Audio Watermarking with The Help of EMD

Vaibhav Raut[1] Prof. Kishore Wane[2]

[1,2]Dhole Patil college of Engineering, Pune

Abstract— In this paper a new audio watermarking technique Empirical Mode Decomposition (EMD) is introduced. The audio signal is first split into frames and each one is then decomposed by EMD to get intrinsic periodic elements known as Intrinsic Mode Functions (IMFs). The combination of watermark and synchronization code is embedded into the last IMF, which is most stable frequency component under different attacks and protects quality of the host signal. So this methodology is applied to any image and audio signals. Experimentation has given the physical marking and its detection to protect audio on different attacks.

Key words: Empirical mode decomposition, Intrinsic Mode Function, quantization index modulation, synchronization code, amplitude modulation, signal to noise ratio, frequency modulation

I. Introduction

Digital audio watermarking has gain a good attention to provide efficient solutions for copyright protection of digital media. Main goals of digital audio watermarking are robustness and imperceptibility. 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. To attain above necessities, pursuing new schemes for watermarking is a very difficult. Different techniques for watermarking with different complexities have been given. For protecting against various attacks on audio, a robust scheme of watermarking is given but it has demerits like bit rate transmission. 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. 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. This scheme breaks down any signal in zero-mean symmetrical envelopes AM-FM modules known as Intrinsic Mode Functions (IMFs). Any signal is expanded by EMD as follows:

$$x(t) = \sum_{j=1}^{c} IMF_j(t) + r_c(t) \qquad (1.1)$$

where C denotes IMF count and r_c (t) denotes final residual

The IMFs are almost orthogonal to each other and have zero means. The quantity of extrema is reduced when it from one to the next mode, and the whole scheme of decomposition is band-pass filtering and cropping. Similarly in other scheme, EMD with Hilbert transform and the watermark is embedded into the IMF covering highest energy. However, why the IMF carrying the highest amount of energy is the best mode to hide the watermark has not been given. In practice an IMF with highest energy can be a highest frequency mode and so it is not robust to attacks. Watermarks inserted into lower order IMFs or high frequency components are most susceptible to attacks. It has been claimed that for robustness, the watermark bits are generally embedded in the perceptually that is low frequency components of the host signal to have better opposition against attacks and imperceptibility. Thus in this project, the watermark are embedded in the extrema of the last IMF. The watermark is linked with a synchronization code to assist its location. The benefit to use the time domain approach in EMD, is the low cost in probing synchronization codes. Audio signal is first segmented into frames where each one is decomposed into IMFs. Bits are added into the extrema of the last IMF such that the watermarked signal inaudibility is confirmed. Experimental results prove that the hidden data are robust for attacks such as noise additions, compressions, requantization, cropping etc. This method is easy and robust as compared to audio watermarking approaches informed recently

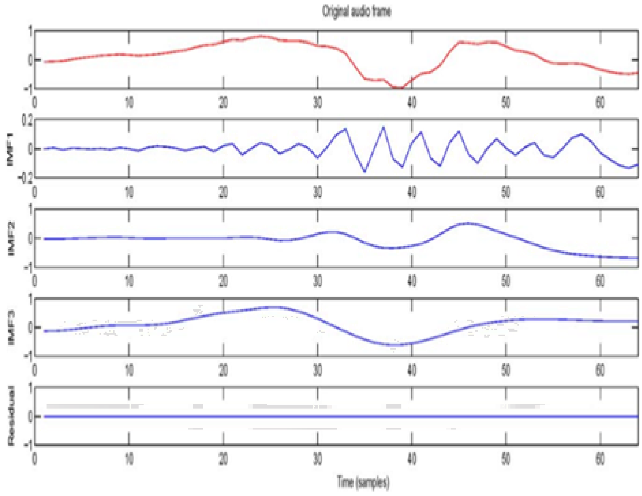


Fig. 1: Decomposition of an audio frame by EMD

II. Proposed Methodology

The idea behind proposed watermarking method is to hide watermark and synchronized code (SC) itself into the original audio signal in the time domain. The input signal is first segmented in frames and then EMD is carried on each frame to extract the respected IMFs (Fig.1). After that a binary data

Sync code	Watermark bits	Sync code

Fig. 2: Data combination structure

sequence containing combination of SCs and watermark bits (Fig.2) is embedded in the extrema of consecutive last-IMFs. A bit (0 or 1) is introduced per extrema. As the number of IMFs and their number of extrema depend on the amount of data of each frame, the number of bits that has to be embedded varies from last-IMF of one frame to the succeeding. sequence containing combination of SCs and watermark bits (Fig.2) is embedded in the extrema of consecutive last-IMFs. A bit (0 or 1) is introduced per extrema. As the number of IMFs and their number of extrema depend on the amount of data of each frame, the number of bits to be embedded changes from last-IMF of one frame to the succeeding. Combination of Watermark and SCs are not all embedded in extrema of last IMF of one frame only. In general, the number of extrema per last-IMF is very small as matched to length of the binary sequence to be embedded. This is totally depends on the length of the frame. If we design the numbers of bits of SC by N1 and watermark by N2, the total length of binary sequence to be embedded equals 2N1+N2 . Thus, these 2N1+N2 bits are spread out on numerous last-IMFs of the consecutive frames. This sequence of bits is introduced P times. Finally, inverse decomposition is carried to the modified extrema to recover back the watermarked audio signal. This is done by superposition of the IMFs of each frame trailed by concatenation of the frames (Fig.3). For extraction of data, the watermarked audio signal is split into frames and EMD carried on each frame (Fig.4). Binary data sequences are thus extracted from each last IMF by probing for SCs (Fig. 5). Fig.6 gives the last IMF before and after watermarking. This figure displays that there is little difference in terms of amplitudes between two modes. EMD is fully data adaptive, thus it is important to assure that the number of IMFs will be same before and after embedding of the watermark (Figs.1,4). In fact, if the numbers of IMFs are different, there is no any assurance that the last IMF always have the watermark information to be extracted. To overcome this issue, the sifting of the watermarked signal is compelled to extract the same number of IMFs as before watermarking. The proposed watermarking scheme does not need host signal is not required for watermark extraction. Overview of the proposed method is detailed as follows:

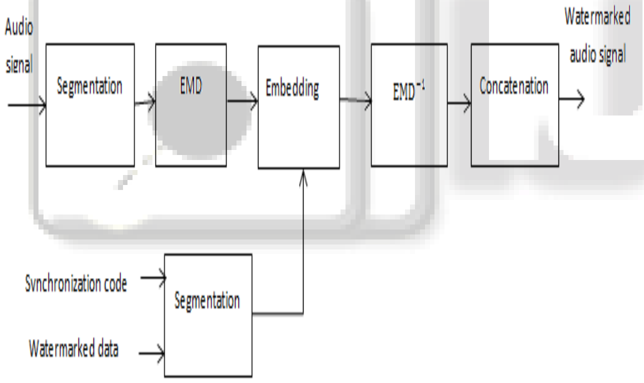


Fig. 3: Watermark embedding

A. Synchronization Code:

To trace the embedding position of the hidden watermark bits in host signal, SC is used. This code remains unaffected by cropping or shifting attacks. Let U be the original SC and V be any unknown sequence of the same length. Sequence V is considered as a SC if and only if the number of different bits between U and V, compared bit by bit, is less or equal to a predefined threshold [3].

B. Watermark Embedding:

Beforehand embedding, a binary sequence is made by combination of SCs and watermark bits which is denoted by – mi ∈ {0,1}, i-th bit of watermark (Fig. 2). Watermark embedding basics are shown in Fig.3 and thorough as follows:

Step 1: Split the audio signal into frames.
Step 2: Decompose every frame into IMFs.
Step 3: Embed the binary sequence { mi }, P times in extrema of the last IMF [13].

$$e_i^* = \begin{cases} \left\lfloor \frac{e_i^*}{S} \right\rfloor . S + \operatorname{sgn}\left(\frac{3S}{4}\right) & if\ m_i = 1 \\ \left\lfloor \frac{e_i^*}{S} \right\rfloor . S + \operatorname{sgn}\left(\frac{S}{4}\right) & if\ m_i = 0 \end{cases} \quad (2.1)$$

where, e_i-extrema of IMF of the host signal and 〖e_i〗^* - watermarked signal. sgn function equals " + " if e_i is a maxima, and " - " if it is a minima. ⌊ ⌋ means the floor function, and S represents the embedding strength selected to preserve the inaudibility constraint.

Step 4: Rebuild the frame changed IMFs and concatenate the watermarked frames to recover the watermarked signal.

C. Watermark Extraction:

For extraction of watermark, host signal is spliced into frames and EMD is implemented on each one as in embedding. The binary data can be extracted using rule given by (3), then search for SCs in that extracted data. This procedure is

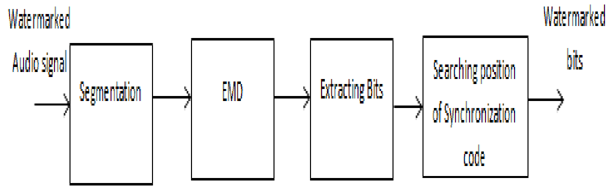


Fig. 4: Watermark extraction

recurring by shifting the selected segment, one sample at time until SC is found. With the location of SC determined, we can then recover the hidden information bits. Let y = { mi*} is binary data to be extracted and U is the original SC. To locate the embedded watermark, examine the SCs in the sequence { mi*} bit by bit. The extraction is achieved without using the original audio signal. Basic steps present in the watermarking extraction, shown in Fig.4, are given as below:

Step 1: Split the watermarked signal in frames.
Step 2: Decompose every frame into the IMFs.
Step 3: Extract the extrema { ei*} IMF .
Step 4: Extract using the following rule [3]:

$$m_i^* = \begin{cases} 1 \ \ if\ e_i^* - \left\lfloor \frac{e_i^*}{S} \right\rfloor . S \ \geq \operatorname{sgn}(S/2) \\ 0 \ \ if\ e_i^* - \left\lfloor \frac{e_i^*}{S} \right\rfloor . S < \operatorname{sgn}(S/2) \end{cases} \quad (2.2)$$

Step 5: Set the index of the extracted data for start, y , to I =1 and select L = N1 samples
Step 6: Calculate the matching between the extracted segment V = y (I : L) and U bitwise. If the matching value greater ,then V is considered as the SC and goes Step 8. Else continue the next step.
Step 7: Increment I by 1 and slide the window further to next L = N1 samples and redo Step 6.
Step 8: Calculate the matching between the other extracted segment, V' = y.(I + N1 + N2 : I + 2N1 +N2) and U bitwise.
Step 9: I ← I + N1+N2, the new I value is equals to bits sequence length, go to Step 10 or else do Step 7 again.
Step 10: For correction, extract P watermarks and make comparison bitwise between these, and finally extract the required watermark. Embedding and extraction of watermark processes are given in Fig. 6.

Fig. 5: Binary Watermark

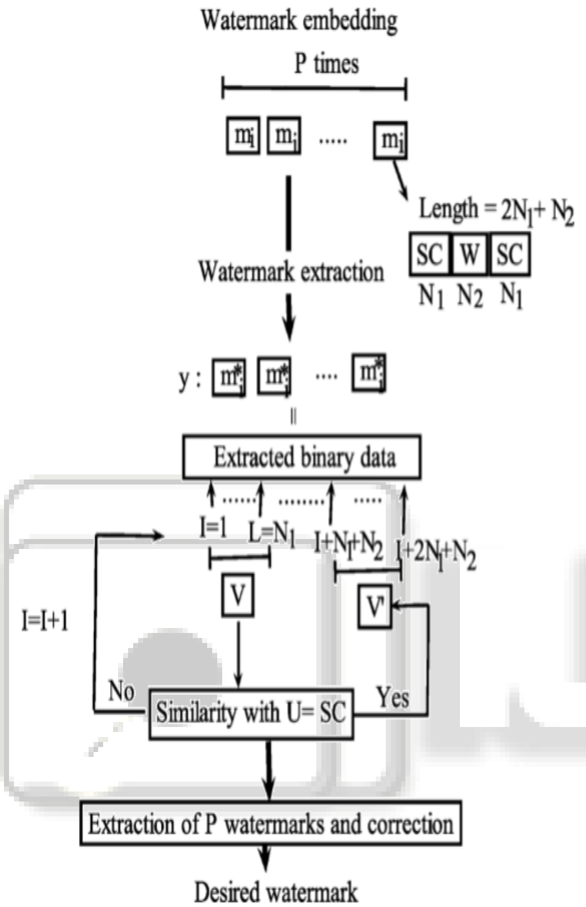


Fig. 6: Embedding and extraction of watermark actual process

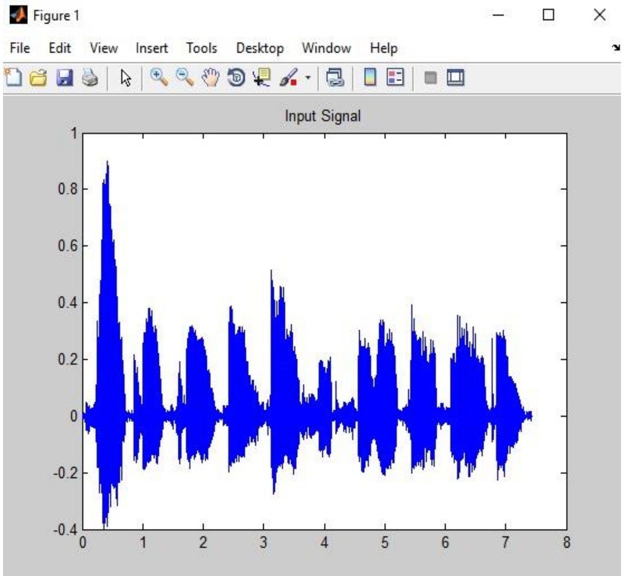
III. QUALITY ANALYSIS

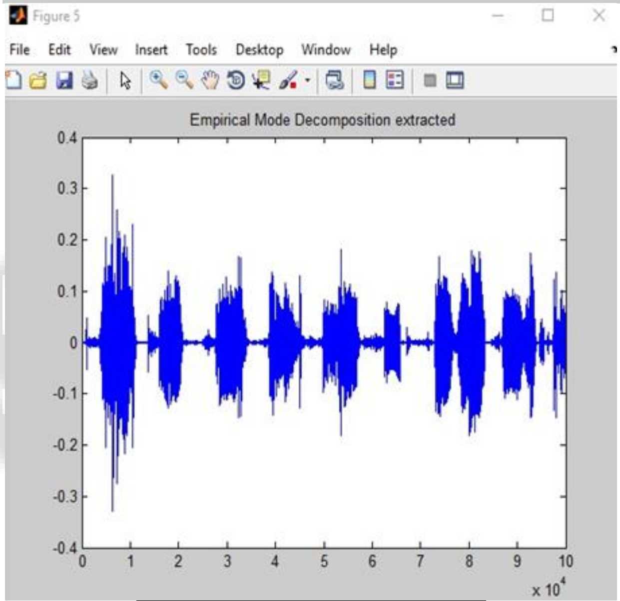
To estimate the performance of proposed method Signal to Noise Ratio (SNR), Bit error rate (BER) in between original and the watermarked audio signals is considered. A watermarked audio signal should maintain more than 18 dB SNR. To evaluate the watermark detection accuracy after attacks ,we used the and the defined as follows:

$$BER(W,W') = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} W(i,j) \oplus W'(i,j)}{M*N} \qquad (3.1)$$

where $\oplus$ denotes XOR operator and M × N denotes the sizes of binary watermark image. W and W' gives original and the recovered watermark separately. Is used to evaluate the watermark detection accuracy after signal processing operations.

IV. RESULT





Audio Signal	BER	SNR
Signal 1	1.543	18.92
Signal 2	1.792	21.38
Signal 3	1.133	19.31

Table 1: Beer and Snr of Watermarked Audio

V. CONCLUSION

In this paper a new watermarking scheme based on the EMD is projected. Watermark is inserted in very low frequency mode. It is added up with synchronization codes, so it protects against various attacks. Data bits of the synchronized watermark are embedded in the extrema of the last IMF of the audio signal so does not hampers the quality of host signal. As it works in real time and as it has adaptive nature, it is a better watermarking scheme. In all audio test signals, the watermark introduced is imperceptible. Experiments gives that the watermarked audio signals are not distinguishable from their original ones. This watermarking method has easy calculations and does not use the original audio signal

REFERENCES

[1] I. 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. Tew fik, “Robust audio watermarking using perceptual masking,” Signal Process., vol. 66, no. 3, pp. 337–355, 1998.
[3] 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.
[4] 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.
[5] D. Kiroveski and S. Malvar, “Robust spread-spectrum audio watermarking,” in Proc. ICASSP, 2001, pp. 1345–1348.
[6] N. E. Huanget 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. IEEECCECE,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.

