and (2) α (u) is defined as,

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{N}} & for\ u = 0 \\ \sqrt{\frac{2}{N}} & for\ u \neq 0 \end{cases}$$

The first transform coefficient is, referred to as DC Coefficient, the average value of the sample sequence. The other transform coefficients are called the AC Coefficients.

4.3 Arnold Transform

Arnold Transform is named after V.I. Arnold [13].It realigns the pixel matrix of an image. Image scrambling based on Arnold Transform changes the position of pixels resulting in a disordered image. Mathematically, it can be defined as:

$$\begin{bmatrix} a' \\ b' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} (\text{mod } N_l)$$

Where *x* and *y* are the coordinates of the original image; *N* is the order of image matrix; *x*' and *y*' are the coordinate of the transformed image. This technology of image scrambling is commonly used due to its periodicity. If the parameters which are used at the time of scrambling are known to us, then we can easily restore the original image by using reverse scrambling. Scrambling can be used for data hiding techniques. So, if it is hacked by some attacker, only the scrambled image is shown. Arnold transform period is given as Ts N2 / 2. We have considered an image of size 32x32 and its Arnold period is 24. Therefore the Private Key can be entered accordingly. Watermark Encoder and Decoder. The encoding and decoding algorithm is explained in the following sections.

5. ENCODING

Encoding is the process of embedding a watermark image on an audio signal in such a way that there are no audible changes in the audio and it seems exactly like the original audio. The process can be represented as:

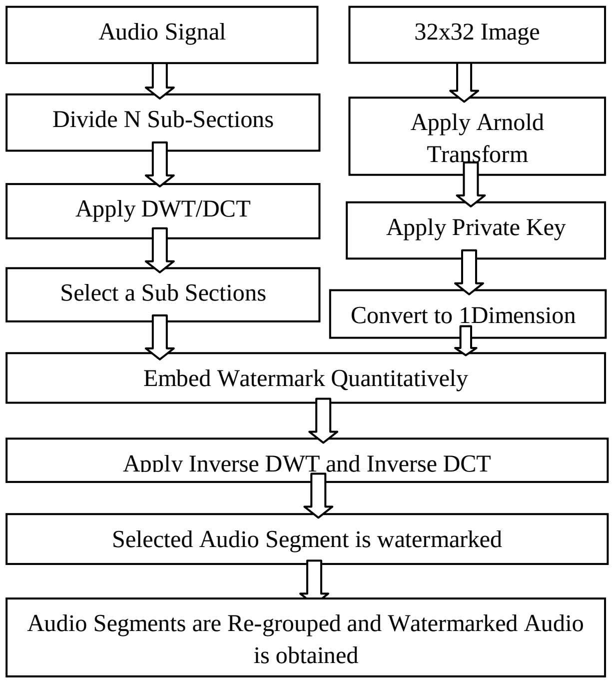


Fig. 4. Watermark Encoding Process

Encoding is the process of embedding a watermark image on an audio signal in such a way that there are no audible changes in the audio and it seems exactly like the original audio.

Encoding Algorithm is explained as:

Step 1: Select the original audio. Here, we have preferred *.wav or *.WAV format.

Step 2: Select the watermark image and convert it into a binary image and resize it to 32x32 matrices.

Step 3: A private key with Arnold transformation is applied on image S and Scrambled Secret Image is obtained.

Arnold Period for a 32x32 image is 24.

Step 4: The watermark is converted into a vector.

Step 5: The audio is divided into the various segments of fixed length.

Step 6: Apply the H-level DWT. H is selected such that length of low frequency coefficients should be MN.

Step 7: Do the DCT on low frequency coefficients.

Step 8: Embed the Watermark bits as per the Quantization Function selected.

Step 9: Apply the Inverse DCT and Inverse DWT on the quantized low frequency coefficients.

Step 10: The divided segments of Audio are re-arranged into a single audio and saved it as a wAudio1, wAudio2 and wAudio3.

6. DECODING

Decoding is done at the receiver's end to extract the embedded watermark image. Schematically it can be represented as:

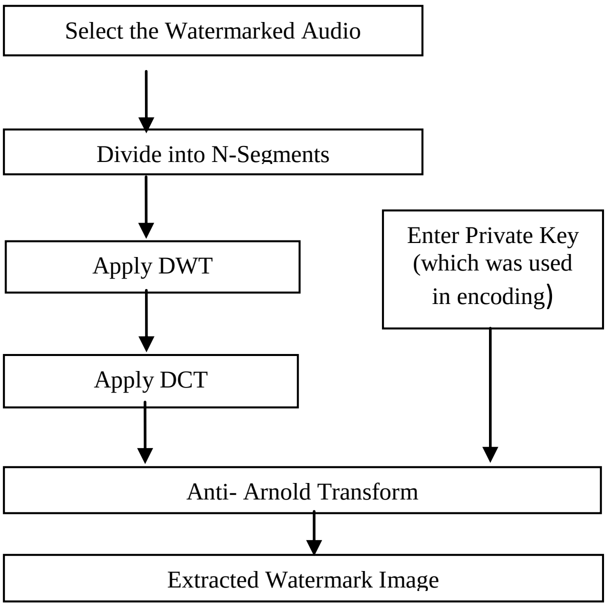


Fig. 5. Watermark Decoding Process

Decoding Algorithm is explained as:

Step 1: Select the watermarked audio which was saved after the encoding process.

Step 2: Divide the audio into various segments as it was done during encoding.

Step 3: Apply DCT and DWT again on the selected segment.

Step 4: Now, as the quantization was used to embed the watermark on the audio, according to that quantization, apply the opposite and extract the watermark.

Step 5: Entered the private key or otherwise the watermark extracted will be a scrambled image.

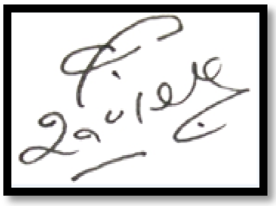
Step 6: Once the watermark is extracted, apply various attacks like Compression, White Gaussian and filtering and check SNR, MSE, NCC and BER.

Step 7: Observe the calculated parameters and draw the conclusion.

7. SIMULATION RESULTS

This section presents the results obtained from implementation of audio watermarking based on DWT/DCT and Arnold Transform in MATLAB R2012a. MATLAB is a high-level language which provides interactive environment for numerical computation, programming and visualization.

Our experimental work considers three coloured images which are converted into 32x32 bitmap images: Signature, Logo and Warning. Three audio signals Audio 1, Audio 2 and Audio 3 are also taken on which watermarking is tested.

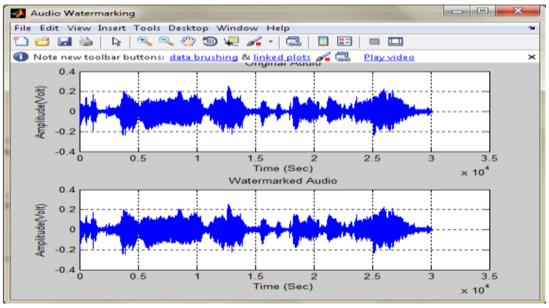


a) Signature **b) Logo**

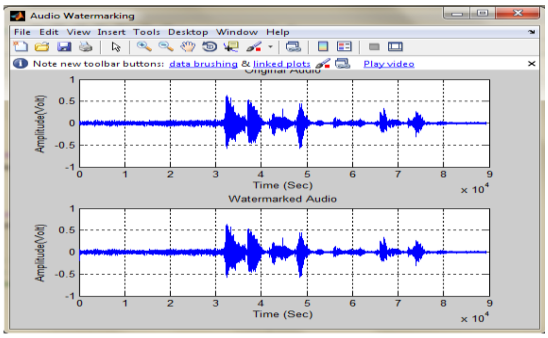
c) Warning

Fig. 6. Watermark Images Used

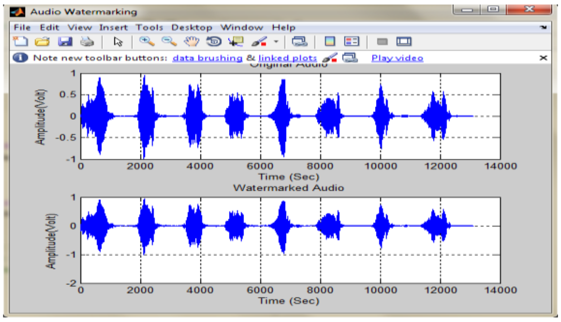
Figure 6 (a-c) shows three coloured images named Signature, College Logo and Warning. These images are converted to 32x32 bitmap images before being applied as watermark images.



a) Audio 1 watermarked with 'Signature'



b) Audio 2 watermarked with 'Logo'



c) Audio 3 watermarked with 'Warning'

Fig. 7. Original and Watermarked Audio

Figure 7 shows the histogram of original audios and watermarked audios. It is observed that histograms of both the audios are same. Hence, the watermarked audios obtained after encoding process look similar to their respective original audios. Therefore, they do not attract the attention of attackers. Hence, watermarked audio is safely transmitted and heard. Though, some attacks like White Gaussian Noise, Filters and Compression are applied here to check the robustness of the watermarked audio and parameters like Signal to Noise Ratio (SNR), Mean Square Error (MSE), Normalised Coorelation Coefficient (NCC) and Bit Error Rate (BER) are calculated here for all the three watermarked images.

7.1 Audio Quality Parameters

7.1.1 Signal to Noise Ratio

Signal to noise ratio is a parameter used to know the amount by which the signal is corrupted by the noise. It is defined as the ratio of the signal power to the noise power. Signal to noise ratio can also be calculated by equation below. Z is the un-watermarked audio signal and *Z'* is the watermarked audio signal. Both *Z* and *Z'* has *Mt* samples.

$$SNR = 10\log\left(\frac{\sum_{a=1}^{M_t} Z^2(a)}{\sum_{a=1}^{M_t}\left(Z(a)-Z'(a)\right)}\right)$$

7.1.2 Mean Square Error

The mean squared error (MSE) of an estimator is one of many ways to quantify the difference between values implied by an estimator and the true values of the quantity being estimated. MSE is a risk function, corresponding to the expected value of the squared error loss or quadratic loss. MSE measures the average of the squares of the "errors." The error is the amount by which the value implied by the estimator differs from the quantity to be estimated. The difference occurs because of randomness or because the estimator doesn't account for information that could produce a more accurate estimate. Where xi is watermarked audio and yi is original audio. N is no. of samples.

$$MSE(x,y) = \frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2$$

7.1.3 Normalized Correlation Coefficient

In order to eliminate the subjective factors and reflect the fairness of copyright protection, we employed the normalized correlation coefficient (NC) to estimate the similarity between the original watermark and the extracted watermark. The normalized correlation coefficient (NC) is defined as equation.

$$NC = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} Y(i,j)Y'(i,j)}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N} Y(i,j)^2}\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N} Y'(i,j)^2}}$$

7.1.4 Bit Error Rate

Bit error rate can be defined as the percentage of bits corrupted in the transmission of digital information due to the effects of noise, interference and distortion.

$$BER = \frac{B_{err}}{M \times N} \times 100\%$$

Where, B_{err} is the number of error bits and refers to the size of the image (totalling the number of bits in the image).

Table 1. Comparison of SNR and MSE values of wAudio1 with and without Attacks

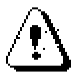
Watermarked Audio 1	Parameter	Value	Watermark Image
No Attack	MSE	0.0000	
	SNR	Inf dB	
	NCC	0.0010	
	BER	0.0000 %	
White Gaussian Noise	MSE	0.0000	
	SNR	110.7156 dB	
	NCC	0.0010	
	BER	0.0000 %	
Compression	MSE	0.0000	
	SNR	150.5663 dB	
	NCC	0.0010	
	BER	2.4414 %	
Low Pass Filtering	MSE	0.0000	
	SNR	132.1964 dB	
	NCC	0.0010	
	BER	0.4883 %	

Table 2. Comparison of SNR and MSE values of wAudio2 with and without Attacks

Watermarked Audio 2	Parameter	Value	Watermark Image
No Attack	MSE	0.0000	
	SNR	Inf dB	
	NCC	0.0012	
	BER	0.0000 %	
White Gaussian Noise	MSE	0.0000	
	SNR	101.7308 dB	
	NCC	0.0012	
	BER	0.0000 %	

Compression	MSE	0.0000	
	SNR	168.3001 dB	
	NCC	0.0012	
	BER	0.0000 %	
Low Pass Filtering	MSE	0.0000	
	SNR	140.6164 dB	
	NCC	0.0012	
	BER	0.0000 %	

Table 3. Comparison of SNR and MSE values of wAudio3 with and without Attacks

Watermarked Audio 3	Parameter	Value	Watermark Image
No Attack	MSE	0.0000	
	SNR	Inf dB	
	NCC	0.0013	
	BER	0.0000 %	
White Gaussian Noise	MSE	0.0000	
	SNR	110.4424 dB	
	NCC	0.0013	
	BER	10.7422 %	
Compression	MSE	0.0000	
	SNR	149.8636 dB	
	NCC	0.0010	
	BER	46.1914 %	
Low Pass Filtering	MSE	0.0000	
	SNR	164.2207 dB	
	NCC	0.0013	
	BER	0.4883 %	

The experimental results have illustrated the robust nature our watermarking scheme. In addition, the watermark can be extracted without the help from the original digital audio signal and can be easily implemented. To check the robustness SNR, MSE, NCC and BER is calculated. After embedding watermark, the SNR of all selected audio signals using the proposed method are above 20 dB which ensures the imperceptibility of the proposed system. This satisfies the IFPI requirement (20 dB).The value of MSE is always less than 1 for each audio file. The value of NCC of most of the audio files was 1 and BER was almost 0. It can be seen that using the proposed algorithm to embed watermark is inaudible and is robust to some common attacks like compression, noise and low-pass filter. We can extract watermark and it has good legibility.

8. CONCLUSION AND SCOPE

The watermarking of audio data is an appropriate mechanism to protect the intellectual property rights. In this scheme we scramble the watermark with Arnold transformation and embed quantitatively it into the audio signal. Before embedding the watermark first we apply DCT and DWT on the audio file. Audio watermarking schemes can prove really valuable for copyright control of digital material. Available studies on audio watermarking are far less than that on image watermarking or video watermarking. Although, audio watermarking studies have advanced considerably during the last a few years. Those studies have contributed much to the progress of audio watermarking technologies. Recent research works have investigated many new audio watermarking techniques which preserve the integrity of the hidden data and its securing. In this work, one such approach has been presented. Also the strengths and weaknesses are portrayed. Also, it can be concluded that the transform domain is preferred over the spatial domain in terms of imperceptibility, non-detection and capacity. However, more sophisticated technologies are required. As every technique is different from another one, comparison among them cannot be done.

9. REFERENCES

[1] Ifra Bilal, Mahendra Singh Roj, Rajiv Kumar and P K Mishra (2014) “Recent Advancement in Audio Steganography”, 2014 International Conference on Parallel, Distributed and Grid Computing (978-1-4799-7683-6/14/$31.00©2014 IEEE, 402)

[2] G. Zeng and Z. Qiu, (2008) “Audio watermarking in DCT; Embedding Strategy and Algorithm”, in Proceedings of 9th International Conference on Signal Processing, (ISP’09), pp 2193-2196.

[3] Ali Al-Haj, Ahmad Mohamma1 and Lama Bata (2011) “DWT–Based Audio Watermarking”, The International Arab Journal of Information Technology, Vol. 8, No. 3, July 2011

[4] X. Y. Wang and H. Zaho, (2006) “A novel Synchronization Invariant Audio Watermarking Scheme Based on DWT and DCT ”, IEEE Transactions on Signal Processing, Vol. 54, No. 12, pp 4835-4840.

[5] E. Ercelebi. And L.Batakci., (2009) “Audio Watermarking Scheme Based on Embedding Strategy in Low Frequency Components with a Binary Image.”, Digital Signal Processing, Vol. 19, No. 2, pp 265-277.

[6] Yan Yang, Rong Huang and Mintau Xu (2009) “A Novel Audio Watermarking Algorithm for Copyright Protection Based on DCT Domain.” 2nd International Symposium on Electronic Commerce and Security ,pp 184-188.

[7] P.K Dhar, M.I.Khan and S. Ahmad (2010) “A New DCT-Based Watermarking Method for Copyright Protection of Digital Audio.”, International Journal of Computer Science & Information Technology(IJCSIT), Vol. 2 No.5.

[8] P.K Dhar, M.I.Khan and J.M.Kim (2010) “A New Audio Watermarking System using Discrete Fourier Transform for Copyright Protection.”, International Journal of Computer Science & Network Security (IJCSNS), Vol. 10 No.6.

[9] P.K Dhar and J.M.Kim (2011) “Digital Watermarking Scheme Based on First Fourier Transformation for Audio Copyright Protection.”, International Journal of Security and its Applications (IJSIA), Vol. 5 No.2.

[10] T.K. Tewari , V.Saxena and J.P Gupta (2012) “A Novel Approach to Generate Watermarks Using Auditory Features for Audio Watermarking.”, Journal of

Theoretical and Applied Information Technology (JATIT), Vol. 35 No.2.

[11] Dr. M. Sathik and S. S. Sujatha (2010) “An Improved Invisible Watermarking Technique for Image Authentication.”, International Journal of Advanced Science and Technology (IJAST), Vol. 24.

[12] M. Sifuzzaman, “Application of Wavelet Transform and its Advantages Compared to Fourier Transform”,Journal of Physical Sciences, Vol. 13, 2009, 121-134.

[13] Lingling Wu et al., “Arnold Transformation Algorithm and Anti-Arnold Transformation Algorithm”,the 1st International Conference on Information Science and Engineering (ICISE2009).

[14] Ravula, R., Rai, S. (2010), “Audio watermarking using transformation techniques”, Electronic thesis and desertion, lauisian state university, etd-10182010-105045.

[15] Singh, J., Garg, P.and De, A.N. (2011), “Audio watermarking based on quantization index modulation using combined perceptualmasking”, © Springer Science + Business Media, LLC.

[16] Yang, S., Tan, W., Chen, Y. and Ma, W. (2010), “Quantization based digital audio watermarking in Discrete Fourier domain”, Journal of Multimedia, Vol. 5, No. 2.