Audio Watermarking using Empirical Mode Decomposition

Charulata P. Talele[1], Dr A. M. Patil[2]

[1]ME Student, Electronics and Telecommunication Department
J. T. Mahajan College of Engineering, Faizpur, Maharashtra, India 425524.
[2]Dr. A. M. Patil, Electronics and Telecommunication Department
J. T. Mahajan College of Engineering, Faizpur, Maharashtra, India 425524.

---***---

Abstract - *In this paper a new adaptive audio watermarking algorithm based on Empirical Mode Decomposition (EMD) is introduced. The audio signal is divided into frames and each one is decomposed adaptively by EMD, into intrinsic oscillatory components called Intrinsic Mode Functions (IMFs). The watermark and the synchronization codes are embedded into the extrema of the last IMF, a low frequency mode stable under different attacks and preserving audio perceptual quality of the host signal. We show the robustness of the hidden watermark for different attacks such as Gaussian Attack, filtering Attack, Cropping Attack, Re-sampling Attack and Pink noise Attack. The comparison analysis shows that our method has better performance than watermarking schemes reported recently.*

Key Words: **Empirical Mode Decomposition (EMD), Intrinsic Mode Functions (IMFs), Discrete cosine transform (DCT), Discrete Wavelet Transform (DWT)…**

1. INTRODUCTION

Watermarking is the method of hiding information bits into a host signal. The information bits or data can be used for identifying the owner, for copyright protection, authentication, copy control etc. The host signal can be audio, video or image. The copyright protection of digital media is achieved by embedding a watermark in the original host signal. If the host signal is an audio signal then the watermarking is audio watermarking. The requirements of digital audio watermarking are imperceptibility, robustness and security. Imperceptibility means the watermark in the audio signal is inaudible to the listener. Robustness refers to the changes in the watermark image under various conditions. Security means the watermark cannot be extracted or deleted without the permission of the owner. There are different methods for audio watermarking. The methods are broadly classified into temporal and spectral watermarking. In temporal watermarking the watermarked data are embedded directly into the host audio signal in the time domain. In spectral watermarking certain frequency transforms are performed to the host audio signal and then embed the watermark info into the transformed frequency domain data block. In temporal watermarking many techniques are adopted such as watermarking in the dual channel audio using echo hiding scheme [1], echo hiding using analysis by synthesis method [2], using time scale modification method [3], using EMD (Empirical Mode Decomposition) method [4]. In spectral watermarking many techniques are used which involves the transformation of the host signal such as DWT [5], DCT [6]. Both temporal and spectral techniques can be combined together for better results as in [7] which use both DWT and SVD for audio watermarking. For achieving a good tradeoff between reliability and robustness while achieving better payload a new method using the decomposition property EMD and compression property of DCT is proposed. The limit of DWT, DCT etc. are they use fixed basis functions which do not match all real signals. So to overcome this problem we use a new decomposition method called as Empirical Mode Decomposition (EMD) which decomposes the signal adaptively according to the value of signals. So there is no apriority basis function for the EMD and is completely data driven. EMD recursively breaks down signals into nearly zero mean symmetric envelops called as Intrinsic Mode Functions (IMFs). The EMD decomposition can be expressed as

$$X_t = \sum_{n=1}^{\infty} IMF_j(t) + \mathrm{rN(t)}$$

Where N is the number of IMFs and rN(t) is the final residual. The IMFs obtained are nearly orthogonal to each other. The EMD is performed in a finite number of modes and the number of extrema decreases when we move from one mode to other. The watermark is embedded in the extrema of the last IMF as in [4]. In the proposed method a color image is used as the watermark. The color image is decomposed into its RGB components and then DCT transformation [8] is applied. By exploiting the energy compaction property of the DCT [9] the image is compressed. Then the compressed image is converted into binary bits and is watermarked into the audio signal. The reverse process is performed to obtain the watermarked color image from the watermarked audio signal.

2. PROPOSED WATERMARKING ALGORITHM

The proposed method uses a binary data as a watermark to embed in an audio signal. The binary data is obtained from a colored image First the audio signal is divided into frames and EMD [10] is calculated on each and every frame. The last IMF obtained from the EMD and the watermark bits are embedded in it. The extrema of the last IMF is calculated and the watermark bits are embedded using Quantization Index Modulation (QIM) [11]. So a bit is inserted at each extrema of the last IMF. Then the inverse EMD is performed and the frames are obtained. Finally the frames are concatenated to obtain the final watermark embedded signal. The binary bits which are to be embedded are obtained from a compressed color image. The image is compressed using the energy compaction property of the DCT. A suitable compression method is chosen so that it gives a good PSNR value. Then it is converted into binary format and the bits for watermark are obtained.

For extracting the watermarked data from the audio the watermarked audio signal is subtracted from the original audio and the QIM coefficients are extracted. The binary bits are demodulated from the QIM coefficients. The binary bits are then converted into the DCT coefficients and inverse DCT transformation is taken. Now the colored image is obtained. Overview of the proposed audio watermarking technique is detailed as follows:

2.1 Watermark Preparation

Before embedding into the audio signal the watermark data should be converted into a binary sequence. Here the watermark is an image, so we have to convert the image into a binary sequence detailed as follows:

- Decompose the colored image into RGB components.
- Each component image is taken separately and is divided into 8x8 blocks of sub images.
- DCT is calculated on each sub block.
- The DCT coefficients are quantized using a quantization matrix. Quantization is achieved by dividing each element of DCT transformed matrix by the element in the quantization matrix and rounding it to nearest integer.
- After quantization most of the coefficients become zero. Using the energy compaction property of the DCT the most relevant coefficients are selected by a zigzag manner. According to the compression ratio needed we select the number of coefficients and discard the other ones.
- The selected coefficients are converted into binary bits to obtain a binary sequence.

2.2 Watermark Embedding

The binary sequence obtained from the watermark preparation phase used for embedding in the host audio signal. Basics of watermark embedding are as shown in Figure 1 and can be detailed as:

- Divide the original host audio signal into frames.
- Decompose each frame into several IMFs using EMD.
- Embed the binary sequence into the extrema of the last IMF by QIM

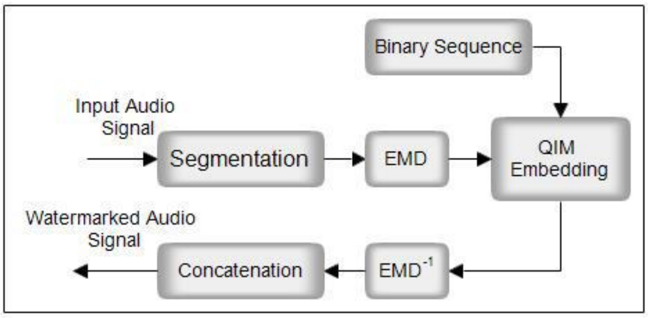

Fig-1: Watermark Embedding

where e_i^* is the modified extrema of the last IMF and eiis the extrema of last IMF of host signal. sgn function is equal to '+' if e_i is a maxima and '-' if ei is a minima. The embedding strength is denoted as S and the value of S is chosen so that the watermark is inaudible.

- Perform the inverse EMD (EMD) on the IMFs with modified last IMF to obtain the frames.
- Concatenate the watermarked frames to obtain the watermarked audio signal.

C. Watermark Extraction and Reconstruction

The binary sequence is extracted from the watermarked audio signal. As the watermark is embedded as QIM coefficients in the audio, we have to extract the QIM coefficients and decode it to binary bits. Basic steps involved in watermark extraction are shown in Figure 2 and detailed as Subtract the watermarked audio signal from the original audio signal to obtain the QIM coefficients. The binary bits are obtained from the QIM coefficients and Linearize the binary bits to form a binary sequence which is the watermark. The colour image watermark is recovered from the RGB components Basics of watermark reconstruction is shown in Figure 3 and detailed as follows:

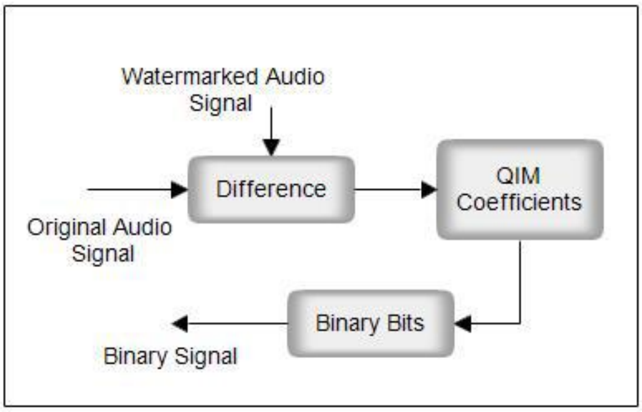

Fig-2: Watermark Extraction

- The binary sequence is first divided into three corresponding R G B components.

- For each component the binary sequence is converted back into its decimal form.
- The coefficients obtained are arranged in an inverse zigzag form, after padded with zeros corresponding to the compression ratio, to obtain 8x8 transformed sub images.
- To obtain the DCT coefficients multiply each 8x8 transformed sub image with the corresponding element in the quantization matrix.
- Inverse DCT transformation is applied and sub images are obtained.
- The sub images are joined to get the R, G, B components and further they are fused together to obtain the final watermarked image.

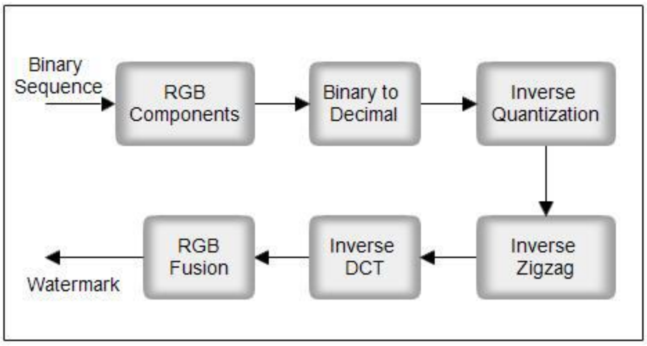

Fig-3: Watermark Reconstruction

3. RESULTS AND PERFORMANCE ANALYSIS

Simulations are performed on different 10 audio files in "WAV" format namely crabby.wav, dadada.wav, duck.wav, formal.wav, helloo.wav, lion2.wav, powers.wav, silyvoice.wav, spock.wav spooky.wav and trek.wav with variable size in Kilo bits, length in seconds and same bit rate 64 Kbps from internet source [2][3]. A binary logo image of Hewlett-Packard (hp) size M_N = 792_792 = 1296 bits (here M=N) with size 52.5 Kb watermark is embedded. Watermark is shown in figure 4. The figure 5, figure 6 and figure 7 shows simulation results of scraby.wav audio file embedding with hp logo without attack and with different attacks namely Gaussian Attack, filtering Attack, Cropping Attack, Resampling Attack and Pink noise Attack. The difference between the original signal and the signal after watermarking shown in figure 5. The difference is not much and the distortion is not audible to the human ear. The performances of any audio watermarking technique should be analyzed based on the imperceptibility of the audio watermark on the capability of the watermark to resist against above mentioned attacks and on the data payload. In this sense, we will measure the performance against attacks by using Bit Error Rate (BER) given for scraby.wav in table 1 [16]. A low BER indicates an accurate detection mechanism for the watermark. BER can be calculated with the help of equation 3.

Fig-4: logo(hp)

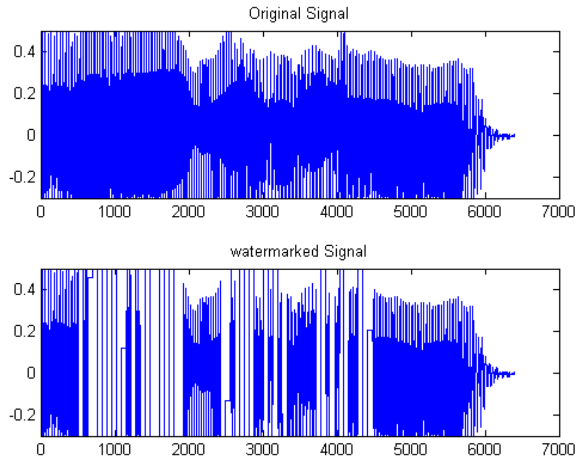

Fig-5: Audio scraby.Wav signal before embedding and after watermark embedding.

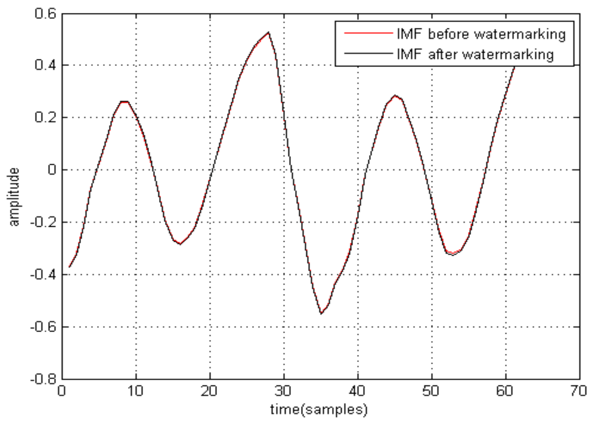

Fig-6: IMF before (red) and after (black) for scraby.wav

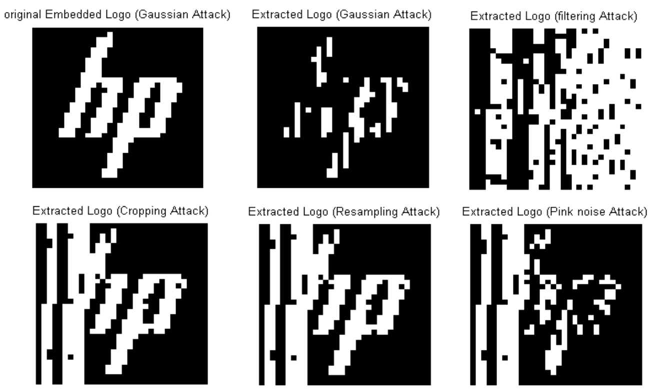

Fig-7: Rsults of extracted hp logo for different attacks

Where is the XOR operator and M N are the binary watermark image sizes. W(i;j) and W0(i;j) are the original and the recovered watermark respectively. is used to evaluate the watermark detection accuracy after signal processing operations.

Table -1: BER results of reconstructed logo for different attacks.

Type of Attack	BER
Normal	0.29297
Gaussian Attack	10.6445
filtering Attack	54.3945
Cropping Attack	15.918
Resampling Attack	15.918
Pink noise Attack	23.0469

Figure 8 shows bar chart representation of BER values for scrbby.wav, dadada.wav, duck.wav, formal.wav and helloo.wav audios as per bar chart in figure 8 and figure 9 it is clear that proposed method is not best suitable for filtering attack because result for filtering attack shows highest BER up to 54.3945 and mean BER of 33.6292 that means 33% of embedded information in audio file is lost during processing in other hand BER result for Gaussian Attack shows good result with mean BER 11.8696 it means proposed method works better on Gaussian attack.

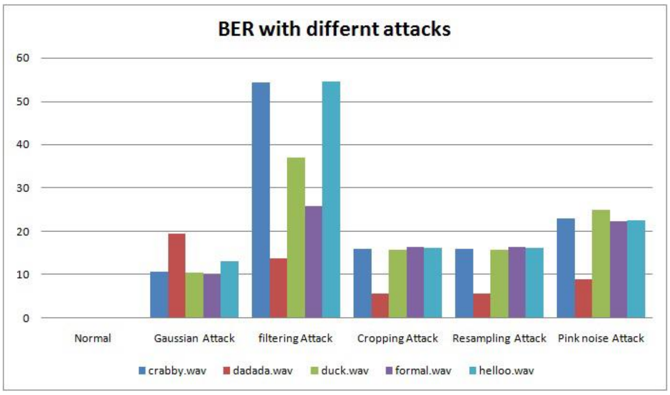

Fig- 8: BER values for 5 different audio samples

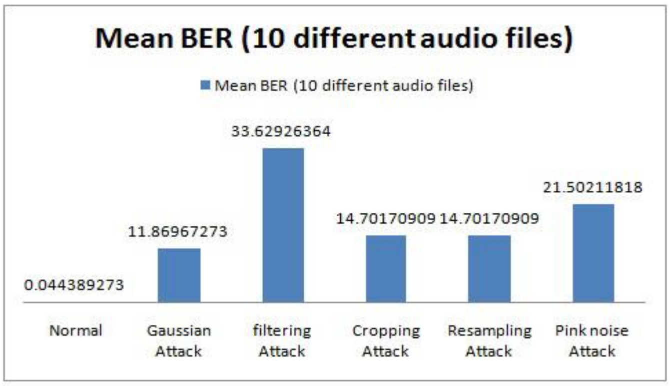

Fig-9: Mean BER values for 10 different audio samples with hp logo.

3. CONCLUSIONS

In this paper, we have proposed a robust audio watermarking algorithm based on empirical mode decomposition. In our algorithm, the intrinsic feature of final residual is selected to embed watermark thus the imperceptibility of watermarking is improved. The experimental results show that the algorithm is robust against 5 kinds of common attacks such as Gaussian Attack, filtering Attack, Cropping Attack, Resampling Attack and Pink noise Attack. Result shows proposed method is not best suitable for filtering attack because result for filtering mean BER of 33.6292 that means 33% of embedded information in audio file is lost during processing in other hand BER result for Gaussian Attack shows good result with mean BER 11.8696 it means proposed method works better on Gaussian attack.

ACKNOWLEDGEMENT

We are thankful to the entire faculty and staff of Electronics and Telecommunication Engineering Department for the help and moral support which went along the way for the successful completion of this research work, our greatest thanks is to all who wished us success especially our parents and family members.

REFERENCES

[1] C. P. Talele, Dr A. M. Patil, "Audio Watermarking via EMD "Review", International Journal of Engineering Science and Computing,Volume 6, Issue No. 8, Pages 2658-2661, August 2016.

[2] http://www.sillyhumor.com/answer/classicanswerin gmachinemessages.html

[3] http://www.externalharddrive.com/waves/animal/in dex.html

[4] I. J. Cox and M. L. Miller, "The first 50 years of electronic watermarking," J. Appl. Signal Process., vol. 2, pp. 126–132, 2002.

[5] M. D. Swanson, B. Zhu, and A. H. Tewfik, "Robust audio watermarking using perceptual masking," Signal Process., vol. 66, no. 3, pp. 337–355, 1998.

[6] S. Wu, J. Huang, D. Huang, and Y. Q. Shi, "Efficiently self-synchronized audio watermarking for assured audio data transmission," IEEE Trans. Broadcasting, vol. 51, no. 1, pp. 69–76, Mar. 2005.

[7] V. Bhat, K. I. Sengupta, and A. Das, "An adaptive audio watermarking based on the singular value decomposition in the wavelet domain," Digital Signal Process., vol. 2010, no. 20, pp. 1547–1558, 2010.

[8] D. Kiroveski and S. Malvar, "Robust spread-spectrum audio watermarking," in Proc. ICASSP, 2001, pp. 1345–1348.

[9] K. Khaldi, M. T.-H. Alouane, and A. O. Boudraa, "Voiced speech enhancement based on adaptive filtering of selected intrinsic mode functions," J. Adv. in Adapt. Data Anal., vol. 2, no. 1, pp. 65–80, 2010.

[10] L. Wang, S. Emmanuel, and M. S. Kankanhalli, "EMD and psychoacoustic model based watermarking for audio," in Proc. IEEE ICME, 2010, pp. 1427–1432.

[11] A. N. K. Zaman, K. M. I. Khalilullah, Md. W. Islam, and Md. K. I. Molla, "A robust digital audio watermarking algorithm using empirical mode decomposition," in Proc. IEEE CCECE, 2010, pp. 1–4.

[12] I. J. Cox, J. Kilian, T. Leighton, and T. Shamoon, "A secure, robust watermark for multimedia," LNCS, vol. 1174, pp. 185-206, 1996.

[13] B. Chen and G. W. Wornell, "Quantization index modulation methods for digital watermarking and information embedding of multimedia," J. VLSI Signal Process. Syst., vol. 27, pp. 7–33, 2001.

[14] W. N. Lie and L.-C. Chang, "Robust and high quality time-domain audio watermarking based on low frequency amplitude modification," IEEE Trans. Multimedia, vol. 8, no. 1, pp. 46–59, Feb. 2006.

[15] I. K. Yeo and H. J. Kim, "Modified patchwork algorithm: A novel audio watermarking scheme," IEEE Trans. Speech Audio Process., vol. 11, no. 4, pp. 381–386, Jul. 2003.

[16] R. Tachibana, S. Shimizu, S. Kobayashi, and T. Nakamura, "An audio watermarking method using a two-dimensional pseudo-random array," Signal Process., vol. 82, no. 10, pp. 1455–1469, 2002.

[17] N. Cvejic and T. Seppanen, "Spread spectrum audio watermarking using frequency hopping and attack characterization," Signal Process., vol. 84, no. 1, pp. 207–213, 2004.

[18] W. Li, X. Xue, and P. Lu, "Localised audio watermarking technique robust against time-scale modification," IEEE Trans. Multimedia, vol. 8, no. 1, pp. 60–69, 2006.

[19] M. F. Mansour and A. H. Tewfik, "Data embedding in audio using timescale modification," IEEE Trans. Speech Audio Process., vol. 13, no. 3, pp. 432–440, May 2005.

[20] Ghobadi and Alireza, "Blind audio watermarking for tamper detection based on LSB." Advanced Communication Technology (ICACT), 2013 15th International Conference on. IEEE, 2013

[21] K. Khaldi and A. Boudraa "Audio Watermarking Via EMD" IEEE Transactions On Audio, Speech, And Language Processing, Vol. 21, No. 3, March 2013