# EFFICIENT APPROACH FOR DIGITAL AUDIO WATERMARKING SCHEME

**Rakshitha. N**
Student
Dept. of ECE
ATMECE
Mysuru, Karnataka

**Niveditha G.B.**
Student
Dept. of ECE
ATMECE
Mysuru, Karnataka

**Niveditha B.S.**
Student
Dept. of ECE
ATMECE
Mysuru, Karnataka

**Girish M**
Asst. Professor
Dept. of ECE
ATME college of Engineering
Mysuru, Karnataka.

*Abstract:* **Digital watermarking technology is interested in solving the problem of copyright protection, data authentication, content identification, distribution, and duplication of the digital media due to the great developments in computers and Internet technology. Recently, protection of digital audio signals has taken a great attention from the researchers. This paper proposed a new audio watermarking scheme based on Discrete Wavelet Transform (DWT), Singular Value Decomposition (SVD), and Quantization Index Modulation (QIM) with synchronization code embedded within double encrypted watermark images or logos into stereo audio signal. In the algorithm, an original audio signal is split into blocks and each block is decomposed into two levels discrete wavelet transform, and then the approximate low frequency sub-band coefficients are decomposed by the SVD transform, obtaining a diagonal matrix. The prepared watermarking and synchronization code bit stream is embedded into the diagonal matrix using Quantization Index Modulation (QIM). After that, we apply ISVD and IDWT to obtain the watermarked audio signal. The watermark can be blindly extracted without knowledge of the original audio signal. Experimental results show that the transparency and imperceptibility of the proposed algorithm is satisfactory, and robustness is strong against popular audio signal processing attacks. High watermark payload is achieved and performance analysis is presented.**

## I. INTRODUCTION

With the fastest development of Multimedia frameworks and Computer networks, there is a requirement of copyright and proprietorship of computerized digital media. In the past two decades, the usage of internet has been increased drastically by introducing digitization. It is found that in 1993, internet used to carry only 1% of information; by 2000, data was increased to 51% and by the year 2007, more than 97% of information was carried by the internet. The digital data may be text, image, audio and video that are carried over open networks and hence there is a need to protect the data. It has become very easy task for unapproved clients to directly access, effortlessly control, manipulate and offer the digitized data. Due to the illicit duplicating and absence of security, numerous distributors are unwilling to distribute their work on web. At the point when a delicate or secret must be conveyed to the destination, privacy ought to be given. To take care of these issues, couples of strategies were developed – cryptography and watermarking.

Watermarking can provide protection to the digital content during its normal usage since the copyright information is hidden within the digital content in such a way that it cannot be removed. This unique feature makes watermarking one of the most promising techniques for digital content protection. A simple example of watermarking is specified patterns which are present in currencies that are visible only when it is held to light. The specialty of the watermark is that it is always present in the host. Such watermark can be used for various applications like indexing, broadcast monitoring, data authentication etc. A lot of work has been carried out on digital watermarking on various media like audio, video and image however this project deals with audio watermarking.

## II. LITERATURE SURVEY

In audio watermarking different schemes are proposed both in time and frequency domain. Some of them are described briefly as follows:

### A. *Robust watermarking using perceptual masking*

In this technique, the watermark information is placed in the original digital audio by directly varying the audio samples and the scheme also make use of temporal and frequency perceptual masking. The masking will guarantee the inaudibility and robustness of watermark. The watermark is constructed by splitting the audio signal into smaller segments and then adding the pseudo random sequence that is shaped by the phenomena of masking. The watermark may consist of owner's personal details and spectral and temporal masking using the masking effects of Human Audio Systems. Thus the noise like watermark is difficult to detect and is inaudible [2].

### B. *Robust spread spectrum audio watermarking Scheme*

One of the most promising watermarking schemes is to embed the spread-spectrum (SS) sequence (SS sequence is taken as watermark here) in the high amplitude regions of a signal. In this technique, the robustness and detection accuracy is increased by introducing components like synchronization search, chip redundancy, cepstrum filtering and chess watermark to the SS watermarking. Let '$x$' be the original audio signal vector that need to be watermarked, then the watermarked vector is obtained by $y = x + w$, where $w$ is the watermark which has elements $wi$(chips) assigned to one of the equiprobable values $wi \in \{ -\Delta, +\Delta \}$. The vector '$x$' is composed of magnitudes of various frames of Modulated Complex Lapped Transform (MCLT). The parameter $\Delta$ is set based on the sensitivity of ear to the changes in amplitude, which should be more than 1dB in this case since the original signal is a vector of the magnitude frequency components in dB. Once the watermark is introduced, the time domain watermarked audio signal is generated by the combination of watermarked vector $y$ with the original phase of $x$. Then these frames which are modified are passed through inverse MCLT. Only the MCLT coefficients between 2-7 kHz sub-bands are modified and considered in the watermark detection process. This decreases the carrier noise effect and sensitivity to compression and down sampling. The disadvantage of this technique is limited transmission rate, so to improve the transmission rate; watermark scheme in the wavelet domain is introduced [3].

*C. A blind audio watermarking scheme for embedding text using Godelization technique*

This paper presents a new blind audio watermarking scheme to hide the information into digital audio by encoding the data using godelization technique. Here .wav file is considered as cover signal to hide the secret information and the watermark data is the text. The data is embedded into the digital samples of the audio based on a secret key. This watermarking algorithm can extract the hidden watermark without the help of the original audio signals. Hence the scheme is referred to as 'blind'. In this method, a two layer approach is used to get the watermarked audio file. In the first layer, the watermark data to be hidden is first encoded using the concept of Godelization. In the second layer, the data encoded in step one isplaced into the last digit of the audio samples based on the secret key and a modified technique of low bit encoding schemes. The results show that the text data after embedding procedure is inaudible and there is not much distortion in the host.

*D. Audio watermarking technique robust against Time-Scale Modification*

This paper proposes a new content dependent robust audio watermarking scheme to resist against synchronization attacks like cropping and time-scale modification. The idea is to select steady high-energy regions first that represent music edges like transitions, note attacks or drum sounds by using different methods and then to embed the watermark in those high-energy regions. The algorithm was applied to a set of audio signals like pop, saxophone, rock and piano (15 seconds, mono, 16 bits/sample, 44.1 kHz).The signal-to-noise ratio is obtained to be 29.5 dB [28].

*E. Efficiently self-synchronized audio watermarking*

In this paper, a new self-synchronization algorithm is implemented for audio data transmission. Here the synchronization code is combined with watermark data thus the embedded signal will have self-synchronization ability. To achieve the robustness against attacks the synchronization code and the information bits are introduced into the low frequency coefficients in transform domain (DWT). Because of the time-frequency localization characteristics of wavelet domain, the computational load in searching the position of synchronization codes will be decreased. The basic idea in this scheme is to first divide the long audio signal into smaller segments and then to apply DWT to get the frequency coefficients. Then the watermark with synchronization code is embedded in the low frequency DWT coefficients. Then inverse DWT is applied to get the watermarked signal. In the data extraction part, the embedded signal is segmented and DWT is applied. The data is extracted from the DWT coefficients of low-frequency sub-band. Once the data is extracted, synchronization code is searched until it is found. With the position of synchronization is found, the hidden information is extracted [4].

## III. METHODOLOGY

The idea of the proposed audio watermarking scheme is to place the watermark along with synchronization code into the original audio in time domain. The host signal is initially segmented into frames and EMD algorithm is applied on each frame to obtain the respective IMFs. Then the binary sequence consisting of SC and watermark bits are embedded into the extrema of set of last IMFs of all frames using Quantization Index Modulation (QIM). A bit 0 or 1 is embedded per extrema. The frames are reconstructed using modified IMFs by applying inverse EMD and the obtained frames are further concatenated to get the watermarked audio signal.

In the extraction process, the watermarked audio is segmented into frames to apply EMD to each frame. Then the watermarked data is obtained by searching the positions of Synchronization code. The number of IMFs should be equal before and after watermarking so to make sure they are same; sifting of watermark signal is forced. The audio watermarking scheme used in this project is considered as blind since the host signal is not used for the extraction process. The block diagram of the proposed system is shown in Figure 1.

The watermarking scheme makes use of,
1. Host signal – The host signal is the audio signal that is of length 4 seconds saved in .wav format. For testing different audio signals like speech, music, pop, jazz, rock and classical are used in the project.
2. Watermark – The watermark might be any pattern of bits or the image of M* N bits that need to be introduced into the host signal.
The overview of the proposed system includes the following steps:

- Watermark pre-processing.
- Embedding of watermark.
- Extraction of watermark.

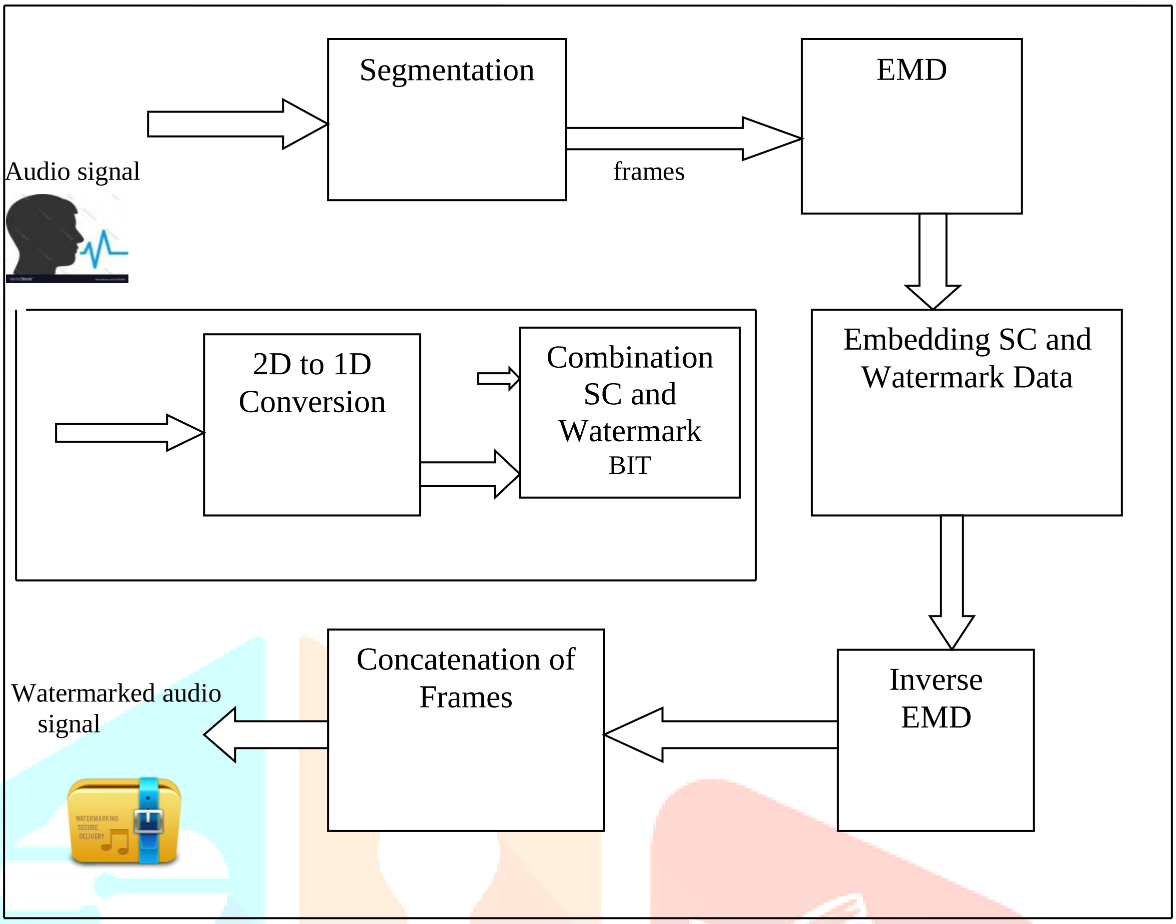


**Figure 1: Block diagram of the proposed system**

### A. Watermark pre-processing

Before embedding, the watermark should be pre-processed so that it can be embedded into the audio signal. It involves 2 steps:

**Conversion of 2D to 1D**

The binary two-dimensional image of size M × N bits is converted into one-dimensional sequence so that it can be embedded into the host audio signal since the audio signal is one-dimensional.

**Combination of SC and watermark**

The synchronization codes are combined with the watermark information bits to form the binary sequence denoted by $bi \in \{0, 1\}$, i-th bit of watermark as shown in Figure 2.

| | | |
|---|---|---|
| Sync-Code | Watermark Bits | Sync-Code |

**Figure 2: Binary sequence to be embedded**

The synchronization code used here is the 16-bit Barker sequence 1111100110101110. The need of synchronization, its properties are explained below.

**Synchronization and Synchronization code**

In many audio watermarking schemes, a synchronization problem occurs where the digital watermark bits are embedded or hidden in the specific places of the host signal. So in order to find the of hidden watermark data, the extraction procedure need to know the position of the embedded bits. Synchronization is the important issue in case of extraction since the host audio signal may be changed by desynchronizing attacks. So any shift in the position of bits may make it unable to extract the watermark data. So the main goal of synchronization methods is to find the new shifted position of bits. In the extraction process, it first tries to align itself with the watermarked block. If it fails, it will be unable to extract the hidden digital watermark in the audio signal and thus synchronization codes. The following three issues to be taken care when we have to choose the synchronization codes:

- The number of synchronization code bits.
- The number of zeros and ones in the code.

- How these synchronization code bits to be embedded into the host audio signal.

**B. Embedding of watermark**

The watermark embedding process involves the steps detailed as follows:

**Segmentation**

The process of splitting an audio signal into number of frames is called segmentation. Since our ear cannot response to very fast change of speech or audio data content, we normally cut the audio data into frames before analysis. To convert any signal into frames, a continuous audio signal is sampled at required sampling frequency. Here the original audio signal of length 4 seconds is sampled at 44.1 kHz and is divided into non-overlapping frames with a frame size of 64 samples each.

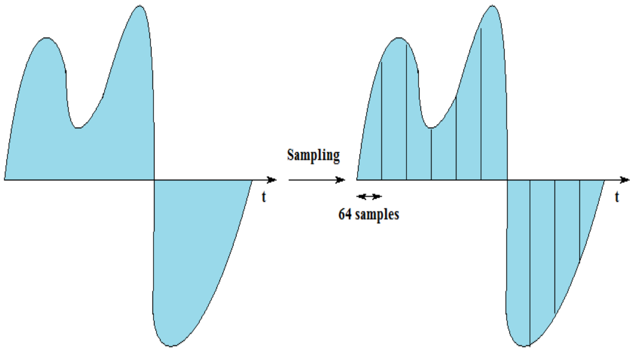


**Figure 3: Segmentation of an audio signal into frames**

**Empirical Mode Decomposition (EMD)**

EMD is conducted on every frame to decompose each frame into its consecutive IMFs [6] [7] [8]. The EMD and steps for extracting IMFs is explained as follows.

Empirical Mode Decomposition (EMD) is a new time domain approach for examining the non-stationary signals. This method was proposed by Norden. E. Huang in 1996. This is a signal processing technique without linearity and stationary assumptions. EMD is a data-driven scheme which does not depend on any basis functions that is predefined. EMD algorithm splits the multi component signals into mono-component signals. That mono-component signal itself is called as Intrinsic Mode Functions (IMFs). To extract the IMFs, the process is implemented as follows:
Consider a signal *x(t)*,
**Step 1** – Find all the extrema (either maxima or minima) of the signal *x(t).*

**Step 2** – Connect all the local maxima by interpolation to get upper envelope $Emax$ and repeat the same procedure to connect all local minima to get the lower envelope$Emin$.

**Step 3** – Find the mean of the upper envelope $Emax$ and lower envelope $Emin$to get *m(t).*

$$(t)=(Emax+Emin)/2 \quad (1)$$

**Step 4** – Find the difference of the original signal *x(t)* and the mean *m(t)* to obtain the proto-mode function (PMF) that is $h1$.

$$(t)-m(t)=h1 \quad (2)$$

Where $(t)$ is the original signal and $(t)$ is the mean.

The PMF should be processed further as shown in equation (1) and (2) to get the IMF. This repeated processing is called as 'sifting'. The sifting process is implemented in the first-step processing of EMD technique that contains the parameters
**Step 5** – Separate IMF from the original signal. That is, subtract first IMF from the original signal to get the residue *r(t).* The residue is considered as the new data and the above steps are repeated to get the other Intrinsic Mode Functions. The procedure is repeated until the residue is monotonic.

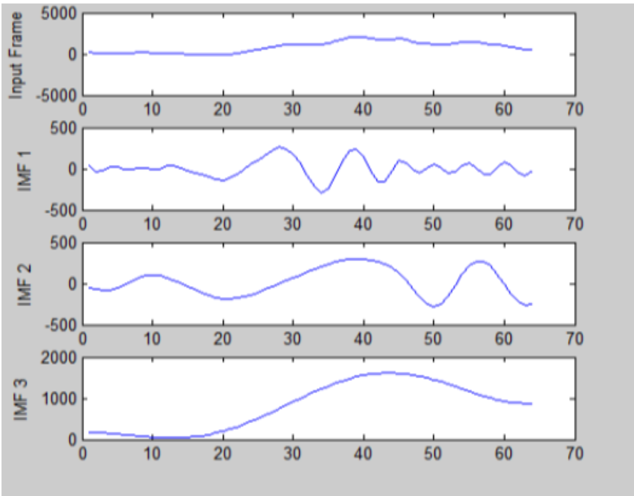


**Figure 4: Decomposition of a frame into its respective IMFs**

To reconstruct the original signal, if we add the entire Intrinsic Mode Functions and residue we will get back the original signal. The decomposition of a frame into corresponding IMFs when EMD algorithm is applied is shown in Figure 4.

**Embedding using QIM**

The binary sequence (watermark bits along with synchronization codes) is embedded 'P' times into the extrema of the last Intrinsic mode function (IMF) by Quantization Index Modulation (QIM) [13]:

Quantization Index Modulation is one of the best techniques to carry out blind watermarking which refuses the host-signal interference. The QIM technique works by first embedding the data by modulating the index with the embedded data and then by quantizing the original signal with the related quantizer.

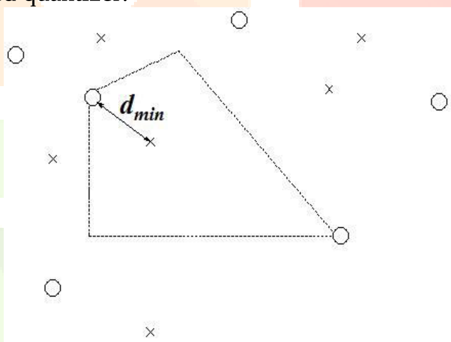


**Figure 5: Quantization Index Modulation for data embedding**

The points × are not same as points 0, thus the reconstruction points are not intersecting. This results in rejection of host-signal interference. The signal value $s$ changes from one × point $(m=1)$ to other or from one 0 point $(m=2)$ to other, as the value of $x$ changes but never changes from one point × to point 0 or from point 0 to point ×. is less. The minimum distance $dmin$ between the reconstruction points of quantizes.

**Reconstruction of frames**

Using the modified IMFs, the frames are reconstructed by applying inverse EMD to the IMFs. The frames are obtained by adding all the IMFs

**Concatenation**

The obtained frames are concatenated to get the watermarked audio signal.

**C. Extraction of watermark**

In the extraction process, the binary data is extracted and the synchronization code in extracted data is searched. This procedure is
In the extraction process, the binary data is extracted and the synchronization code in extracted data is searched. This procedure is repeated by shifting the selected window (segment) only one sample at once until the synchronization code (SC) is found. Once the SC is found, we can easily extract the information bits which are followed by SC. Let $y=\{bi\}$ denotes the binary watermarked data that is extracted and 'A' denotes the original synchronization code, to locate the binary watermark bits we search the SC initially in the sequence $bi$ bit-by-bit. The basic steps in the watermark extraction are detailed as follows:

**Segmentation**

The watermarked audio signal is split into equal number of non-overlapping frames similar to the embedding procedure.

**EMD**

Each frame is decomposed into associated IMFs by applying EMD similar to the process in embedding shown in Fig 4.4.

**Extraction of bits**

Once the IMFs are obtained, the bits can be extracted from the extrema of last IMF of frames. First extrema $\{ei*\}$ of IMF is found and from the extrema$ei*$ binary bits $mi*$ are extracted
where$ei*$ is the extrema of last IMF, $S$ is is the embedding strength taken as 0.98 and $mi*$ is the binary bits to be extracted.

**Searching of SC**

The searching involves the following steps
**Step 1:** Set the start index of the extracted binary data y = $\{bi\}$ to I = 1 and select $L = N1$ samples (Sliding window size).
**Step 2:** Compare the similarity between the extracted data segment $B = y\ (I: L)$ and A (Synchronization Code) bit by bit.
If the value after the similarity is ≥ , then $B$ is taken as the Synchronization Code (SC), then go to step 3.
**Step 3:** Increase the value of $I$by 1 digit, then slide the sliding window to the next $L = N1$ and repeat step 2
**Step 4:** Compare the second extracted segment $B'=y\ (I+N1+N2:I+2N1+2N2)$ and $A$ (SC) bit by bit as in step 2.
**Step 5:** If the new value $I$, $I \leftarrow I+N1+N2$ is equal to sequence length of bits, go to next step (Extraction of P watermarks) or go to step 3.

**Extraction of P watermarks**

The 'P' watermarks which are embedded are extracted and comparison is made between these bits for correction and finally the watermark bits are extracted.

## IV. RESULT AND DISCUSSION

To perform simulations on audio signal the following parameters are considered. These parameters are chosen to compromise between robustness, imperceptibility and payload.
Sampling frequency - 44.1 KHz
Binary logo (watermark) - M*N = 50*50
Synchronization Code - Barker sequence 1111100110101110
Embedding strength S - 0.98
Threshold- 4
Number of samples per frame – 64 samples

*Data payload*

Data payload refers to the number of bits that are embedded per second into the audio signal. If number of watermarks bits that need to be hidden is *'M'* and the length of audio signal is *'L'*, then the payload or data embedding rate *'P'* is given by,

$$P=M\div L \qquad (8)$$

For testing 4 seconds audio signal is taken and number of bits to be embedded is 50*50=2500 bits. Then the data payload is 625 bps.

*Signal to Noise Ratio (SNR)*

The Signal to Noise Ratio is used to estimate the quality and imperceptibility of the watermarked audio signal with respect to the original host signal. The Signal to Noise Ratio of a watermarked audio signal should be more than 20dB as per the recommendation of International Federation of the Photographic Industry (IFPI). Higher the SNR value, better the degree of hidden watermark imperceptibility. The Signal to Noise Ratio can be obtained.

*Bit Error Rate (BER)*

Bit Error Rate is defined as the ratio of bits that are extracted incorrectly to the number of bits that are embedded. To examine the detection accuracy of the watermark after signal processing attacks.
*Results for different audio signals*
The above procedure is repeated for other audio signals like song, pop, jazz, rock and classical and their results are explained below.
**Song:** The input and the watermarked audio are shown in Figure 6.

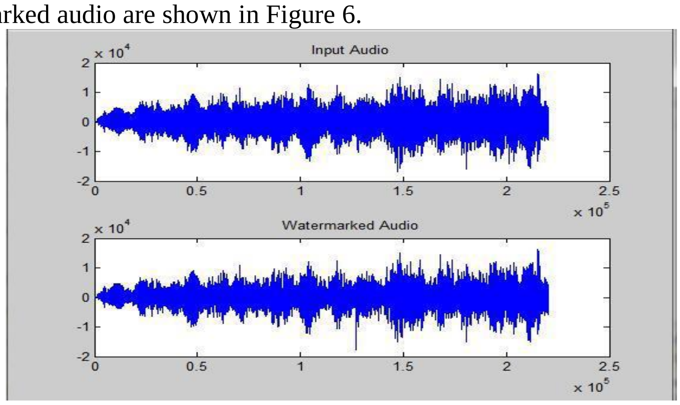


**Figure 6: Input and the watermarked audio (song)**

**Classical**

The input and the watermarked audio are shown in Figure 7.

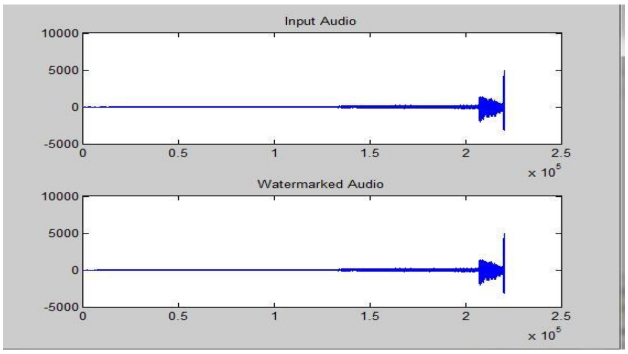


**Figure 7: Input and the watermarked audio (classical)**

## V. CONCLUTION AND FUTURE WORK

### *A. Conclusion*

Watermarking is the process of introducing any predefined data into images in such a way that the degradation of quality of host is minimized and is imperceptible. Many digital watermarking algorithms have been proposed both in time and frequency domain. In this project, a time domain scheme based on Empirical Mode Decomposition (EMD) algorithm is introduced. Here the secret watermark data along with synchronization code is embedded by using Quantization Index Modulation (QIM) in the very low frequency mode (last IMF) achieving good robustness against several attacks. Because of synchronization code, the watermark will have the ability to resist against shifting and cropping. This technique has higher payload and robustness that is demonstrated by the simulations that is carried out. And it is seen that the watermarked audio is indistinguishable from the original audio. The proposed system doesn't use original signal for the extraction procedure, hence it is blind and it involves easy calculations.

### *B. Future work*

1. Encryption can be added with the proposed system to provide authentication or security.

2. In this watermarking method, embedding strength S is kept constant. To improve the performance, it can be made adaptive to the type and magnitude of original signal used.

3. To further improve the performance, the characteristics of psychoacoustic model and Human Auditory system can be included.

4. It can also be checked if the method works out for different sampling rate with same payload.

## REFERENCE

[1] J. Cox and M. L. Miller, "The first 50 years of electronic watermarking," J. Appl. Signal Process., vol. 2, pp. 126–132, 2002.
[2] 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.
[3] D. Kiroveski and S. Malvar, "Robust spread-spectrum audio watermarking," in Proc. ICASSP, 2001, pp. 1345–1348.
[4] 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.
[5] 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.
[6] N. E. Huang et al., "The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis," Proc. R. Soc., vol. 454, no. 1971, pp. 903–995, 1998.
[7] K. Khaldi, A. O. Boudraa, M. Turki, T. Chonavel, and I. Samaali, "Audio encoding based on the EMD," in Proc. EUSIPCO, 2009, pp. 924–928.
[8] K. Khaldi and A. O. Boudraa, "On signals compression by EMD," Electron.Lett., vol. 48, no. 21, pp. 1329–1331, 2012.
[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.