

## EMD Based Audio Watermarking With TSM Attack

Kishor H.Bamne<sup>1</sup>, V. M. Harnal<sup>2</sup>, V. M. Jain<sup>3</sup>

<sup>1,2,3</sup> P. G. Dept., MBES College of Engineering, Ambajogai, india, 431517

---

**Abstract:** Audio files are capable of hiding the data and identity of the owner which is called Watermark. We present a watermarking procedure to embed for copyright protection into digital audio by directly modifying the audio samples then audio signals are divided into no. of samples each sample is decomposed adaptively by the method of Empirical Mode Decomposition (EMD).those samples called as a 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. We show the robustness of the hidden watermark for Time-scale modification attack. TSM is a challenging attack and can be used for watermarking. TSM is the process of either compressing or expanding the time-scale between two extrema (successive maximum and minimum pair) of the audio signal. The comparison analysis shows that audio watermarking on no attack and TSM attack watermarking.

**Keywords:** Empirical Mode Decomposition (EMD), Intrinsic Mode Functions (IMFs), audio watermarking, Time-scale modification (TSM), synchronization code.

---

### I. Introduction

In digital media the embedding of watermarking in audio is for copyrights protection and authentication digital media by embedding a watermark in the original audio signal. Main requirements of digital audio watermarking are imperceptibility, robustness and data capacity. Digital watermarking has been proposed as a means to identify the owner or distributor of digital data. Watermarking is the process of encoding hidden copyright information in digital data by making small modifications to the data samples. It hides information in the data in such a way that the basic appearance of the data is not destroyed.

Digital watermarking has gotten a lot of consideration in the writing to give effective answers for copyright assurance of digital media by installing a watermark in the first sound sign [2]–[6]. Digital audio watermarking has gain a good attention to provide efficient solutions for copyright protection of digital media. The watermark should be inaudible within host audio so that its quality is maintained but it should be robust to the attacks or signal distortions given to the host data. At the end, watermark should be easy to extract and to prove ownership [6]. For protecting against various attacks on Audio. A robust scheme of watermarking is given but it has demerits like bit rate transmission [6]. To eliminate bit rate problem, watermarked schemes in the domain of wavelet has also been proposed. Watermarking in wavelet domain has fixed basis functions, thus they cannot guarantee to match all real time signals [4]-[5]. To conquer these drawbacks, Empirical Mode Decomposition (EMD), this new signal decomposition technique has been introduced. It works on both stationary as well as non-stationary signals. EMD does not require a priori choice of filters or any basis functions [7]. This scheme breaks down any signal in zero-mean symmetrical envelopes AM-FM modules known as Intrinsic Mode Functions [1].

Previously different methods have been proposed for audio watermarking but some problems are arises like robustness, Imperceptibility and data capacity .Now we are proposed a new algorithm in audio watermarking for the copyright protection. That is Empirical Mode Decomposition (EMD). EMD - based time-frequency analysis, called Hilbert-Huang Transform (HHT), this is only one of many applications made possible by EMD. The final result and ideas in time domain applications using EMD apply to two-dimensional signals, such as images, as well as audio [7]. EMD decomposes the spatial frequency components into a set of IMFs (Intrinsic Mode Functions) where the highest spatial frequency component of each spatial position is in the first IMF and the second highest spatial frequency component of each spatial position is in the second IMF, etc. An IMF is defined as a function in which the number of extrema points and the number of zero crossings are the same or differ by one [8]

Synchronization attacks are a serious problem to any watermarking schemes. Audio processing such as random cropping and time-scale modification (TSM) causes displacement between embedding and detection in the time domain and is hence difficult for watermark to survive. The basic idea of the algorithm is to change the length of the intervals between salient points of the audio signal to embed data. The intervals are quantized and the data is embedded in the quantization indices [10]. TSM is a serious attack to audio watermarking; very few algorithms can effectively resist this kind of synchronization attack. According to the Secure Digital Music

Initiative (SDMI) Phase-II robustness test requirement, a practical audio watermarking scheme should be able to withstand pitch-invariant TSM up to  $\mp 4\%$  [9].

## II. Proposed Watermarking Algorithm

The decomposition starts from finer scales to coarser ones. Any signal  $x(t)$  is expanded by EMD as follows:

$$x(t) = \sum_{j=1}^C IMF_j(t) + rc(t) \tag{1}$$

Where  $C$  is the quantity of IMF's  $rc(t)$  and signifies the last leftover. The IMFs are about orthogonal to one another, and all have almost zero means [1].

The basic idea of the proposed watermarking system is to hide the data into the original audio signal a watermark (secret data) with a Synchronized Code (SC) in the time domain format. The input signal which is original audio signal is first segmented into samples after that algorithm EMD is conducted on every samples to extract the associated IMFs. Then all the samples are converted into binary data sequence consisted of SCs and informative watermark bits (Fig. 1) is embedded in the extreme of a set of consecutive last-IMFs. All bit (1 or 0) is inserted per extreme. Finally, inverse transformation function is applied to the modified extrema to recover the watermarked audio signal by superposition of the IMFs of each sample followed by the concatenation of the sample in Figure 2. [1].

The proposed EMD watermarking scheme is blind, that is, the original signal (host signal) is not required for watermark extraction. These are the basically in 3 steps those are,

- 2.1. Synchronization of code
- 2.2. Watermark embedding
- 2.3. Watermark extraction

### 2.1. Synchronization of code.

To locate the embedding position of the hidden watermark bits in the host signal a SC is used. This code is unaffected by cropping and shifting attacks [4]. Let  $U$  be the original SC and  $V$  be an unknown sequence of the same length. Sequence  $V$  is considered as a SC if only the number of different bits between  $U$  and  $V$ , when compared bit by bit, is less or equal than to a predefined threshold  $\tau$  [1].

### 2.2. Watermark embedding.

Synchronization of code are combined with watermark bits from a binary sequence before embedding. It is denoted by  $n_i \in \{0, 1\}$ ,  $i$ -th bit of watermark. Basics of our watermark embedding are shown in Figure 2 and detailed as follows [1].

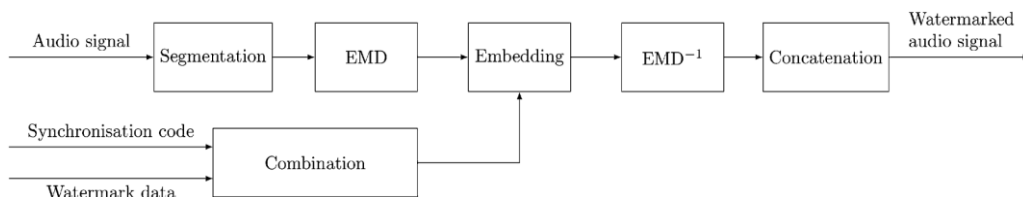


Fig.1. Watermark embedding presses [1].

Steps to Embedded:-

- Step1: Divide the original host signals to the no. of samples
- Step2: Each sample is decomposed into IMFs (intrinsic mode function).
- Step3: Embed  $T$  times the binary sequence  $n_i$  into extrema of the last IMF ( $IMF_Z$ ) by QIM.

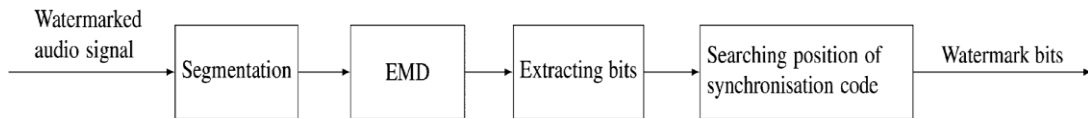
$$e_j^* = \begin{cases} \left[ \frac{e_j}{H} \right] * + \operatorname{sgn}\left(\frac{3H}{4}\right) & \text{if } n_j = 1 \\ \left[ \frac{e_j}{H} \right] * + \operatorname{sgn}\left(\frac{H}{4}\right) & \text{if } n_j = 0 \end{cases} \tag{2}$$

Where  $e_j$  and  $e_j^*$  are the extrema of  $IMF_Z$  of the original host audio signal and the watermarked signal respectively. If “Sgn” function is equal to “[+]” then it is a maxima, and equal to “[−]” then it is a minimal.  $S$  denotes the embedding strength.

Step4: Reconstruct the samples by using inverse EMD modified  $IMF_Z$  and concatenate the watermarked frames to retrieve the watermarked signal [1].

**3.3. Watermark extraction.**

Host signal is splitted into samples and EMD is performed on each one, as in embedding for the watermark extraction of binary data using rule given by (3).then find out the SCs in the extracted data. This procedure is as follows.



**Fig.2.** Watermark Extraction [1].

Continuously repeated by shifting the selected segment one sample at a time until a SC is found. With the position of SC determined, after that we can extract the hidden data i.e. information bits, which given as

- Step1: Divided the watermarked signal into no. of samples.
- Step2: Decompose each and every sample into IMFs.
- Step3: Extract the extrema  $e_j^*$  of  $IMF_z$
- Step4: Extract using the following equation.

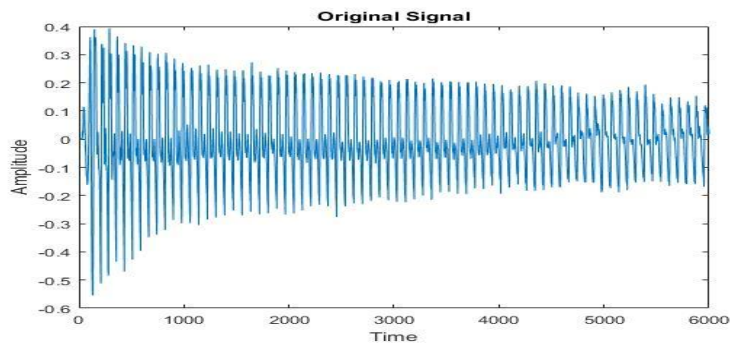
$$e_j^* = \begin{cases} 1 & \text{if } e_j^* - \left\lfloor \frac{e_j^*}{H} \right\rfloor \cdot H \geq \text{sgn}(H/2) \\ 0 & \text{if } e_j^* - \left\lfloor \frac{e_j^*}{H} \right\rfloor \cdot H \leq \text{sgn}(H/2) \end{cases} \quad (3)$$

- Step5: Set the count of the extracted data, y, to I=1 and Select= N1 samples (sliding window)
- Step6: Evaluate the similarity between the extracted segments  $V = y(I:L)$  and U bit by bit. If the similarity Value is  $\geq \tau$ , is taken as the SC and go to step8. Otherwise
- Step7: Increase the count, I by 1 and slide the window to the next  $L=N1$  samples and repeat step 6.
- Step8: Similarity between the second extracted segments  $V = y(I+N1:N2:I+2N1+N2)$  and U may evaluate bit by bit.
- Step9:  $I \leftarrow I+N1+N2$ , of the new I value is equal to sequence length of bits, go to Step 10 else repeat.
- Step10: Watermarks is extracted P times and make comparison in bit by bit between these watermark, for correction, and finally extract the desired watermark [1].

**III. simulation results**

Simulations are performed on audio signals sampled at 42.1 kHz. The embedded watermark, W, is a binary logo of size 1632 bits. We convert this 2D binary image into 1D sequence in order to embed it into the audio signal. The synchronization code used is a 16 bit Barker sequence. Each audio signal is divided into frames of size 64 samples and the threshold is set to 4.

Time scaling pitch-invariant TSM up to +4% TSM. These parameters have been chosen to have a good compromise between imperceptibility of the watermarked signal, payload and robustness. Following figure shows the simulation result of TSM attack on watermarked audio signal after applying the pitch-invariant 4% time speed increment.



**Fig.3.** Input audio signal.

For data extraction, the watermarked audio signal is splitted into frames and EMD is applied to each frame showed in figure 3.

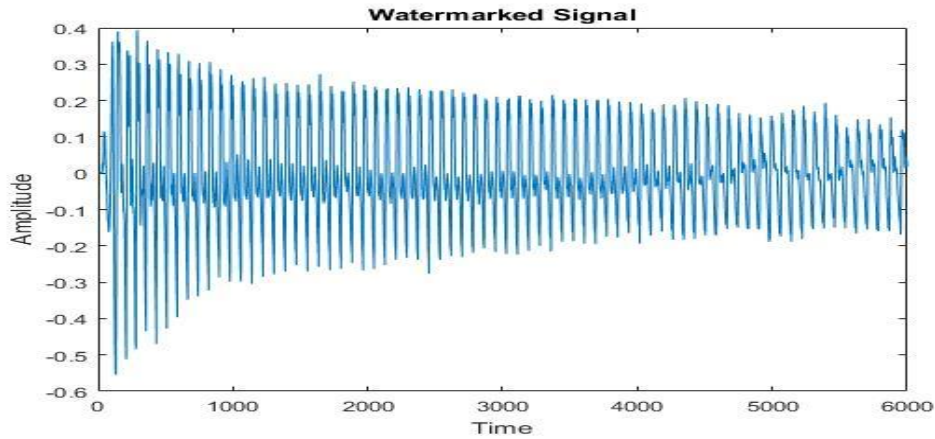


Fig.4. watermarked audio signal.

figure 4 shows its watermarked version, watermarked signal is visually indistinguishable from the original one.

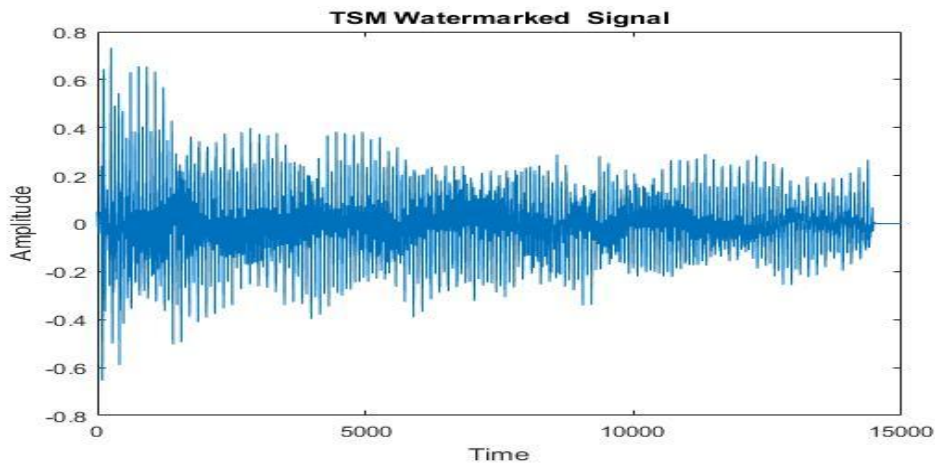


Fig.5.TSM watermark audio signal.

Figure 5 shows TSM watermark it is the process of either compressing or expanding the time-scale between two extrema (successive maximum and minimum pair) of the audio signal. By using TSM we can change the length of the intervals between salient points of the audio signal to embed data. The intervals are quantized and the data is embedded in the quantization indices.

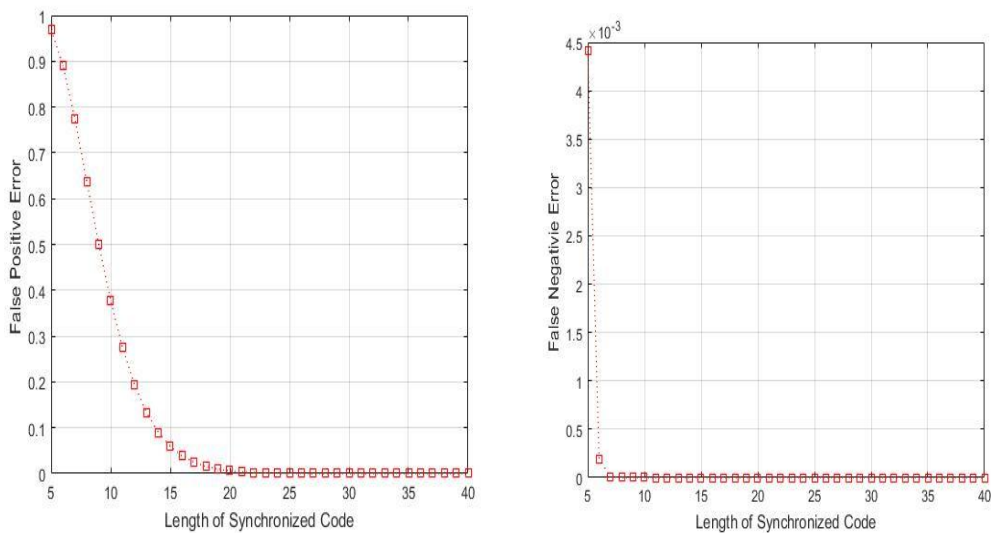


Fig.6.  $P_{FPE}$  &  $P_{FNE}$  versus synchronization code length.

A large NC indicates the presence of watermark while a low value suggests the lack of watermark. Two types of errors may occur while searching the SCs: the False Positive Error (FPE) and the False Negative Error (FNE). These errors are very harmful because they impair the credibility of the watermarking.  $P_{FPE}$  is the probability that a SC is detected in false location while  $P_{FNE}$  is the probability that a watermarked signal is declared as un-watermarked by the decoder [1].

Following table shows the comparison analysis of audio watermarking on no attack and TSM attack

**Table 1.** Comparison of watermark On Attack & No attack.

Parameter	No Attack	TSM Attack
1.SNR Value	56.8754	56.8754
2.ODG Value	-0.8	-0.7
3.BER value	0.484	0
4.NC value	0.6047	1

#### IV. Conclusion

In this paper, for achieving a good performance watermark is added with the last IMF. In addition of synchronization code helps to resist the data loss during shifting and cropping. The input signal is first segmented into frames and EMD is conducted on every frame to extract the associated IMFs. For data extraction, the watermarked audio signal is split into frames and inverse EMD is applied to each frame. Binary data sequences are extracted from each last IMF by searching for SC.

The proposed algorithm has greater robustness against common attacks such as Time scale modification attack than recently proposed algorithms. In MP3 and wave file compression, it produces a better performance compared to existing and old audio watermarking methods. Using of EMD algorithm results have better performances that self-adaptive decomposition of the audio signal, low false positive and false negative error probability rates and easy calculations.

#### Acknowledgements

We convey our sincere thanks to the principal Dr. B. I. Khadakhavi, dean of P.G. department Dr. B .M. Patil and staff of MBES's College of Engineering, Ambajogai for help in carrying out this research work at the institute.

#### References

- [1] KAIS KHALDI AND ABDEL-OUAHAB BOUDRAA "AUDIO WATERMARKING VIA EMD" IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 21, NO. 3, MARCH 2013.
- [2] I. J. Cox and M. L. Miller, "The first 50 years of electronic watermarking," J. Appl. Signal Process., vol. 2, pp. 126–132, 2002.
- [3] 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.
- [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] D. Kiroveski and S. Malvar, "Robust spread-spectrum audio watermarking," in Proc. ICASSP, 2001, pp. 1345–1348.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] Mohamed F. Mansour and Ahmed H. Tewfik, "Data Embedding in Audio Using Time-Scale Modification" IEEE Transactions on Speech and Audio Processing, Vol. 13, No. 3, May 2005.

IOSR Journal of Electronics and Communication Engineering (IOSR-JECE) is UGC approved Journal with Sl. No. 5016, Journal no. 49082.

Kishor H.Bamne EMD Based Audio Watermarking With TSM Attack." IOSR Journal of Electronics and Communication Engineering (IOSR-JECE), vol. 12, no. 5, 2017, pp. 74-78.