

AUDIO WATERMARKING WITH ANGLE QUANTIZATION BASED ON DISCRETE WAVELET TRANSFORM AND SINGULAR VALUE DECOMPOSITION

¹Divya Chadar, ²Shailja Shukla

¹M. Tech. Scholar, Control system, Jabalpur Engineering College, Jabalpur, India.

² Prof and HOD of Computer Science Department, Jabalpur Engineering College, Jabalpur, India.

Email: divvyachadar21@gmail.com, shailja270@gmail.com

Article History: Submitted on 18th January, Revised on 1st February, Published on 3rd March 2018

Abstract. An audio watermark is a unique electronic identifier embedded in an audio signal, typically used to identify ownership of copyright. Proposed work is a new method of audio watermark hiding inside another bigger cover standard audio cover. The method includes 'harr' wavelet based Discrete Wavelet Transform decomposition of frequencies hence the audio samples of watermark gets hidden only those parts of cover audio where human ears are less sensible according to Human Auditory System. Proposed method also includes the Singular Value Decomposition, which is required for making our method robust against the various communication of processing attacks like compression, filtering, fading or noise addition. Proposed work is also using the concept of angular modulation which initially modifies the audio watermark in to provide extra security and also extra robustness in communication. The design is been develop on MATLAB 2013b version and verification of design o the same.

Keywords. DWT: Discrete Wavelet Transform, SVD: Singular Value Decomposition , AQ- angle quantization, HAS-human auditory system.

INTRODUCTION

Audio watermarking is a well-known technique of hiding data through audio signals. It is also known as audio watermarking and has received a wide consideration in the last few years. So far, several techniques for audio watermarking have been discussed in literature by considering different applications and development positions. A perceptual property of human auditory system (HAS) help to hide multiple sequences of audio through a transferred signal. However, all watermarking techniques face to a problem: a high robustness does not come with a high watermark data rate when the perceptual transparency parameter is considered as fixed. Furthermore, selection of a suitable domain, cover, and considering the problems associated with data-hidden techniques must be considered for designing the path to achieve a data-hidden purpose. The remainder of this chapter is organized as follows: Transmission channel for audio watermarking is discussed. Different audio watermarking attacks are explained. Various audio watermarking techniques are compared. Some basic problem with watermarking as Network surveillance and monitoring systems will not flag messages or files that contain watermarked data. Therefore, if someone attempted to steal confidential data, they could conceal it within another file and send it in an innocent looking email. Another issue is Large amount of data has to be transmitted which arises suspiciousness to the intruder. The digital watermarking is a technique to embed information of difference from the digital contents in itself so that a human cannot perceive it without almost affecting the quality of contents using perception properties such as the human vision and hearing. In general, audio watermarking methods should satisfy four requirements: inaudibility, robustness, data payload, and blindness. There is a trade-off among these requirements. For example: increasing data payload in a watermarking system results in quality degradation of the watermarked signal and decrease of robustness against signal processing attacks. It is difficult to achieve a good trade-off among inaudibility, robustness, and data payload. Therefore, the solution to this problem is the main purpose to study digital watermarking at present. Moreover there are different type of attacks which reduces the SNR and enhance the MSE because of this attacks BER increases, list of attacks are as follow:-

Filter: some time filter may remove the watermark content from the audio which may cause bad quality watermark after extraction from cover.

Noise: in long distance communication noise may affect watermark content which may cause bad quality watermark after extraction from cover.

Lossy compression: compression may remove watermark samples which may cause bad quality watermark after extraction from cover.

All above attacks can be resolved with SVD because SVD hides every sample of watermark at singular component of cover which does not change easily, SVD also hides samples at distributed locations.

METHODOLOGY

Proposed work is a new method of audio watermark hiding inside another bigger cover standard audio cover. The ratio of size of the watermark and cover is 1:48, means 1 watermark audio sample will be hidden between 48 samples of cover audio. The method includes 'harr' wavelet based DWT decomposition of frequencies hence the audio samples of watermark get hidden only those parts of cover audio where human ears are less sensible according to HAS. Proposed method also includes the SVD decomposition which is required for making our method robust against the various communication or processing attacks like compression, filtering, fading or noise addition. Proposed work is also hiding each sample of watermark into SVD decomposed U, S and V elements, hence three copies of watermark samples get hidden and the SVD decomposition performs on 4x4 blocks of DWT decomposed cover audio so each sample gets hidden at 4x4=16 locations and 3 copies of each means total 3x16=48 locations, and in communication if due to any kind of attacks some of these watermark samples get modified still we can reconstruct the original watermark out of maximum unchanged watermark samples. Proposed work is also using the concept of angular modulation which initially modifies the audio watermark in to provide extra security and also extra robustness in communication. It basically converts the audio samples from rectangular to polar formats which need to be reconstructed at the receiver end and it completely modifies the watermark audio hence if somehow any intruder extracts the hidden data (i.e. watermark) from the cipher audio, still intruder will not have the correct audio watermark. This concept makes proposed method more secure.

Figure 1 shown below is proposed block diagram for audio watermark hiding inside the cover audio file, steps are as follow:-

Step 1: input Cover audio file and generate its samples into MATLAB environment.

Step 2: DWT decomposing of cover audio file up to two levels

Step 3: convert 1D audio file into 2D audio matrix

Step 4: SVD performed on 4x4 blocks of 2D audio matrix and develop U, V and S which are also with the same 4x4 size

Step 5: input Watermark audio file and generate its samples into MATLAB environment

Step 6: perform logical XOR with the KEY after scaling the samples and then rescale the samples to its original level.

Step 7: perform Angle quantization and generate quantized audio samples

Step 8: quantized watermark hiding for embedding watermark Interchange each sample of quantized audio with any one out of 16 samples of U, S, V The position of the replacing sample is decided based on the key with modulo operation.

Step 9: perform Inverse SVD on the modified U, S and V

Step 10: develop new 4x4 blocks of 2D audio matrix

Step 11: reshape again into the 1D audio file

Step 12: perform 2 level IDWT and finally what we achieve is the cipher audio file which will be similar to cover but watermark will be hidden inside.

Proposed Algorithm for watermark hiding:

Let X is the cover Audio of Mx2 size, with M number of samples at right ear and M number of samples at left ear. Watermarking process will be same for left and right ear samples.

M/4 is need to be reshape in a perfect square shape for the size of NxN, and it is possible only if M/4 is a perfect square, so number of samples 'M' in the cover must be chosen as M/4 is a perfect square or we can select samples as required and remaining will be add after the watermark hidden.

For example: if number of samples in cover audio are M then,

$$q = \sqrt{M1}$$

$$q1 = \text{absolute of } (q)$$

$$M2 = 4 * q1^2$$

$$M3 = M - M2$$

$$M1 = \frac{M}{4}$$

M2 are the sample in cover audio, where M2/4 will be a perfect square, and M3 are the extra samples which will not participated in the process of watermark hiding. These samples will be included at the last after watermark hiding.

DWT applied on cover Audio 'A' which has M2x1 samples. Proposed work use simple 'harr' wavelet for decomposition of Cover audio as work is interested in Frequency factorization only and 'harr' wavelet is best in frequency resolution.

Harr filter coefficients for LPF are L0=1 and L1=1

Harr filter coefficients for HPF are H0=1 and H1=-1

$$x(n)_L = \frac{1}{\sqrt{2}} \sum_{k=1,2}^n x(k) + x(k+1) \quad (1)$$

$$x(n)_H = \frac{1}{\sqrt{2}} \sum_{k=1,2}^n x(k) - x(k+1) \quad (2)$$

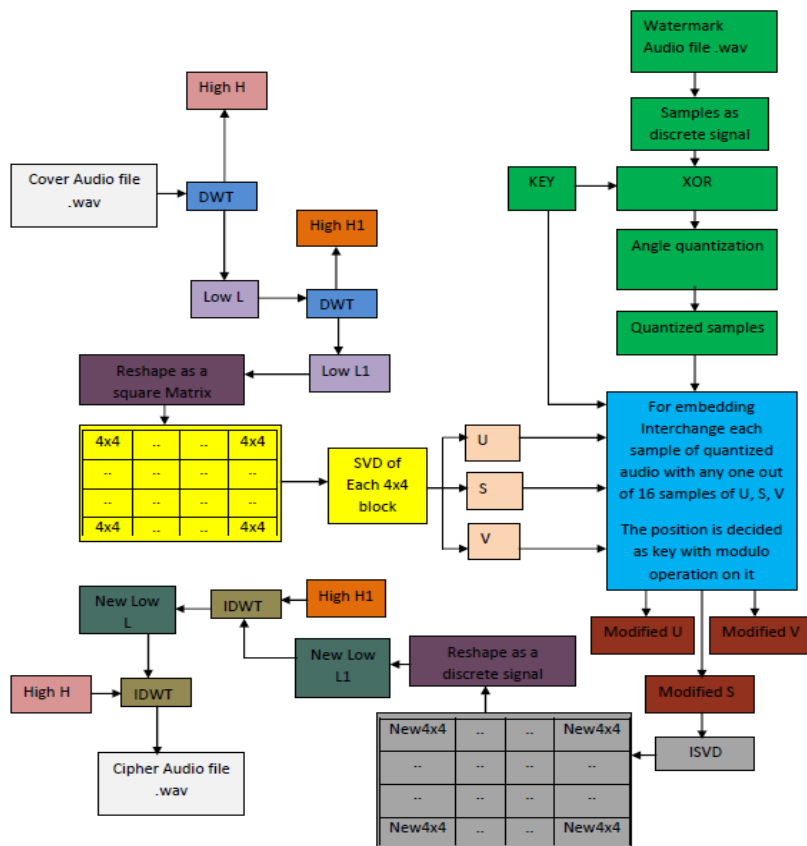


Figure 1. Watermark hiding method

$$x(n)_L = \frac{1}{\sqrt{2}} \sum_{k=1,2}^n x(k) + x(k+1) \quad (1)$$

$$x(n)_H = \frac{1}{\sqrt{2}} \sum_{k=1,2}^n x(k) - x(k+1) \quad (2)$$

Equation (1) and (2) are the level one DWT decomposition

$$x(n)_{L1} = \frac{1}{\sqrt{2}} \sum_{k=1,2}^n x(k)_L + x(k+1)_L \quad (3)$$

$$x(n)_{H1} = \frac{1}{\sqrt{2}} \sum_{k=1,2}^n x(k)_L - x(k+1)_L \quad (4)$$

Equation (3) and (4) are the level two DWT decomposition

Size of $x(n)_L$ is $M/2$ size and $x(n)_{L1}$ is $M/4$ and for left and right ear both

$$(x(n)_{RL1})_{NxN} = \text{Reshape} \left(x(n)_{L1} \frac{M}{4} \times 1 \right) \quad (5)$$

SVD taken of 4×4 block of $x(n)_{RL1}$ means will have total $(NxN)/16$ SVD blocks.

maximum number of watermark samples

$$\text{that can be hide in cover} = \frac{NxN}{16} = \frac{M/4}{16} = \frac{M}{64}$$

Here we are explaining the Calculation of SVD for $x(n)_{RL1}$ only although it is been computed for all 4×4 block of $x(n)_{RL1}$. Let $B1$ is the first 4×4 block of $x(n)_{RL1}$ the singular value decomposition of an 4×4 real matrix $B1$ is a factorization of the form USV^T , where U is an 4×4 real matrix, S is a 4×4 rectangular diagonal matrix with non-negative real numbers on the diagonal, and V is an 4×4 real or complex unitary matrix. The diagonal entries σ_i of S are known as the singular values of $B1$. The columns of U and the columns of V are called the left-singular vectors and right-singular vectors of $B1$, respectively.

$$W1 = B1x B1^T$$

$$(W1 - \sigma_i I) = 0 \quad (6)$$

For a unique set of eigenvalues to determinant of the matrix $(W1 - \sigma_i I)$ must be equal to zero. Thus from the solution of the characteristic equation, $|W1 - \sigma_i I| = 0$ we obtain eight singular values of σ_i where $i = 1, 2, 3, 4$

$$S = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 \\ 0 & 0 & 0 & \sigma_4 \end{bmatrix}$$

And if values of σ_i again put into equation (20) we obtain x_1, x_2, x_3 and x_4 value

$$(W1 - \sigma_i I)x = 0 \quad (7)$$

$$U = \begin{bmatrix} x_1 & -x_2 & -x_3 & -x_4 \\ x_4 & x_1 & -x_2 & -x_3 \\ x_3 & x_4 & x_1 & -x_2 \\ x_2 & x_3 & x_4 & x_1 \end{bmatrix}$$

And if

$$W2 = B1^T x B1$$

$$(W2 - \sigma_i I)x = 0 \quad (8)$$

And if values of σ_i again put into equation (21) we obtain x_1, x_2, x_3 and x_4 value

$$V = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ -x_4 & x_1 & x_2 & x_3 \\ -x_3 & -x_4 & x_1 & x_2 \\ -x_2 & -x_3 & -x_4 & x_1 \end{bmatrix}$$

U, S and V computed for each 4×4 block of $x(n)_{RL1}$, As explain above.

$$U_{x(n)_{RL1}}, S_{x(n)_{RL1}}, V_{x(n)_{RL1}} = \text{SVD}(x(n)_{RL1}) \quad (9)$$

On the other hand let Y is the watermark Audio with Px2 size with P number of samples at right ear and P number of samples at left ear. Watermarking process will be same for left and right ear samples.

First logical XOR with the 8 bit key 'K'

$$Z = (Y \text{ xor } K) \quad (10)$$

Angle-Quantization is a scheme to embed the watermark information into the angle of signal on polar coordinates using Quantization method. The motive behind the angel quantization is to modify the watermark before actual hiding for extra security, This scheme operation is represented as

Let Z1 and Z2 are any sample of Z

$$\theta_1 = \tan^{-1} \left(\frac{Z_1}{Z_2} \right) \quad (10)$$

$$r_1 = \sqrt{Z_1^2 + Z_2^2} \quad (11)$$

Where θ_1 is an angle Z2 and Z1, and r1 is radius.

Provided θ_1 and r1 gets hidden instead of Z1 and Z2 this scheme has a property which is unchanging against amplitude scaling because it implements polar coordinates transform.

If K is the Key then

$$T = K \bmod \theta_1(\text{degree})$$

T shows the position of sample in U, S and V which is to be replace with θ_1 first then second U, S and V will be replace by r1 and this process keep on until all samples finish of watermark as

$$MU = (U_{(T,T)} = r_1)$$

$$MS = (S_{(T,T)} = r_1)$$

$$MV = (V_{(T,T)} = r_1)$$

$$MU = (U_{(T,T)} = \theta_1)$$

$$MS = (S_{(T,T)} = \theta_1)$$

$$MV = (V_{(T,T)} = \theta_1)$$

ISVD taken as

$$(x(n)_{MRL1}) = MU * MS * MV^T \quad (12)$$

$$(x(n)_{ML1})_{\frac{M_2}{4} \times 1} = \text{Reshape}((x(n)_{MRL1})_{N \times N}) \quad (13)$$

IDWT

$$x(k)_{ML} = \frac{x(n)_{ML1} + x(n)_{H1}}{\sqrt{2}} \quad (14)$$

$$x(k+1)_{ML} = \frac{x(n)_{ML1} - x(n)_{H1}}{\sqrt{2}} \quad (15)$$

$$x(k)_M = \frac{x(n)_{ML} + x(n)_H}{\sqrt{2}} \quad (16)$$

$$x(k)_M = \frac{x(n)_{ML} - x(n)_H}{\sqrt{2}} \quad (17)$$

$X(n)$ is the final ciphered audio file which have the watermark audio hidden in original cover audio.

Figure 2 shown below is proposed block diagram for audio watermark hiding inside the cover audio file, steps are as follow:-

Step 1: input Cipher audio file and generate its samples into MATLAB environment.

Step 2: DWT decomposing of cipher audio file up to two levels

Step 3: convert 1D cipher audio file into 2D cipher audio matrix

Step 4: SVD performed on 4x4 blocks of 2D cipher audio matrix and develop U, V and S which are also with the same 4x4 size

Step 5: extract watermark samples which were replaced in U, S and V at the time of watermark hiding, the potion will be obtain with the same KEY at the time of watermark hiding.

Step 6: find maximum match samples in U, S and V and perform acerage of all 48 copies of each samples and the average value will be consider as final sample

Step 7: perform inverse angle quantization and then perform logical XOR with the KEY.

Step 8: the sample remain after all above process are the final recovered watermark audio.

Proposed Algorithm for watermark extraction: Let X is the cipher Audio of $M \times 2$ size, with M number of samples at right ear and M number of samples at left ear. DWT applied on cipher Audio 'X' which has $M \times 1$ samples, Proposed work use simple 'harr' wavelet for decomposition of Cipher audio.

$$x(n)_L = \frac{1}{\sqrt{2}} \sum_{k=1, +2}^n x(k) + x(k+1) \quad (18)$$

$$x(n)_H = \frac{1}{\sqrt{2}} \sum_{k=1, +2}^n x(k) - x(k+1) \quad (19)$$

Equation (1) and (2) are the level one DWT decomposition

$$x(n)_{L1} = \frac{1}{\sqrt{2}} \sum_{k=1, +2}^n x(k)_L + x(k+1)_L \quad (20)$$

$$x(n)_{H1} = \frac{1}{\sqrt{2}} \sum_{k=1, +2}^n x(k)_L - x(k+1)_L \quad (21)$$

Equation (3) and (4) are the level two DWT decomposition

Size of $x(n)_L$ is $M/2$ size and $x(n)_{L1}$ is $M/4$ and for left and right ear both

$$(x(n)_{RL1})_{N \times N} = \text{Reshape} \left((x(n)_{L1})_{\frac{M}{4} \times 1} \right) \quad (22)$$

SVD taken of 4x4 block of $x(n)_{RL1}$ means will have total $(N \times N)/16$ SVD blocks.

U, S and V computed for each 4x4 block of $x(n)_{RL1}$, As explain above.

$$U_{x(n)_{RL1}}, S_{x(n)_{RL1}}, V_{x(n)_{RL1}} = \text{SVD}(x(n)_{RL1}) \quad (23)$$

If K is the Key then

$$T = K \bmod \theta 1(\text{degree})$$

T shows the position of sample in U, S and V which is to be replace with $\theta 1$ first then second U, S and V will be replace by $r1$ and this process keep on until all samples finish of watermark as,

$$r1t1 = U_{(T,T)}$$

$$r1t2 = S_{(T,T)}$$

$$r1t3 = V_{(T,T)}$$

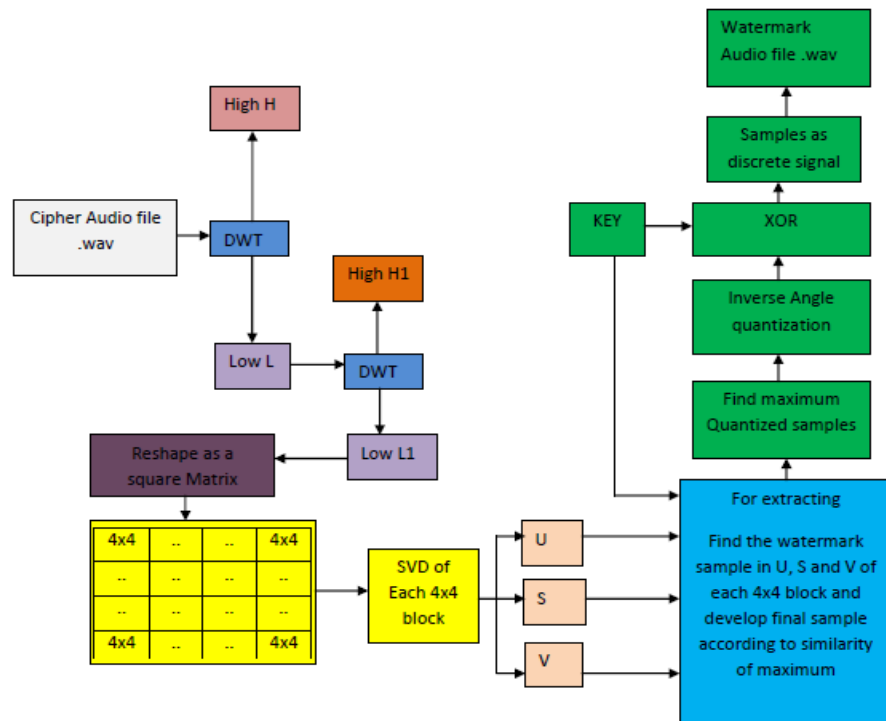


Figure 2. Watermark de-hiding method

$$r1 = \frac{r1t1 + r1t2 + r1t3}{3}$$

$$\theta1t1 = U_{(T,T)}$$

$$\theta1t2 = S_{(T,T)}$$

$$\theta1t3 = V_{(T,T)}$$

$$\theta1 = \frac{\theta1t1 + \theta1t2 + \theta1t3}{3}$$

Inverse Angle-Quantization performed on r1 and $\theta1$

Let Z1 and Z2 are any sample of Z

$$z1 = r1 * \cos(\theta1)$$

$$z2 = r2 * \sin(\theta1)$$

Where Z2 and Z1 are the samples of the watermark.

If 8 bit key is 'K' then

$$Z = (z1 \text{ xor } K)$$

Z is the final extracted watermark sample

RESULTS

The simulation results are been obtain at MATLAB 2013b

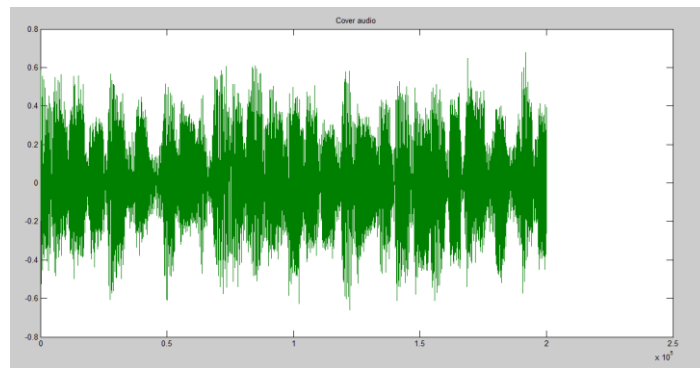


Figure 3. Cover audio

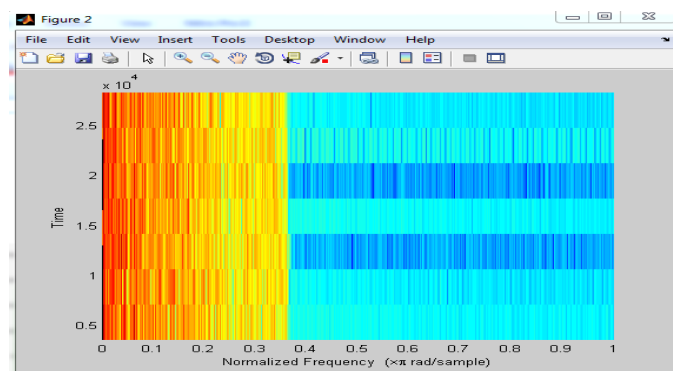


Figure 4. Cover audio spectrogram

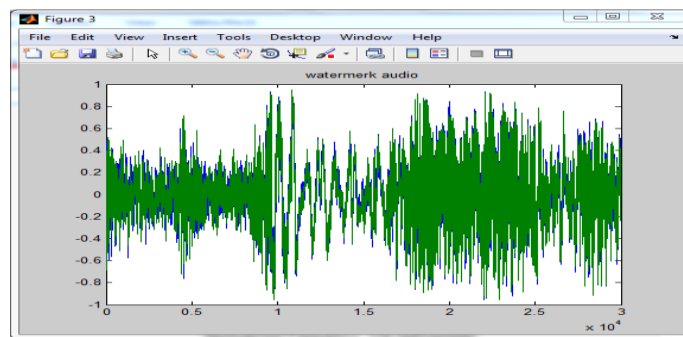


Figure 5. Watermark audio

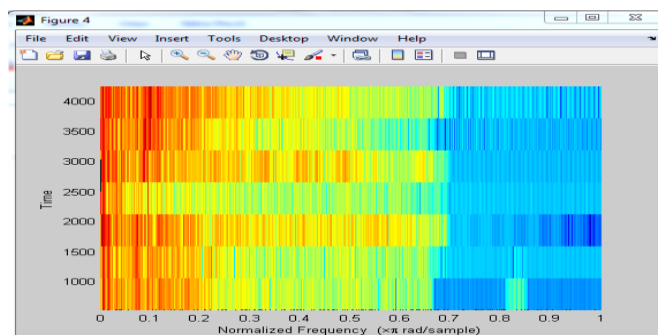


Figure 6. Watermark audio spectrogram

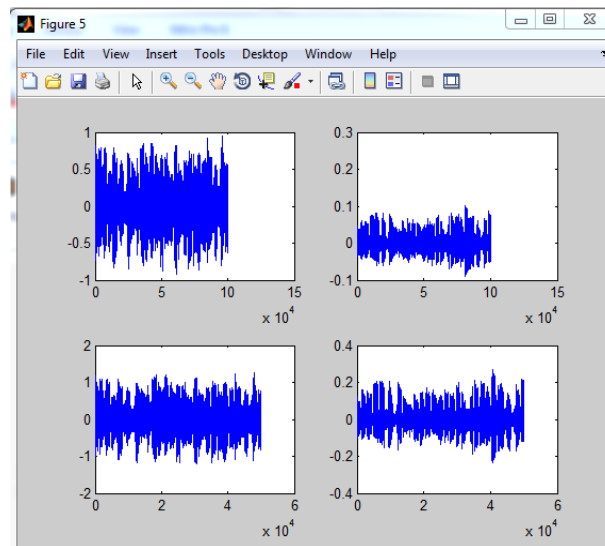


Figure 7. DWT decomposition of cover audio (a) L component (b) H component (c) LL component (d) LH component

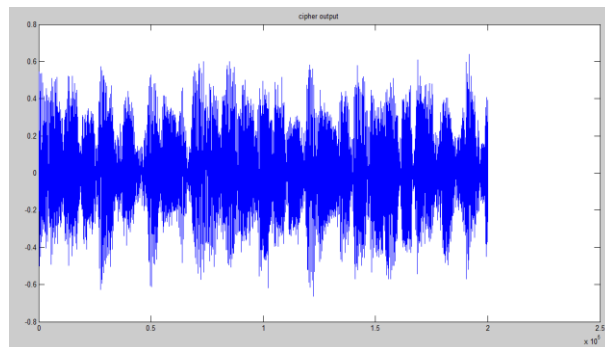


Figure 8. Cipher audio

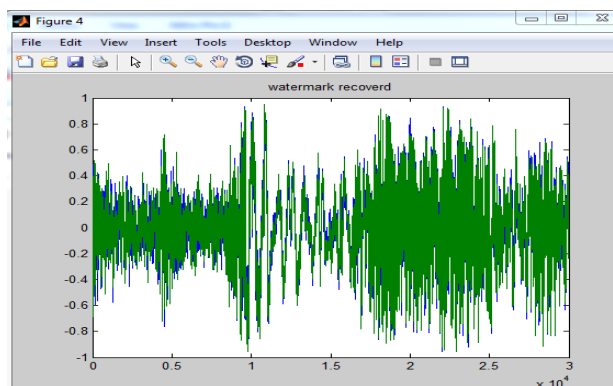


Figure 9. Extracted watermark audio

Figure 3 to figure 9 above shows the simulation graphs obtain while execution of the proposed work, it can be observe that the cover audio and the cipher audio waveforms plots are almost equals and the original watermark and extracted watermark audio are also equals.

Table 1 below is the results observe for the proposed work, the MSE, BER and SNR is been obtain for the different test conditions here the results developed for the MATLAB standard audio signals 'jazz', 'pop' and 'classic' and the length of watermark is of one second and the length of cover is 4 second. Above test scenario is the standard test scenario which is been use for the MATLAB based audio signal processing study.

The simulation results obtain at the MATLAB and the parameters for the proposed work are:

$$\text{MSE} = (D - C)^2 / \text{Len}$$

$$\text{SNR} = 10 \log_{10} (256^2 / \text{MSE})$$

$$\text{BER} = \text{sum of (Data } \textit{xor} \text{ Cipher)} / (\text{Len} \times 8)$$

Table 1. SNR, MSE and BER observation for different test cases

Cover size	Watermark Audio	SNR	MSE	BER
4-second, 784 kb, Jazz-2	1-second, 120 kb	43.3411	0.0185	0.0077
4-second, 784 kb, Pop-2	1-second, 120 kb	43.8548	0.0181	0.0075
4-second 784 kb Classic-2	1-second 120 kb	42.2158	0.0167	0.0082
22-second 3.81 Mb Jazz-1	2-second 392 kb	49.243	0.0094	0.0061
22-second 1.91 Mb Pop-1	2-second 392 kb	44.1846	0.0172	0.0069
22-second 7.63 Mb Classic-1	2-second 392 kb	50.5739	0.0135	0.0058

Table 2 below show the SNR observation for the proposed work for MATLAB standard jazz, pop and classic audio cover and 1sec watermark audio

Table 2. Comparative results: SNR

AUTHOR	Audio type	SNR
Neethu V et al [2]	Pop	34.11
	Jazz	37.79
	Classic	35.8
Masato Ogura et al [1]	Pop	30.0
	Jazz	30.4
	Classic	30.7
Proposed work	Pop	43.8548
	Jazz	43.3411
	Classic	42.2158

It can be clearly observe from the figure above that the proposed work SNR value is higher than available work ant all test conditions. The average BER observe by Masato Ogura et al is 0.6 and proposed work average BER is for the 0.0079 which again less in comparison.

CONCLUSION

The original objective of the thesis work was to develop an optimised technique for hiding audio inside cover audio also to reduce the amount of audio signal while data transmission, the objective are completed as was expected. The problem with watermarking is that it requires lots of data means another bigger audio for sending some small watermark audio, proposed work achieved that same size of watermark can be transmitted with small size of cover audio as achieved SNR is better than available work 43.524.

This paper also describes the problems associate with audio watermarking and problems in available work , observe from SNR value is higher than available work ant all test conditions. The average BER observe by Masato Ogura et al is 0.6 and proposed work average BER is for the 0.0079 which again less in comparison.

REFERENCES

- [1] Masato Ogura, Yosuke Sugiura and Tetsuya Shimamura, "SVD Based Audio Watermarking Using Angle-Quantization", IEEE International Conference on Electrical, Computer and Communication Engineering (ECCE), February 16-18, 2017.
- [2] Neethu V, R.Kalaivani, "Efficient and Robust Audio Watermarking for Content Authentication and Copyright Protection", 2016 IEEE International Conference on Circuit, Power and Computing Technologies [ICCPCT].
- [3] Ivana Milas, Bozidarka Radovic, Danilo Jankovic, "A New Audio Watermarking Method With Optimal Detection, 5th IEEE Mediterranean Conference on Embedded Computing ", MEeo 2016 Bar, Montenegro.
- [4] Michael Arnold, "AUDIO WATERMARKING: FEATURES, APPLICATIONS AND ALGORITHMS", 2000 IEEE International Conference on Multimedia and Expo, 2000. ICME 2000. DOI: 10.1109/ICME.2000.871531.
- [5] Kais Khaldi and Abdel-Ouahab Boudraa, "Audio Watermarking Via EMD", IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 21, NO. 3, MARCH 2013.
- [6] M A namatollhi, A /vorkupipat, C roslaes, "Digital watermarking techniques and trends", 2017 XXV, 203p, 42 illus, ISBN:978-981-10-2094-0, Springer .
- [7] Masoud Nosrati, Ronak Karimi, Mehdi Hariri, "Audio Watermarking: A Survey on Recent Approaches", World Applied Programming, Vol (2), No (3), March 2012. 202-205, ISSN: 2222-2510.
- [8] Harish Kumar, Anuradha, "Enhanced LSB technique for Audio Watermarking", IEEE-20180, ICCCNT'12, 26th_28th July 2012, Coimbatore.
- [9] NEDELJKO CVEJIC "ALGORITHMS FOR AUDIO WATERMARKING AND WATERMARKING", Kuusamonsali (Auditorium YB210), Linnanmaa, on June 29th, 2004, ISBN 951-42-7383-4.

- [10] Arnold, M. 2000. "Audio watermarking: features, applications, and algorithms". In IEEE International Conference on Multimedia and Expo (II). Citeseer.
- [11] Yeo, I.-K., and H.J. Kim. 2003. "Modified patchwork algorithm: A novel audio watermarking scheme". IEEE Transactions on Speech and Audio Processing 11(4): 381–386.
- [12] Cox, I.J., et al. 1997. Secure spread spectrum watermarking for multimedia. IEEE Transactions on Image Processing 6(12): 1673–1687.
- [13] Xu, C., et al. 1999. "Applications of digital watermarking technology in audio signals". Journal of the Audio Engineering Society 47(10): 805–812.
- [14] Oh, H.O., et al. 2001. "New echo embedding technique for robust and imperceptible audio watermarking". In 2001 IEEE international conference on acoustics, speech, and signal processing, 2001. Proceedings. (ICASSP'01). IEEE.
- [15] Kim, H.J., and Y.H. Choi. 2003. "A novel echo-hiding scheme with backward and forward kernels". IEEE Transactions on Circuits and Systems for Video Technology 13(8): 885–889.
- [16] Ko, B.-S., R. Nishimura, and Y. Suzuki. 2005. Time-spread echo method for digital audio watermarking. IEEE Transactions on Multimedia 7(2): 212–221.
- [17] Nian Cai, Nannan Zhu, Shaowei Weng, Bingo Wing-Kuen Ling, "Difference angle quantization index modulation scheme for image watermarking, Signal Processing". Image Communication 34 (2015) 52–60, 0923-5965/& 2015 Elsevier.