ISSN: 2278-4632
Vol-09 Issue-9 No. 1 September 2019

EMPIRICAL MODE DECOMPOSITION FOR AUDIO WATERMARKING

Dr. THODUPUNURI SRINIVAS, ***Professor,***

Department of Electronics and Communications Engineering,

MOTHER THERESA COLLEGE OF ENGINEERING AND TECHNOLOGY, PEDDAPALLY, TS.

ABSTRACT: Empirical Mode Decomposition (EMD) is utilized in this work to present a new technique for adding adaptive watermarks to audio recordings. EMD fluidly separates the sound source into oscillating IMFs (Inherent Mode Functions) from each frame. The watermark and time codes are added at the end of the Intrinsic Mode Function (IMF). Multiple assaults have little influence on the quality of the IMF's original signal because of its extremely low frequency. Pink noise, resampling, cropping, filtering, and the Gaussian attack have no effect on the concealed marking. Our watermarking approach outperformed other recently disclosed strategies.

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

1. INTRODUCTION

Watermarking is a technique for masking the addition of data to a primary transmission. The information can be used for a variety of purposes, including verifying ownership, preventing unauthorized use, validating user names, and restricting the number of copies that can be made. The host content is available in the form of still photographs, videos, or audio files. For copyright reasons, the host data in digital video is watermarked. If the host signal is an audio transmission, the watermarking signal will be as well. Digital music watermarking must be safe, difficult to detect, and long-lasting in order to be effective. If the watermark is not audible in the audio transmission, the recipient cannot decode it. The image of a watermark's endurance can be measured by how little it changes when exposed to various conditions. A trademark is considered inviolable if it cannot be changed or withdrawn without the owner's permission. There are numerous methods for "watermarking" audio recordings. Methods are often classified into one of the following categories: The time-domain audio stream is watermarked with both temporal and spectral temporal watermarking in order to record data. Specific frequencies of the original audio stream are modified in spectral watermarking, and watermark data is attached to the modified data block in the frequency domain. Echo hiding, time scale change, empirical mode decoding (EMD), and dual-channel audio watermarking with echo hiding are all examples of temporal watermarking techniques. During spectral watermarking, the host signal goes through many transformations, including the Discrete Wavelet Transform (DWT) and the Discrete Cosine Transform (DCT). Discrete Wavelet Transform (DWT) and Singular Value Decomposition (SVD) are used in audio watermarking to improve outcomes by combining spectral and temporal approaches. It keeps its dependability and durability while improving carrying capacity. It is proposed to combine in an innovative way DCT's information-reduction skills and EMD's division capabilities. The basic functions utilized by DWT, DCT, and other comparable algorithms do not precisely match to all conceivable real-world signals. MD, or empirical mode decomposition, is a good solution for this problem because it decomposes the signal based on its value. This suggests that the EMD lacks a tool for prioritizing jobs based solely on data. The empirical mode decomposition (EMD) method produces intrinsic mode functions (IMFs), which are equal bands with near-zero mean values. The EMD decomposition is described as follows: The number of intrinsic mode functions (IMFs) is N, and the identical final residue component is rN(t). The learned intrinsic mode functions (IMFs) do not always coincide. When switching modes, there are a limited amount of EMD modes and extremes to choose from. The most current IMF report, as viewed, has some discolouration. The stamp in the suggested technique is a full-color image. In preparation for the Discrete Cosine Transform (DCT), the image is first converted to a series of Red, Green, and Blue (RGB) values. For this, the energy compression function of the Discrete Cosine Transform (DCT) is used. Before being incorporated to the audio recording, the image is transformed to digital data. By rearranging the stages, you can transform a watermarked audio stream into a watermarked color image.

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

The ultimate residue is represented by rN(t), while the number of IMFs is represented by N. There is significant overlap among the discovered IMFs. The EMD has a limited number of modes, and the number of extremes reduces as we cycle through them. The IMF's final two extremes feature a marking. The stamp, as expected, is a color photograph. The DCT transformation is applied after dividing a picture into RGB components. The image is reduced using the DCT's energy compaction device. Following that, the image is compressed, watermarked in the audio signal, and transformed to binary data. To obtain the watermarked color image from the watermarked audio signal, the processes are executed in reverse order.

2. PROPOSED WATER MARKING ALGORITHM

The proposed method entails adding a binary data signature to an audio broadcast. The conversion of a picture to binary code. First, discrete audio data segments known as "frames" must be constructed. The Empirical Mode Decomposition (EMD) for each image is then determined. The watermark bits are saved while EMD is building a fresh IMF. After locating the last two ends of the IMF, Quantization Index Modulation (QIM) is employed to add the watermark bits to the data. As a result, a new component has been added to both sides of the final IMF. A reversed EMD is used to order the frames.

The final output is made up of individual frames including the watermark signal. A compressed color image is used to extract the required binary bits. We minimize the size of the image by using the energy compaction function of the Discrete Cosine Transform (DCT). A high Peak Signal-to-Noise Ratio (PSNR) is achieved by selecting the best compression method. After that, the data is transformed to binary and the signature is deleted. The QIM coefficients are obtained after separating the watermarked audio signal from the source audio. This is the location where you can get watermarked music. Using QIM factors to combine the missing binary digits. The technique is repeated after the binary data are transformed to DCT coefficients. The shot I have is in full color. The following is a detailed description of the suggested audio watermarking method:

Watermark Preparation

The sequence must be translated into binary before watermark data can be applied to an audio stream. The appearance of the logo is depicted in this binary picture.

The image's red, green, and blue components must be removed.

8x8 subimages are formed by merging pieces of the parent image.

For each subblock, the Discrete Cosine Transform (DCT) is computed.

To quantify the DCT numbers, quantization matrices are used. You can round the fraction to the nearest integer by dividing each cell of the DCT-converted matrix by its corresponding cell in the quantization matrix.

When factors are quantized, they are almost worthless. The coefficients are organized in a zigzag pattern using the DCT's energy compaction function. The desired compression ratio determines the number of factors, and any surplus factors are deleted.

The numerals are used to form a sequence after being converted to binary.

Watermark Embedding

Using the binary code generated during the watermark's processing, the watermark is added to the host's audio stream. The fundamental principles underlying watermark embedding are depicted in Figure 1. The parts that follow go into deeper detail on these themes.

The presenter's principal audio stream must be separated.

Multiple IMFs can be produced from a single image using empirical mode decomposition (EMD).

The binary series is appended at the highest and lowest points to the final Intrinsic Mode Function (IMF) using the Quadratic Interpolation Method (QIM).

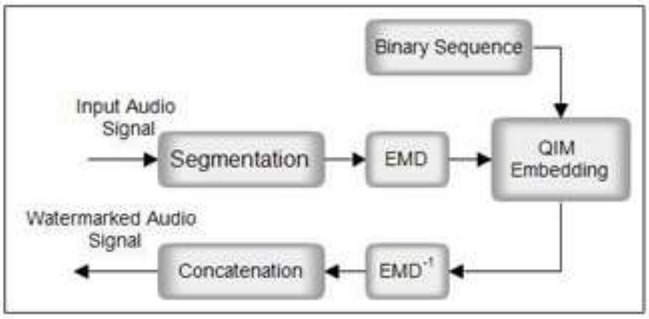

Fig-1: Watermark Embedding

- An The preceding IMF extremum variable ei represents the host signal. The variable e* denotes an altered IMF extremum. When ei is the largest number, a plus sign (+) is used, and when it is not, a minus sign (-) is used. S selects an embedding intensity number that eliminates human hearing of the watermark. To obtain the frames, perform an inverse Empirical Mode Decomposition (EMD) on the most recent IMF.
- Combine the assembled watermarked frames to obtain the watermarked audio stream.

Watermark Extraction and Reconstruction

Using the watermarked audio output, the binary code is discovered. We must first convert the QIM values extracted from the audio to binary before we can retrieve the watermark. Figure 2 depicts and describes the primary steps involved in stamp removal. QIM coefficients can only be extracted from an unwatermarked primary audio stream. The stamp binary sequence is created by aligning and joining the binary integers of the QIM coefficients. In order to create the watermark, we divide the color image into its RGB components. Figure 3 depicts and discusses the fundamentals of watermark restoration.

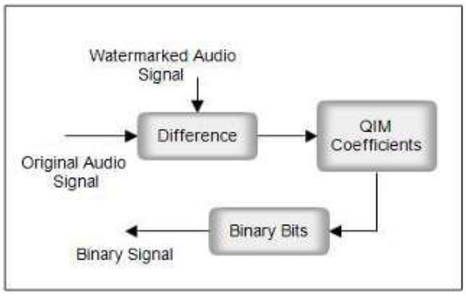

Fig-2: Watermark Extraction

- Removing the watermark
- The binary code is made up of three distinct parts: red (R), green (G), and blue (B).
- The format of each component is converted from binary to decimal.
- The coefficients are assembled in an inverted zigzag pattern and then zeros are added to them based on the compression ratio to create 8x8 transformed sb images.
- To obtain the DCT coefficients, multiply the quantization matrix element by each of the modified 8x8 subimages.
- Subimages can be generated using the Inverse Discrete Cosine Transform (DCT) algorithm.
- When the subimages are combined, the red (R), green (G), and blue (B) components of the watermarked image are obtained.

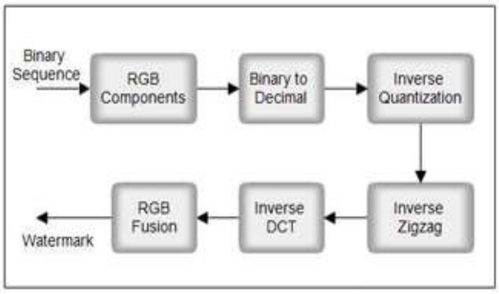

Fig-3: creating watermarks

3.RESULTS AND PERFORMANCE ANALYSIS

The test employs ten different "WAV" audio files. The collection contains files such as "crabby.wav," "dadada.wav," "duck.wav," "formal.wav," "helloo.wav," "powers.wav," "sillyvoice.wav," "spock.wav," "spooky.wav," and "trek.wav," among others. These files' lengths are indicated in seconds, and their sizes are expressed in kilobits.

Each has a throughput of 64 Kbps. Every portion

ISSN: 2278-4632
Vol-09 Issue-9 No. 1 September 2019

of an internet video is available in digital form. On the binary image of the Hewlett-Packard (hp) logo, a watermark reads M_N = 792792 pixels, which is 1296 bits. The file has a size of 52.5 kilobytes. Figure 4 shows a seal. After the HP logo was inserted, the scraby.wav audio file was subjected to many attacks, including Gaussian, Filtering, Cropping, Resampling, and Pink Noise. The results are shown in the graph below. The transformation of non-watermarked signals into watermarked signals is depicted in Figure 1.

Because it is so far away, this alteration cannot be heard.

The utility of an audio watermarking system can be determined by criteria such as the amount of data it can hold, how well it protects against known breaches, and how difficult it is to discover. The Bit Error Rate (BER) will be calculated to determine the security of the scraby.wav file in Table 1. Watermark detection systems that are effective have a low Bit Error Rate (BER). To find the BER, use the third method.

Fig-4: HP is a corporate name.

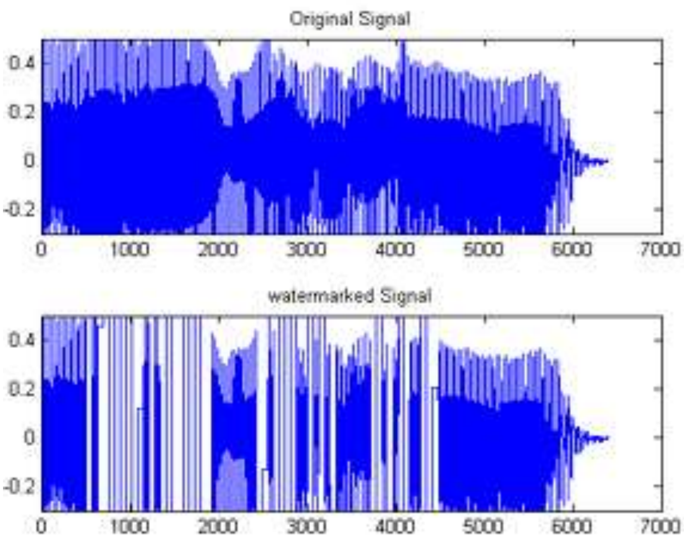

Fig-5: Dress casually.Both before and after the imprint, a Wav signal was recorded.

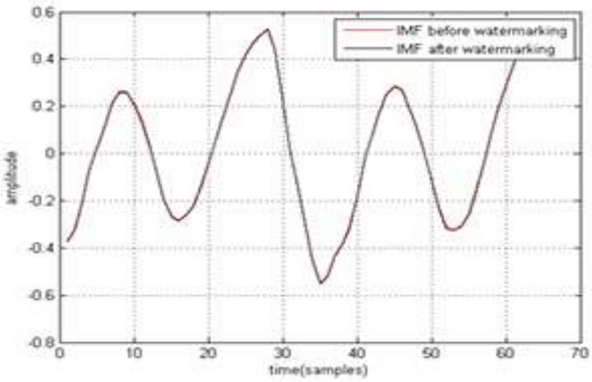

Fig-6: It depicts the IMF before (red) and after (black) scraby.wav.

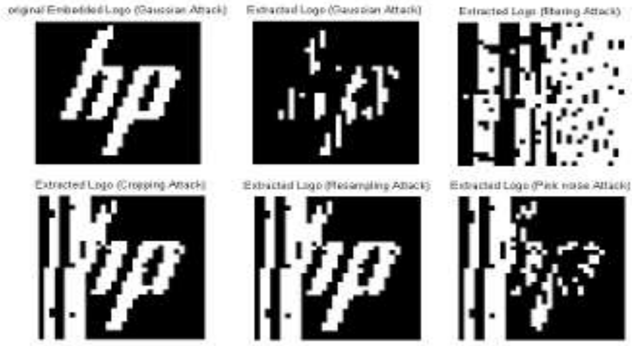

Fig-7: HP logo extraction examples for various assaults

XOR is employed in this case to incorporate binary watermark pictures of M and N bytes. Watermarks appear as W(i;j) before healing and as W0(i;j) after healing. The goal of this test is to see how effectively watermark recognition follows signal processing actions.

Table -1: The BER values of the rebuilt logo for each attack are shown in Table 1

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

The BER values for the wave files scrbby.wav, dadada.wav, duck.wav, formal.wav, and helloo.wav are shown in Figure 8. Figures 8 and 9 show that the proposed assault prevention approach is ineffective. This is obvious considering that the BER values for filtering assaults range from 54.3945 to 33.6292. As a result, roughly one-third of the entered data is lost.

ISSN: 2278-4632
Vol-09 Issue-9 No. 1 September 2019

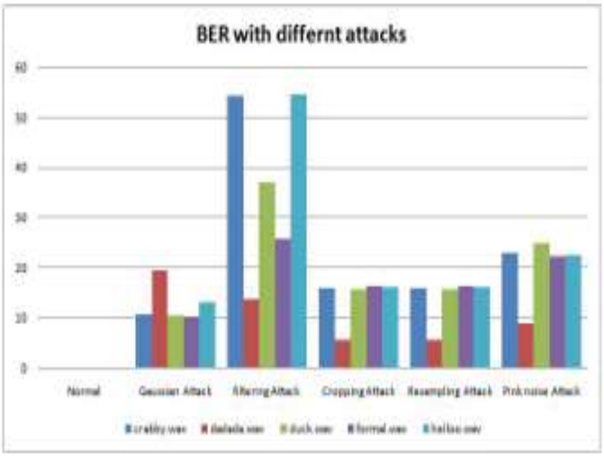

Fig- 8: BER for five different audio samples

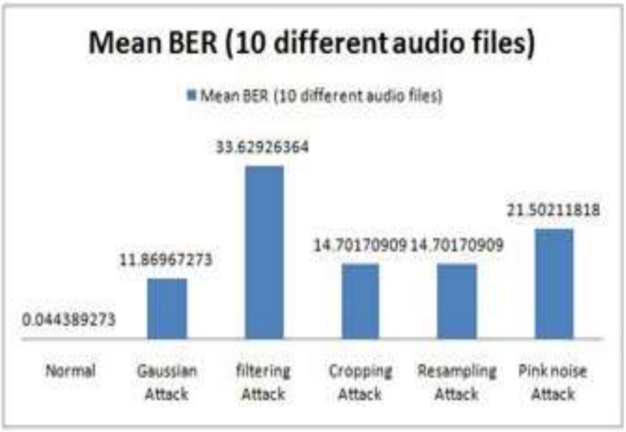

Fig-9: The BER for 10 separate audio samples containing the HP logo was averaged.

4. CONCLUSIONS

Empirical mode decomposition is used in this study to explain how to properly categorize audio recordings. Our solution employs a watermark that conceals the watermarking procedure by employing an end residue characteristic. The approach has been tested and found to be immune to the five most prevalent types of attacks: resampling, cropping, filtering, Gaussian noise, and pink noise. Because 33.0% of the original audio data was lost during conversion, the Bit Error Rate (BER) of 33.6292 suggests that the proposed technique for preventing attacks is ineffective. The filtering results indicate this. The average BER of a Gaussian attack is 11.8696, which is fairly good. Based on this information, the proposed technique has a better chance of success than the Gaussian attack.

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/classicans werin gmachinemessages.html
3. http://www.externalharddrive.com/waves/anim al/in dex.html
4. 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.