



Robust Spread Spectrum Watermarking Technique Based on Empirical Mode Decomposition

Jeebananda Panda¹, Karan Raj Gera², Asok Bhattacharyya³

¹Deptt. of Electronics and Communication Engineering, Delhi Technological University, INDIA

²Deptt. of Electronics and Communication Engineering, S.V National Institute of Technology, INDIA

³Deptt. of Electronics and Communication Engineering, Delhi Technological University, INDIA

Abstract: In this paper, an audio watermarking scheme based on Empirical Mode Decomposition (EMD) is proposed that uses spread spectrum technique for the embedding process. The audio signal is separated into segments, which, by EMD are decomposed into their respective frequencies that are represented by intrinsic mode functions (IMFs). These functions are fully described by their extrema. Having the lowest frequency, the spread spectrum technique is performed over the extrema points of the last IMF. The technique involves Discrete Cosine Transform (DCT) of the extremas, that are rearranged in descending order before the watermarking bits are added to them using additive embedding technique. The embedding process is carried out on the last IMF while the extraction procedure is performed over the lower order IMFs using non-blind technique. Through rigorous simulations, the scheme is proved to be robust against various attacks like additive noise, cropping, high pass and low pass filtering, and wiener filter..

Keywords: Additive Watermarking, Empirical Mode Decomposition, Intrinsic Mode Function, Spread Spectrum, Non Blind Technique

I. Introduction

Recent improvement and growth of the cyber world has generated the capability and motivation for electronic watermarking as means of copyright protection [1]. The watermarking techniques are categorized into fragile or robust, and blind and non-blind embedding techniques. This paper concerns robust watermarking technique that uses non-blind technique for the extraction of the same. Robustness means that the algorithm is resilient to some audio distortions such as cropping, filtering, additive noise and mp3 compression. At this stage, the scheme possesses tolerable immunity against the first three attacks, but near future is expected to see the scheme to work its way out of the last attack too.

An audio watermarking scheme needs to bear some basic features: 1) Imperceptibility: the quality of the audio signal should remain intact after the watermarking process; 2) Robustness: the watermarking technique should be robust against various attacks; 3) Capacity: the watermarking scheme should bear enough capacity to embed a certain amount of watermarking bits into the audio signal; 4) Security: security implies that the watermark can be detected only by the authorized person. Often, these featured compliment each other which calls for the need for proposition of different watermarking techniques to reach an optimum result.

Recently, EMD [2]-[4] has gained popularity as a method to extract the constituent frequencies from a signal as compared to the wavelet transform [5]. In [5], many multi-resolution sub-bands are created when Discrete Wavelet Transform (DWT) is performed over the host audio signal; and the watermarking bits are embedded in the highest resolution sub-band. However, its basic functions are fixed and may not match the shape of the considered data series in every instant in time limiting its usage and growing interest in EMD. EMD, on the other hand divides the signal into a set of functions called the Intrinsic Mode Functions (IMF), by a process known as the sifting process. According to the decomposition rule, any signal can be expanded as:

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

Where C is the number of IMFs and $r_c(t)$ denotes the final residual.

Moreover, introduction of an embedding technique called the spread spectrum [6] increases the robustness of the scheme. In [6], Cox et. al. constructed the watermarks based on the principle that they be identically distributed (i.i.d) Gaussian random vector that are inserted into significantly spectral components of the data, with minimal detection. The basic concept involved the Discrete Cosine Transform of the image signal [7]. In [7], it is shown that DCT gives a good approximation, Ψ , to the eigenvectors of the Toeplitz matrices that plays

an important role in signal processing (especially image processing). However, we would perform DCT on the extrema points of the last IMF of the host audio signal, whose coefficients are provided by the following equations:

$$G_x(0) = \frac{\sqrt{2}}{M} \sum_{m=0}^{M-1} X(m) \quad (2)$$

$$G_x(k) = \frac{2}{M} \sum_{m=0}^{M-1} X(m) \cos \frac{(2m+1)k\pi}{2M} \quad k = 1, 2, \dots, M-1 \quad (3)$$

where $X(m)$ is the signal on which the transform is to be performed.

To reduce the risk of decoding the scheme, the sequence is then arranged in descending order with the original sequence saved as a secret key, which is then embedded with the watermarking bits treated in the same manner. In spread spectrum watermarking, a narrow-band signal is transmitted over a much larger bandwidth such that the signal energy presented in any signal frequency is undetectable. Thus, the watermark is spread over many frequency bins so that the energy in one bin is undetectable. Thus, to destroy it, noise of high amplitude is required to be added to all frequency bins. This type of watermarking is robust, since to be confident of eliminating a watermark, the attack must attack all possible frequency bins with modifications of considerable strength.

II. Proposed Watermark Embedding Algorithm

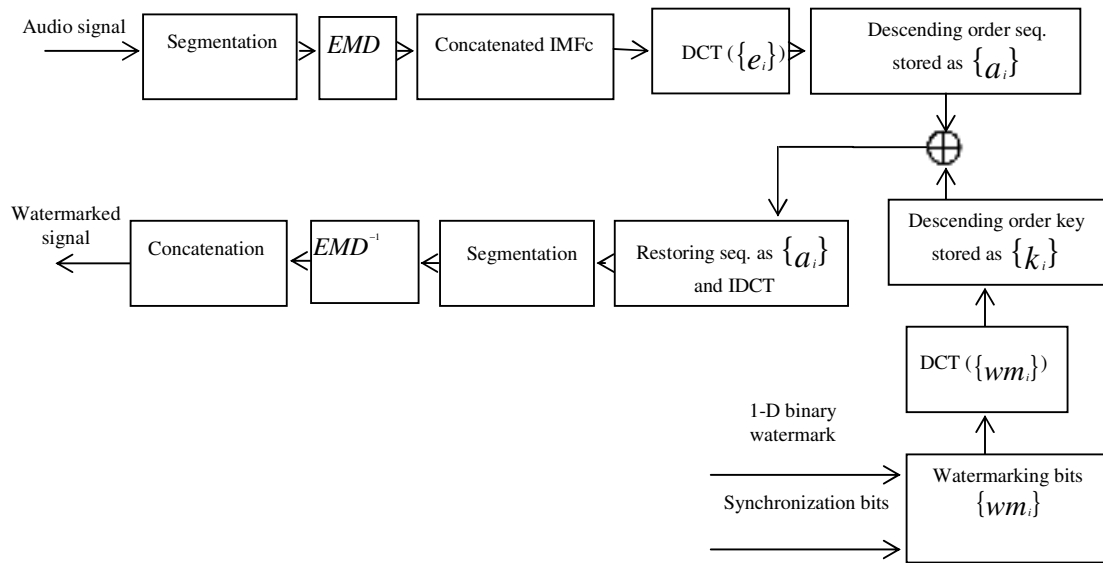


Fig 1. Watermarking embedding algorithm

The basic concept as stated in [9], behind watermarking algorithm is that the image is first converted into its binary form so that it can be represented by a matrix $M \in \{0,1\}$. It is then converted into a 1-D array $A \in \{0,1\}$, which is combined with the synchronization bits to form the watermark array $A \in \{0,1\}$. The synchronization code is a stream of bits $\in \{0,1\}$ chosen, to differentiate between two consecutive message bits. The present sequence of the watermark array is stored before it is arranged in descending order. This key serves as a security confirmation when extracting the watermark. The extrema points of the last IMF are also treated in the same manner before embedding the watermarking bits. The basic steps to embed the watermark, as shown in fig. 1, are explained as follows:

1. The original audio signal is to be divided into segments.
2. The Empirical Mode decomposition should be performed separately on all the segments.
3. Perform DCT on segments of the extrema points of the last IMFs concatenated together as IMFc, of the

size of watermarking bits. Store the sequence as an array $\{a_i\}$ before arranging it into descending order as $\{e_i\}$.

4. The same procedure is carried out on the watermarking bits and the key is stored as an array $\{k_i\}$.
5. Embed times $\{k_i\}$ into maximas in $\{e_i\}$ and/or minimas in $\{e_i\}$ by additive embedding [9]:

$$\{e_i^*\} = e_i + \text{sgn}(S^* k_i) \quad (4)$$

where sgn function denotes '+' if e_i is a maxima, else it represents a '-' operation. S is the embedding strength [10] fixed to an optimum value.

6. The embedded points e_i^* are arranged in the original order using the key $\{a_i\}$ before IDCT (inverse DCT) is performed on e_i^* .
7. Separate the concatenated watermarked IMFc into its respective segments.
8. Reconstruct segments by EMD^{-1} using the modified last IMFs of each segment and concatenate the watermarked segments to retrieve the watermarked signal.

III. The Watermark Extraction Algorithm

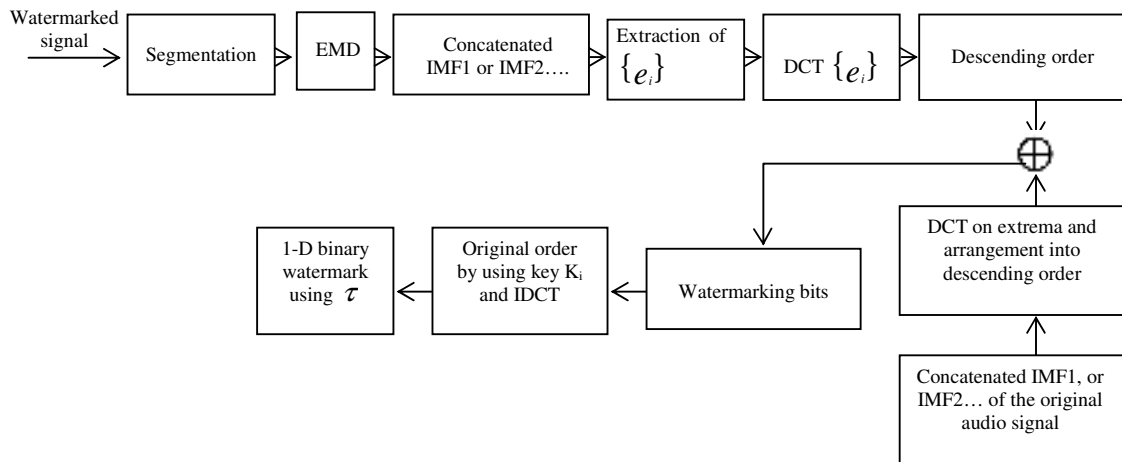


Fig 2 . Watermark extraction algorithm

The proposed extraction process is based on non-blind method, and is carried out on the lower order IMFs. This is done due to the fact that the dependency of the selected frequencies on standard deviation causes the magnitude of the watermarked spikes to move up to the higher frequencies from where it can be successfully extracted. As the deviation reduces as the IMF order increases, the strength of the watermark also reduces. The extraction involves using the secret key to arrange the extracted watermarking bits in the original order. The basic steps involved in the extraction process, as shown in fig. 2, are explained as follows:

1. The watermarked signal is to be divided into segments.
2. IMFs should be obtained by performing EMD on all the segments.
3. The IMF order of all the segments, from which the extrema points $\{e_i^*\}$ need to be extracted, need to be concatenated.

4. Perform DCT on segments $\{e_i^*\}$ of the size of the watermarking bits before arranging it into descending order. Also perform the same on the corresponding IMF of the original audio signal to get $\{e_i\}$.
5. Obtain $\{m_i^*\}$ from $\{e_i^*\}$ by:

$$m_i^* = e_i^* - e_i \quad (5)$$

6. Arrange $\{m_i^*\}$ into the original order $\{m_i^{**}\}$ by using the secret key provided to authorized personnel.
7. Perform IDCT on $\{m_i^{**}\}$ before setting a threshold value τ by:

$$\tau = \frac{\max(m_i^{**}) + \min(m_i^{**})}{2} \quad (6)$$

8. Obtain the binary 1-D array χ_i of the image by:

$$\chi_i = \begin{cases} 1 & \text{if } m_i^{**} \geq \tau \\ 0 & \text{if } m_i^{**} \leq \tau \end{cases} \quad (7)$$

9. Recreate the 2-D binary image by $\{\chi_i\}$.

IV. Experimental Analysis

This Technique has been formulated over a 2 sec wav audio signal sampled at 44.1 khz which gives us 88200 samples to embed the image. Each segment contains 441 samples, which divides the audio signal into 200 segments. The binary image to be watermarked is taken as of $M \times N = 10 \times 20 = 200$ bits (fig. 3). This 2-D image is converted into a 1-D array in order to embed it into the audio signal. The synchronization code is chosen to be [0 1 1 1 1 1 0] based on the concept of byte-stuffing in synchronous transmissions in serial medium.

As stated earlier, spread spectrum allows all energy, corresponding to any one of the frequency bins, to be dissipated across many frequency bins thus making the scheme quite robust. The process performed over the maxima values is been shown in fig. 4.

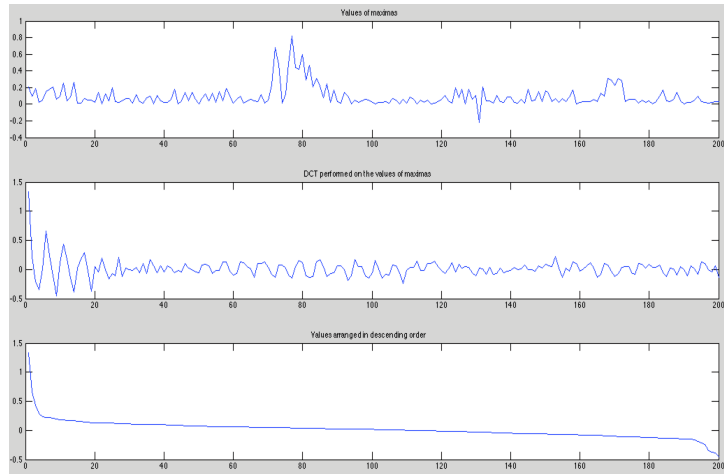


Fig. 4. DCT performed over the maxima values and then arranged into descending order.



Fig 3 . Binary watermark

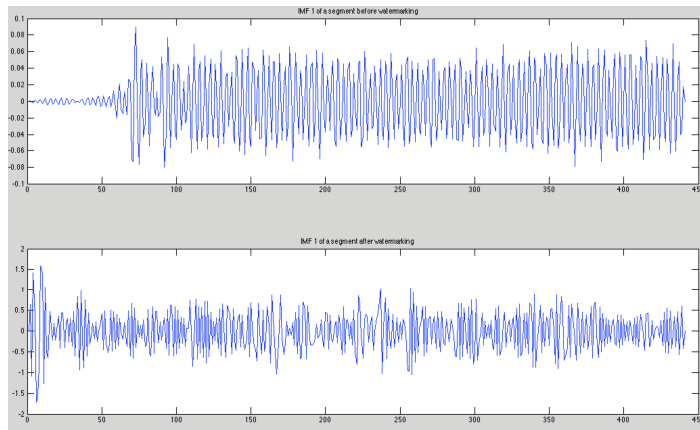


Fig. 5 IMF 1 before and after watermarking one segment

Moreover, the embedded bits can be observed across IMF 1 and IMF 2 clearly as shown in fig. 5 and fig. 6. However, the percentage similarity of the extracted watermark from IMF 3 came out to be only 14.81%, which may be neglected.

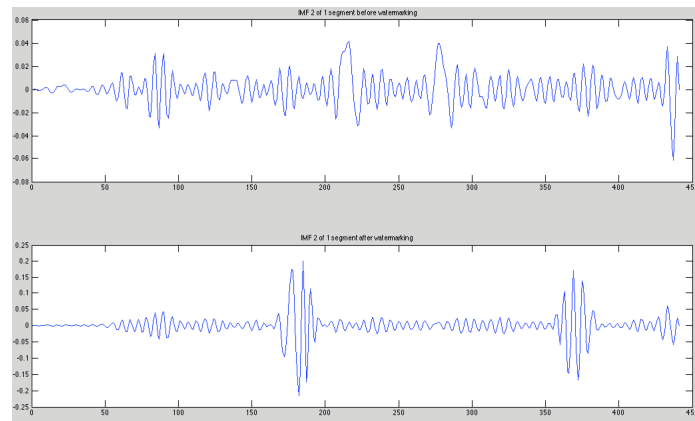


Fig. 6 IMF 2 before and after watermarking one segment

Managing the imperceptibility and the robustness of the scheme, 0.73 came out to be the best possible value of the embedding coefficient that allowed us to extract 100% watermark from both IMF 1 and IMF 2. This value is much less than that proposed in [8]. Also, the value of the embedding strength can be reduced to 0.52 in case of 100% extraction from only IMF 1. The range and percentage similarities of the extracted watermarks are been shown in table no.1.

Table 1: Percentage similarity of the extracted watermark for different values of S

Value of S	IMF order	Percentage similarity	Extracted watermark
0.52	IMF 1	100	
	IMF 2	81	
0.73	IMF 1	100	
	IMF 2	100	

V. Results

To evaluate the robustness of the proposed scheme, different attacks have been performed:

- 1) **Noise** – White Gaussian Noise (WGN) is added to the signal to make the Signal to Noise Ratio (SNR) as 20 dB.
- 2) **Filtering** – The signal is filtered through the low pass as well as a high pass filter.
- 3) **Cropping** – Segments of 441 samples are removed at about 13 locations and are replaced with those segments contaminated with WGN.
- 4) **Compression** – The watermarked signal is compressed and eventually decompressed using MP3 compression at 320 kbps.

Table 2 and 3 show the results of the extraction from IMF 1 and IMF 2 after attacks on the watermarked signal. Given all the attacks, including WGN, detection over 50% is possible.

Table 2: Percentage similarity of the extracted watermark for different attacks from IMF1

Attack type	Percentage similarity	Extracted watermark
No attack	100	
AWGN (20 dB)	55.75	
Cropping	68.19	
High pass filter	72.37	
Wiener filter (10 dB)	91.60	

Table 3: Percentage similarity of the extracted watermark for different attacks from IMF2

Attack type	Percentage similarity	Extracted watermark
No attack	100	
AWGN (20 dB)	100	
Cropping	100	
Low pass filter	84.61	
Wiener filter (10 dB)	75.59	

Detection through multiple IMFs gives us the advantage over the filtering attack. If a particular frequency corresponding to an IMF is masked, the watermark can be successfully extracted from the other frequency bearing the magnitude change. The scheme has even proved to be robust enough against Wiener filter attack.

Moreover, the spread spectrum technique has been proved to be efficient for a range of values of the embedding strength that provides an option to decide an optimum value considering the compromising factor between imperceptibility and robustness. Also, by table no. 3, the scheme has proved that it is robust even in the case of attacks being performed on the lower order IMFs. Table no. 4 gives the experimental results proving the scheme to be robust against the additive white Gaussian noise attack for range of values. The technique extracts the watermark from the signal, adulterated with noise up to 4dB SNR, with 100% percentage similarity. The experimental results for the same are shown in table no. 4.

Table 4: Percentage similarity of the extracted watermark for different values of SNR after AWGN

SNR after AWGN	Percentage similarity	Extracted watermark
3	98.18	
4	100	
5	100	
10	100	
15	100	
20	100	

VI. Conclusion

In this paper, a new scheme to embed the watermark has been proposed. Watermark is embedded in the lowest frequency, and also, the embedding process is through spread spectrum technique thus achieving good performance against various attacks. Extensive simulations have been performed and the scheme has been proved to be robust. However, this scheme did not give positive results for mp3 compression attack. Our future work concerns the same. Moreover, effort is being carried out for embedding the watermark with varying values of embedding strengths, which would allow better inaudibility.

References

- [1] I. J. Cox, J. Kilian, T. Leighton, and T. Shamon, "A secure, robust watermark for multimedia," LNCS, vol. 1174, pp. 185–206, 1996.
- [2] 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.
- [3] Schlurmann, T. "The empirical mode decomposition and the Hilbert Spectra to analyse embedded characteristic oscillations of extreme waves." *Actes de Colloques-IFREMER* 32 (2001).
- [4] Khaldi, Kais, et al. "Audio encoding using Huang and Hilbert transforms." *Communications, Control and Signal Processing (ISCCSP), 2010 4th International Symposium on*. IEEE, 2010.
- [5] Al-Haj, A., L. Bata, and A. Mohammad. "Audio watermarking using wavelets." *Networked Digital Technologies, 2009. NDT'09. First International Conference on*. IEEE, 2009.
- [6] Cox, Ingemar J., et al. "Secure spread spectrum watermarking for multimedia." *Image Processing, IEEE Transactions on* 6.12 (1997): 1673-1687.
- [7] Ahmed, N., T. Natarajan, and K. R. Rao. "Discrete Cosine Transform." *Selected Papers on Visual Communication: Technology and Applications* 13 (1990): 145.
- [8] Kais Khaldi and Abdel-Ouahab Boudraa, "Audio Watermarking Via EMD," *IEEE Trans. Audio Speech Language Process.*, vol. 21, no. 3, pp. 675–680, Mar. 2013.
- [9] Khokhar, Hafiz Malik Ashfaq, Rashid Ansari, and Marco Salvemini. "Blind Detection for Additive Embedding Using Underdetermined ICA." *Multimedia, 2006. ISM'06. Eighth IEEE International Symposium on*. IEEE, 2006.
- [10] Vuckovic, Vesna. "Embedding Strength Criteria for AWGN Watermark, Robust Against Expected Distortion." *Computing and Informatics* 29 (2010): 357-387.