ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

AUTHENTICATION AND SELF-SYNCHRONISATION OF AUDIO SIGNAL BY WATERMARKING USING EMPIRICAL MODE DECCOMPOSITION

Miss. Shruti S Kolekar [*], Mr. K.R.Desai

[*1]PG Scholar, [*2]Asst. Professor, [*1,2] Department of Electronics & telecommunication , Bharati Vidyapeeth's College of Engineering, Kolhapur, India

ABSTRACT

Digital multimedia data is transmitted for numerous real time and non real time applications. In this process undesired intruders may cause the duplication, copying, modification. At this point authentication becomes a zenith issue. To avoid further illegal redistribution and misuse of digital data numerous techniques have been put forth in past decades. The research paper brings to light, number of methods from past years in field of data secrecy. Empirical mode decomposition is a method that provides efficient and resilient transmission of audio signal. Implementation of synchronization sequence causes self-synchronization. This paper brings to notice of the analysts a technique for authentication of such signals based on numerous data evaluation parameters.

KEYWORDS: Self synchronisation, synchronization sequence, Resilience.

INTRODUCTION

Diverse multiple access techniques have made the transmission of digital media very easy. This has given rise to numerous security concerns associated with the data. We consider audio signal particularly, in our context. The signal may be of artistic or intellect importance. The need of authentication came into picture when the fear of illegal distribution, copying and modification was encountered. Steganography and Cryptographic methods are quite superficial measures for these issues. Steganography conceals the secret data from world. Cryptographical methods offer mere encryptions and decryptions and prevent detection of secret data from eavesdropper. Piracy is a threat that can at the times be expensive in terms of huge financial losses and security vulnerability. Watermarking an audio sample can help to prevent these losses by providing security after decryption. Digital watermarking is a promising alternative to above methods that can help prevent detection, duplication and modification of vital data. Digital watermarking is a technique that can provide copyright protection [1]-[4]. Generally a binary watermark is embedded in original host audio signal. Basically watermark should be imperceptible to listener. It must provide sufficient data capacity. It should assure robustness to different distortions offered to the audio signal. Distortions may be attacks like additive noise, MP3 compression, Filtering, cropping and requantization. A robust technique may limit transmission rate. A trade off was to be made in earlier techniques [2]-[4]. Audio watermarking is a technique, particularly dealing with embedding of secret information in host audio signal. Past decades have experienced the rise of interest in digital watermarking [1]. Muzak Corporation first implemented owner identification. The notch filter was used to filter the 1 kHz audio signal. Morse code was used to encode identification information. The major intention was to differentiate their recordings from contemporary similar ones.

Temporal and frequency perceptual masking [2] was developed to guarantee robustness and inaudibility. This technique explored masking in human auditory system. The Self coordinating audio authentication scheme for audio data transmission [3] uses a discrete wavelet approach. This novel technique embeds one synchronization sequence and a part of informative data in DWT coefficients. The scheme observes robustness of embedded data to different

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

manipulations and attack. A combination of singular value decomposition [4] and DWPT (Discrete Wavelet Packet Transformation) can provide better imperceptibility and robustness with satisfactory payload capacity. Quantization Index Modulation [6] is introduced to provide good rate, less distortion and sufficient robustness performance. Their low complexity variation called dither modulation is very advantageous against square error distortion. It uses few scalar quantizers and adders. The information-embedding capacities for the case of a Gaussian host signal and additive colored Gaussian noise attacks are observed. In this scheme a post processing of distortion compensation is encountered. The digital watermarking can be employed in time domain [7]. The method maintains the quality of digital watermark as per Human auditory system. Generally a group of samples is scaled for sample amplitudes. This is employed to save the wave envelope in time domain. There is a statistical method to implement audio watermarking in transform domain [8]. The transform can be DCT, DFT and DWT. Patchwork algorithm is a technique implemented over image which includes pseudo randomly chosen patches. Secret information is inserted in time domain. Time domain is vulnerable to attacks. It was modified to implement over audio signal in frequency domain. The embedding factor is calculated based on sample mean and variance. Use of sign function and Patch size consideration can guarantee desired robustness. Spread spectrum (SS) and QIM are information modulation techniques. SS watermarks [9] are embedded in frequency domain. So they can be audible in blocks that contain silent periods. The robustness of these techniques is examined with respect to watermark estimation attacks. Psycho-acoustic Frequency Masking (PAFM) block repetition coding can enable robustness. Audio authentication can face challenge due to time scale modification (TSM) [10] and cropping. To prevent this we embed watermark in steady local regions with high energy. The watermark can surmount distortions. In the new scheme introduced in [11] the length of interval between salient points of host signal is changed to embed the data. Salient points are peaks of wavelet co efficient envelope. There is quantization of intervals and watermarks are embedded in quantization indices. It is suitable for marketing, broadcast and playback monitoring applications. The scheme offers robustness against MP3 lossy compression, low pass filtering and TSM. EMD and Hilbert Spectral analysis [12] are quite advantageous in non-stationary data analysis. IMFs [13], [14] form the basis of decomposition. They are complete and orthogonal. It allows the tedious complex data to be converted into a version in which we can define instantaneous frequencies

Digital watermarking can be done in spatial and frequency domain. Of these, frequency domain technique has proved to be more reliable and stable. The digital watermark to be embedded is a binary logo image. It can be of any form as desired by the user and depending on application. The watermark that we consider has to satisfy few requirements:

1. Lack of Perception:

The watermark should be inaudible or invisible to the end user. The quality of the host signal should not be hampered

2. Resilience:

The watermarked signal should be resistant to the attacks that it faces in transmission medium. Watermark may encounter different attacks like compression, filtering and other signal processing manipulations.

3. Detection complexity:

The watermark must be capable to retain necessary information on desired user's side. Unambiguous retrieval of copyright information should be accompanied.

4. Invisible for intrusion:

The watermark must be secret and unavailable to the intruder.

5. Embedding capacity:

It gives the rate of embedding the secret information in the host signal.

6. Deliberate blindness:

The watermark should not require reference to the original audio signal.

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

7. Evaluation intricacy:

The process required for embedding the watermark must be optimally complex.

The trade off between the previously mentioned factors is quite essential as increasing one factor can affect the overall watermarking scheme. Watermarking of audio signal that we consider is using empirical mode decomposition [5], [15]. This is a scheme wherein the analysis of the non stationary signals is done on adaptive basis. The system is well useful for the real time signals which have changing nature at different instant of time. The problem of meager transmission bit rate can be treated using the wavelet approach. If such signals are treated with wavelet approach there can requirement of prior choice of filter and filter co efficients. The basis function has to be fixed and the scheme is classical kernel based. Moreover the empirical mode decomposition is dependent on the incoming data. The scheme has no such demerits as above and is reliable for treating the audio signals. The scheme divides the incoming sample into symmetric, orthogonal, zero mean AM-FM envelopes. These envelopes are called as intrinsic mode functions (IMFs) [5]. IMFs exhibit peculiar features: Number of zero crossings equals number of extrema or differs by one. Local maxima and minima at any point differ by zero in mean value at any point.

The sample is divided from smaller fractions to comparatively larger fractions. This is as shown in fig 1. For introducing authentication in audio signals the audio sample is first divided in frames. On applying EMD respective IMFs are obtained. The watermark (WM) bits and synchronization sequence (SS) is introduced in extrema of last IMFs. The watermark (WM) bits and synchronization sequence (SS) is introduced in extrema of last IMFs. As the signals travel through different media they undergo different deviations and losses due to alterations, attenuation and losses. Thus synchronization sequence plays a dominant role in evaluating WM position. This facilitates efficient extraction of desired WM bits.

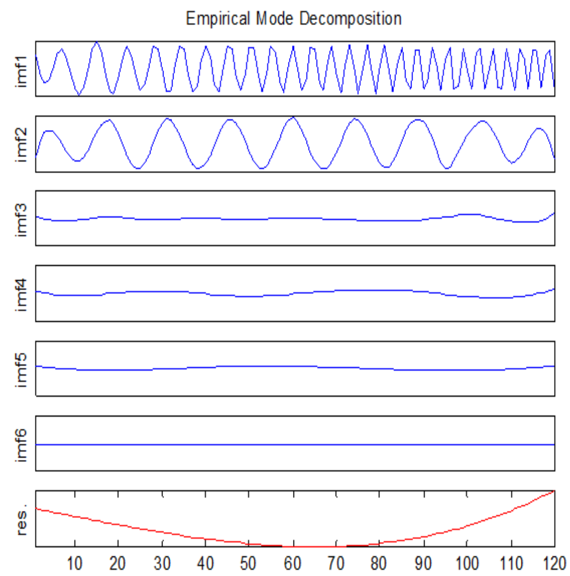

Fig 1: EMD in general

We can represent expansion of a signal in using EMD, in general terms as

$$y(t) = \sum_{i=1}^{k} \text{IMF i } (t) + rk(t) \qquad (1)$$

Where, y(t) – any audio signal

K – IMF count, $r_k(t)$ – final residual

SYSTEM UNDER CONSIDERATION

In the considered framework we first divide an audio signal into frames and apply a special function known as EMD over each frame to obtain IMFs. Watermark together with Synchronization sequence is hidden in extrema of last IMF, in time domain. A bit is inserted in extrema [5]. This can help us assure imperceptibility and robustness. There is no transform of domain. Local extrema can define and recover IMFs [5], [14] fully. Higher orders (low frequency) are perfect places to embed watermark. They are signal dominated and their change may degrade the signal. Quantization Index Modulation (QIM) [6] is a category of watermarking that we have chosen. It makes watermark more robust and blind. QIM parameters are selected such that they can provide inaudible watermark. Time domain coupled with EMD can lower the expense to search Synchronization sequence. The required framework

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

employs following chief modules: Synchronization sequence (SS), embedding and extraction module.

1. Synchronization sequence (SS):

It helps to provide position of embedded watermark. It assists in resisting distortion and cropping attacks. To differentiate synchronization sequence from any unknown sequence the difference between both is calculated. If bit by bit difference is less than or equal to some predefined level [3]. Synchronization sequence and authentication watermark sequence are combined to give a binary sequence m_i as in Figure 2. The desired data sequence structure has binary bit value b_j =0 or 1.

2. Watermark (WM) insertion module:

Audio signal is the input to the system. It is divided into frame segments. Each segment is applied EMD to form IMFs [14].

Sync-sequence	Authentication watermark	Sync-sequence

Fig 2: Data sequence structure b_j

Synchronization sequence and Watermark are combined to give binary sequence b_j. It is embedded E times in extrema of last IMF (IMF_k) [6].

$$xj = \left\lfloor \frac{xj}{s} \right\rfloor .s + \mathrm{sgn}\left(\frac{3s}{4}\right) \text{ if } bj=1$$

$$= \lfloor xj/s \rfloor .s + \mathrm{sgn}\left(\frac{s}{4}\right) \text{ if } bj=0 \qquad (2)$$

Here x_j* and x_j denote extrema of IMF_k of host and watermarked signal. ⌊ ⌋ function to indicate floor of entity. Sgn has positive value if x_j is maxima else it is negative. S is embedding strength. Frame is reconstructed by taking inverse of EMD. Final concatenation of watermarked frames reveals watermarked audio signal.

3. Extraction of Watermark:

During extraction as in fig 4, the input watermarked signal is divided into frame segments. EMD is applied over each division to obtain IMFs correspondingly. This is same as done during embedding. Find extrema {x_j*} of IMF_k. To locate the embedded watermark we find the synchronization sequences in the sequence {bj*} bit by bit. Let z = {bj*} be data to be extracted and v be initially used SS. Determine binary sequence according to a rule in [3]. When the position of Synchronization sequence is found, the hidden information bits can be extracted. From the so far extracted content, we find SS. This process is repeated until Synchronization sequence is found, by shifting the selected window, one sample at time. Sliding window size is decided by considering initial value to 1 and last value to l1. Extracted signal and original SS is compared for correlation. If the difference between sequences taken bit by bit is greater than threshold the count is incremented. The window is further slide by L1 samples.

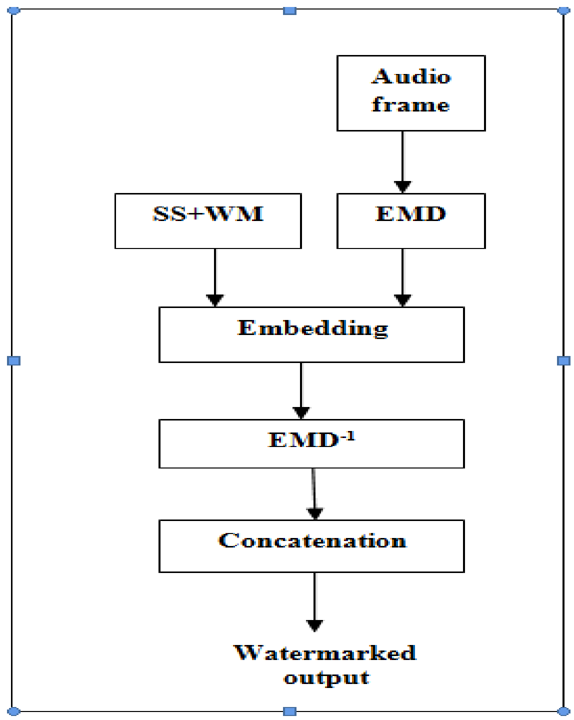

Fig 3: WM insertion module

Again correlation is determined progressively. Extract E watermarks and examine their correlation bit by bit for checking. The original audio signal is not needed during this process.

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

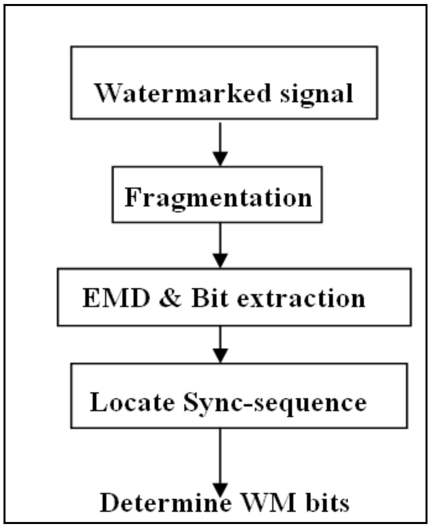

Fig 4: Steps to extract WM bits

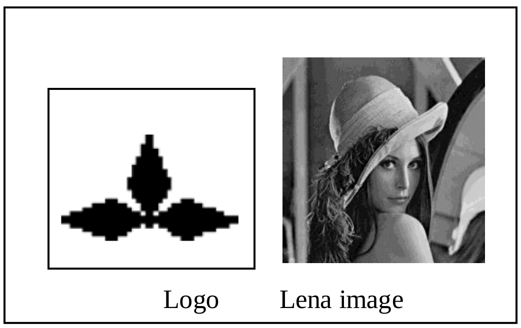

Fig 5: Watermarking images

RESULTS AND DISCUSSION

To analyze different parameters we have considered GUIs using MATLAB software as shown in fig 6, 7 and 8. The input audio signal is sampled at approximately 44 kHz. Frame length in size of 256 samples. Embedding strength is taken as 0.01. If the sample rate is insufficient we can up sample to 44 kHz and further process is carried out. The audio signal is iterated in steps and finally residual is obtained. The final component of imf after watermarking is quite similar to that before watermarking. For finding different data parameters like BER (Bit Error Ratio), SNR (Signal to noise ratio), FPE (False positive error), FNE (False negative error) and NC (Normalized Correlation) [15], following GUIs are implemented.

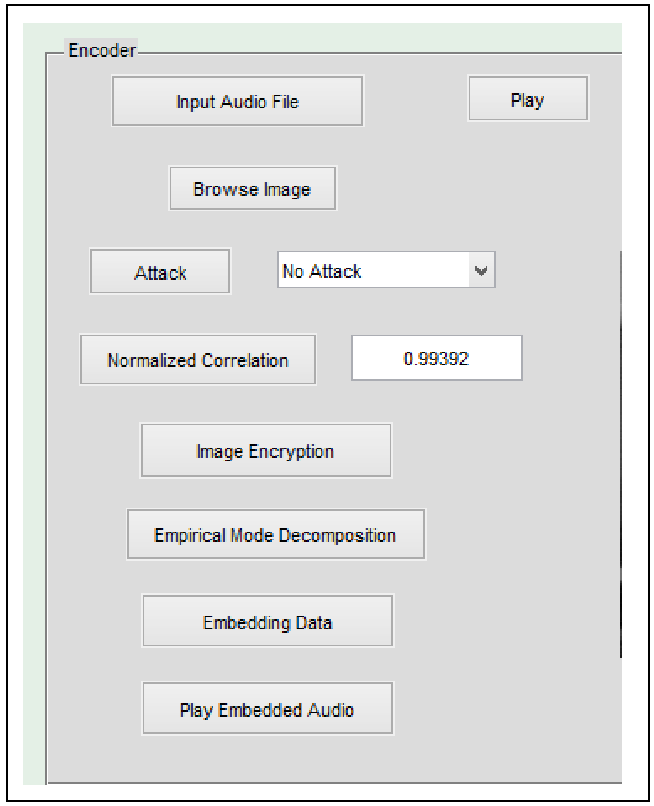

Fig 6: GUI section at encoding side

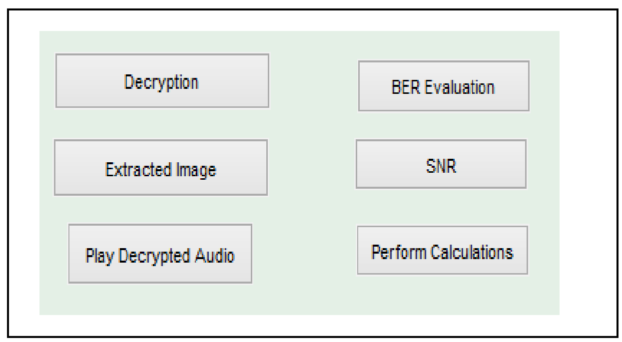

Fig 7: GUI section for decryption

From experimental analysis we obtain following values for two different types of parameters as in table 1. BER can be defined as a measure to analyze transmission errors in bits. SNR of the output determines the amount of desired signal. The method considered in this context provides SNR up to 19. FPE and FNE are the types of errors encounters while evaluating position of Sync-sequence. FPE is supposed to be occurred if a fake sync sequence is detected.

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

Table no 1: Parameters obtained

Parameters	Values obtained
BER	1.646
SNR	17.90
FPE	0.9
FNE	0.04
NC	0.99

FNE arises when required sync sequence is lost. In this error the part of desired WM bits may be lost. Different attacks are applied to the audio signal and different effects are examined for robustness analysis. Each type of attack affects the signal accordingly. But despite of attack the system continues to provide security to audio data.

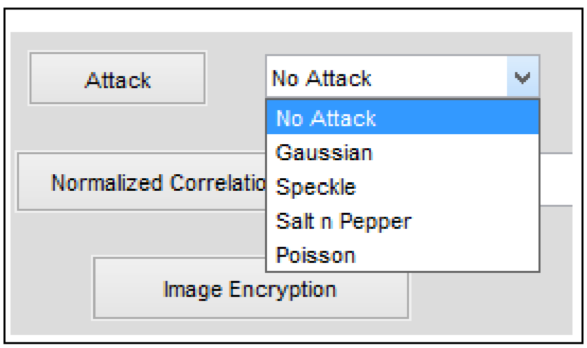

Fig 8: GUI showing types of attacks

Fig 9 and 10 show that there are few losses incurred in this considered process. We cannot guarantee the process to be full proof against attacks. The different attacks applied have diverse effects on the output.

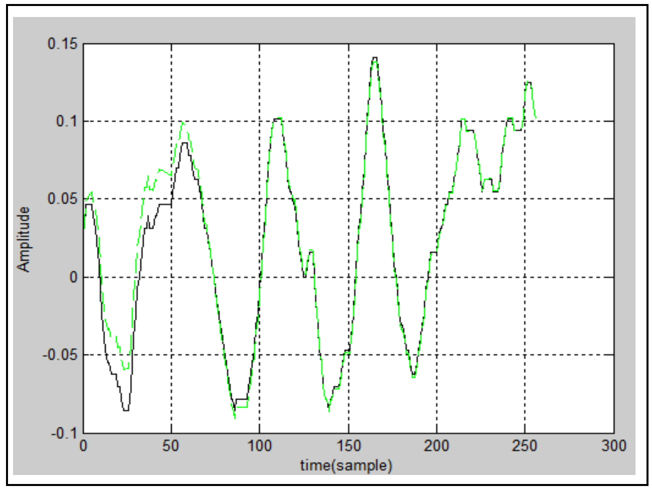

Fig 9: Final imf before (black)and after (green) authentication

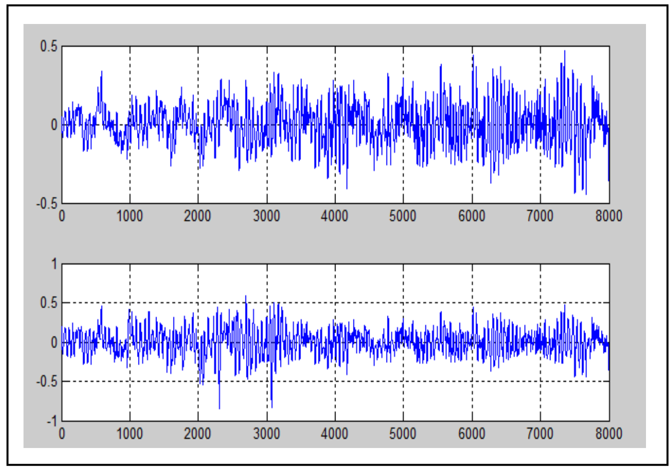

Fig 10: Audio signal before (first) and after (second) authentication

CONCLUSION AND FUTURE SCOPE

In this paper we come across novel technique of audio signal authentication. The method shows that watermarking does not cause any declination in overall quality of original host signal. Every technique has to be novel from previous counterparts. No technique is foolproof to attacks. There may be lacuna in one or the other factor. If we suppose a technique to provide utmost robustness it may hamper data payload capacity and vice a versa. Each basic requirement has to be addressed so carefully that other factors will not be affected adversely. As seen through results empirical mode decomposition seems to be a quite reliable method that can help balance the basic requirements of watermarking.

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

In coming future the novel techniques need to investigate use of variety of sampling frequency that can support same data payload. The method to employ different sound versions like jazz, pop, classic, rock etc for same sound sample needs to be assuring minimum bit error rate and other random signal properties. The embedding strength needs to be adapted as per incoming signal. The audio signal is analog by default so while further signal processing it needs to be digitized. It must be crucially investigated that whether this conversion has any effect on final output. The losses incurred need to be investigated in magnitude and type. The channel of transmission is not perfect so the technique needs to be investigated which can give the analysts the type and intensity of losses. We generally consider an image as watermark but other forms of watermarks can also be investigated to make overall process blind and unpredictable to the intruders and hence guaranting desired secrecy.

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. Tewfik, "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] Hwai-Tsu Hu, Hsien-Hsin Chou, Chu Yu and Ling-Yuan Hsu, "Incorporation of perceptually adaptive QIM with singular value decomposition for blind audio watermarking," Hu et al. EURASIP Journal on Advances in Signal Processing , 2014

[5] 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.

[6] 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.

[7] 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.

[8] I.-K. Yeo and H. J. Kim, "Modified patchwork algorithm: A novel audio watermarking scheme," IEEE Trans. Speech Audio Process., vol. 11, no. 4, pp. 381–386, Jul. 2003.

[9] D. Kiroveski and H. S. Malvar, "Spread-spectrum watermarking of audio signals," IEEE Trans. Signal Process. , vol. 51, no. 4, pp. 1020–1033, Apr. 2003.

[10] 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.

[11] M. F.Mansour and A. H. Tewfik, "Data embedding in audio using time-scale modification," IEEE Trans. Speech Audio Process., vol. 13, no. 3, pp. 432–440, May 2005.

[12] 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.

[13] N. Cvejic and T. Seppanen, "Spread spectrum audio watermarking using frequency hopping and attack characterization," Signal Process., vol. 84, no. 1, pp. 207–213, 2004.

[14] Sangjin Cho, and Yejin Seo, "Extraction of Significant Features using Empirical Mode Decomposition and Its Application," in Proc. WCECS, vol. 1, 2013

[15] Kais Khaldi and Abdel- Ouahab Boudraa, "Audio watermarking via EMD", IEEE Trans. Audio, Speech, and language processing, VOL. 21, NO. 3, MARCH 2013.