ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 6, June 2015

Audio Watermarking using Colour Image Based on EMD and DCT

Suhail Yoosuf[1], Ann Mary Alex[2]

P. G. Scholar, Department of Electronics and Communication, Mar Baselios College of Engineering and Technology, Trivandrum, India [1]

Faculty, Department of Electronics and Communication, Mar Baselios College of Engineering and Technology, Trivandrum, India [2]

Abstract: Watermarking is the process of embedding information into a signal (e.g. audio, video or pictures) in a way that is difficult to remove. If the signal is copied, then the information is also carried in the copy. An audio watermark is a unique electronic identifier embedded in an audio signal, typically used to identify ownership of copyright. Audio watermarking is done using Empirical Mode Decomposition (EMD) where the audio signal is divided into segments and then decomposed into different IMFs (Intrinsic Mode Functions).The watermark is then applied to the last IMF as binary bits using QIM (Quantization Index Modulation). The watermarked colour image is obtained using DCT (Discrete Cosine Transform) compression method and converted to binary bits. In the watermark extraction phase the watermark is obtained as binary bits and converted back into image using IDCT (Inverse Discrete Cosine Transform) technique. The watermark is embedded into the IMF preserving audio perceptual quality of the host signal. By varying watermarking intensity the robustness of the hidden watermark for additive noise and MP3 compression is enhanced and the payload of the watermarked audio signal is increased.

Keywords: Empirical mode decomposition, intrinsic mode function, audio watermarking, image compression, discrete cosine transform.

I. INTRODUCTION

Watermarking is the method of hiding information bits into a host signal. The information bits or data can be used for identifying the owner, for copyright protection, authentication, copy control etc. The host signal can be audio, video or image. The copyright protection of digital media is achieved by embedding a watermark in the original host signal. If the host signal is an audio signal then the watermarking is audio watermarking. The requirements of digital audio watermarking are imperceptibility, robustness and security. Imperceptibility means the watermark in the audio signal is inaudible to the listener. Robustness refers to the changes in the watermark image under various conditions. Security means the watermark cannot be extracted or deleted without the permission of the owner.

There are different methods for audio watermarking. The methods are broadly classified into temporal and spectral watermarking. In temporal watermarking the watermarked data are embedded directly into the host audio signal in the time domain. In spectral watermarking certain frequency transforms are performed to the host audio signal and then embed the watermark info into the transformed frequency domain data block.

In temporal watermarking many techniques are adopted such as watermarking in the dual channel audio using echo hiding scheme [1], echo hiding using analysis by synthesis method [2], using time scale modification method [3], using EMD (Empirical Mode Decomposition) method [4].

In spectral watermarking many techniques are used which involves the transformation of the host signal such as DWT [5], DCT [6]. Both temporal and spectral techniques can be combined together for better results as in [7] which use both DWT and SVD for audio watermarking. For achieving a good tradeoff between reliability and robustness while achieving better payload a new method using the decomposition property EMD and compression property of DCT is proposed.

The limit of DWT, DCT etc. are they use fixed basis functions which do not match all real signals. So to overcome this problem we use a new decomposition method called as Empirical Mode Decomposition (EMD) which decomposes the signal adaptively according to the value of signals. So there is no apriori basis function for the EMD and is completely data driven. EMD recursively breaks down signals into nearly zero mean symmetric envelops called as Intrinsic Mode Functions (IMFs). The EMD decomposition can be expressed as

$$x(t) = \sum_{j=1}^{N} IMF_j(t) + r_N(t) \qquad (1)$$

where N is the number of IMFs and $r_n(t)$ is the final residual.

The IMFs obtained are nearly orthogonal to each other. The EMD is performed in a finite number of modes and the number of extrema decreases when we move from one mode to other. The watermark is embedded in the extrema of the last IMF as in [4].

In the proposed method a colour image is used as the watermark. The colour image is decomposed into its RGB components and then DCT transformation [8] is applied.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 6, June 2015

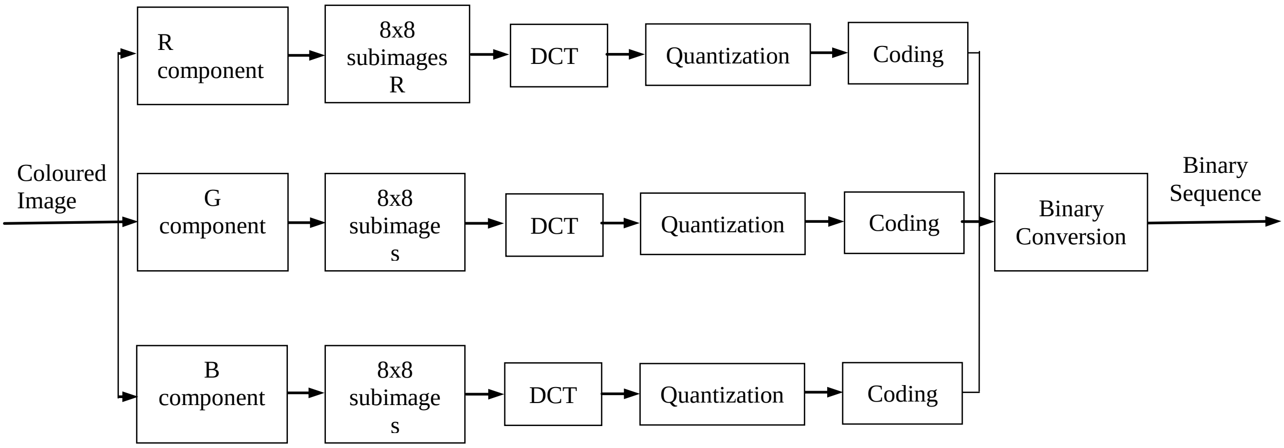

Fig.1. Watermark Preparation

By exploiting the energy compaction property of the DCT [9] the image is compressed. Then the compressed image is converted into binary bits and is watermarked into the audio signal. The reverse process is performed to obtain the watermarked colour image from the watermarked audio signal.

II. PROPOSED METHOD

The proposed method uses a binary data as a watermark to embed in an audio signal. The binary data is obtained from a coloured image First the audio signal is divided into frames and EMD [10] is calculated on each and every frame. The last IMF obtained from the EMD and the watermark bits are embedded in it. The extrema of the last IMF is calculated and the watermark bits are embedded using Quantization Index Modulation (QIM) [11].So a bit is inserted at each extrema of the last IMF. Then the inverse EMD is performed and the frames are obtained. Finally the frames are concatenated to obtain the final watermark embedded signal. The binary bits which are to be embedded are obtained from a compressed colour image. The image is compressed using the energy compaction property of the DCT. A suitable compression method is chosen so that it gives a good PSNR value. Then it is converted into binary format and the bits for watermark are obtained.

For extracting the watermarked data from the audio the watermarked audio signal is subtracted from the original audio and the QIM coefficients are extracted. The binary bits are demodulated from the QIM coefficients. The binary bits are then converted into the DCT coefficients and inverse DCT transformation is taken. Now the coloured image is obtained. Overview of the proposed audio watermarking technique is detailed as follows:

A. Watermark Preparation

Before embedding into the audio signal the watermark data should be converted into a binary sequence. Here the watermark is an image, so we have to convert the image into a binary sequence. Basics of watermark preparation are shown in Fig. 1. and detailed as follows:

Step 1: Decompose the coloured image into RGB components.

Step 2: Each component image is taken separately and is divided into 8x8 blocks of sub images.

Step 3: DCT is calculated on each sub block using the formula to obtain the DCT coefficients.

$$D(i,j) = \frac{1}{\sqrt{2N}} C(i)C(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} p(x,y) \cos\left[\frac{(2x+1)i\pi}{2N}\right] \cos\left[\frac{(2x+1)j\pi}{2N}\right]$$

$$where\ C(u) = \begin{cases} \frac{1}{\sqrt{2}} & if\ u = 0 \\ 1 & if\ u > 0 \end{cases} \qquad (2)$$

Step 4: The DCT coefficients are quantized using a quantization matrix. Quantization is achieved by dividing each element of DCT transformed matrix by the element in the quantization matrix and rounding it to nearest integer.

$$C_{ij} = round\left(\frac{D_{ij}}{Q_{ij}}\right) \qquad (3)$$

where D_{ij} is the transformed image matrix and Q_{ij} is the quantization matrix.

Step 5: Next step is coding. After quantization most of the coefficients become zero. Using the energy compaction property of the DCT the most relevant coefficients are selected by a zigzag manner as shown in Fig. 2. According to the compression ratio needed we select the number of coefficients and discard the other ones.

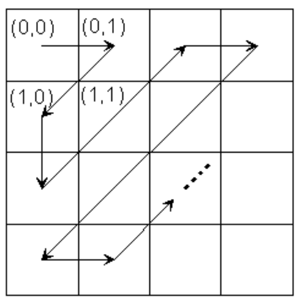

Fig.2. Zigzag Coding

Step 6: The selected coefficients are converted into binary bits to obtain a binary sequence.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 6, June 2015

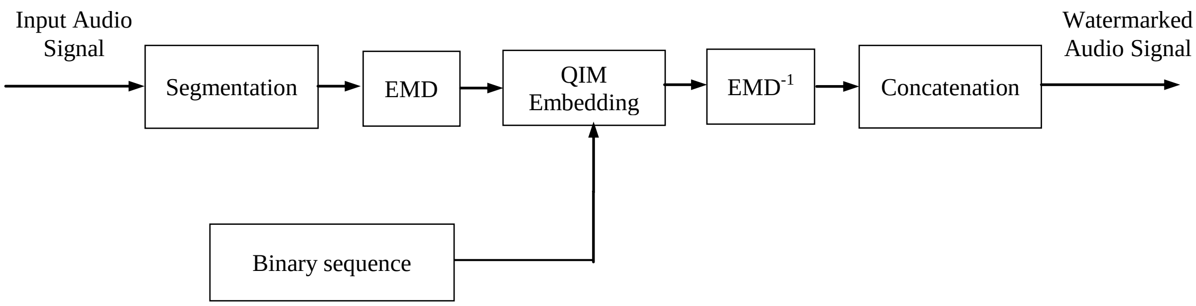

Fig.3. Watermark Embedding

B. Watermark Embedding

The binary sequence obtained from the watermark preparation phase used for embedding in the host audio signal.

Basics of watermark embedding are as shown in Fig. 3. and can be detailed as:

Step 1: Divide the original host audio signal into frames.

Step 2: Decompose each frames into several IMFs using EMD.

Step 3: Embed the binary sequence into the extrema of the last IMF by QIM

$$e_i^* = \begin{cases} e_i + sgn\left(\frac{3S}{4}\right) & if\ b_i = \\ e_i + sgn\left(\frac{S}{4}\right) & if\ b_i = \upsilon \end{cases} \quad (4)$$

where e_i^* is the modified extrema of the last IMF and e_i is the extrema of last IMF of host signal. *sgn* function is equal to '+' if e_i is a maxima and '-' if e_i is a minima. The embedding strength is denoted as *S* and the value of *S* is chosen so that the watermark is inaudible.

Step 4: Perform the inverse EMD (EMD^{-1}) on the IMFs with modified last IMF to obtain the frames.

Step 5: Concatenate the watermarked frames to obtain the watermarked audio signal.

C. Watermark Extraction

The binary sequence is extracted from the watermarked audio signal. As the watermark is embedded as QIM coefficients in the audio, we have to extract the QIM coefficients and decode it to binary bits.

Basic steps involved in watermark extraction are shown in Fig. 4. and detailed as:

Step 1: Subtract the watermarked audio signal from the original audio signal to obtain the QIM coefficients.

Step 2: The binary bits are obtained from the QIM coefficients.

Step 3: Linearize the binary bits to form a binary sequence which is the watermark.

D. Watermark Reconstruction

The watermarked colour image is reconstructed from the extracted binary sequence. The binary sequence contains the information of R, G and B components of the image.

The colour image watermark is recovered from the R G B components Basics of watermark reconstruction is shown in Fig. 5. and detailed as follows:

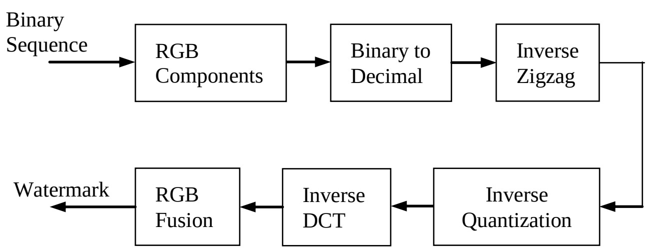

Fig.5. Watermark Reconstruction

Step 1: The binary sequence is first divided into three corresponding R G B components.

Step 2: For each component the binary sequence is converted back into its decimal form.

Step 3: The coefficients obtained are arranged in an inverse zigzag form, after padded with zeros corresponding to the compression ratio, to obtain 8x8 transformed sub images.

Step 4: To obtain the DCT coefficients multiply each 8x8 transformed sub image with the corresponding element in the quantization matrix.

Step 5: Inverse DCT transformation is applied and sub images are obtained.

Step 6: The sub images are joined to get the R, G, B components and further they are fused together to obtain the final watermarked image.

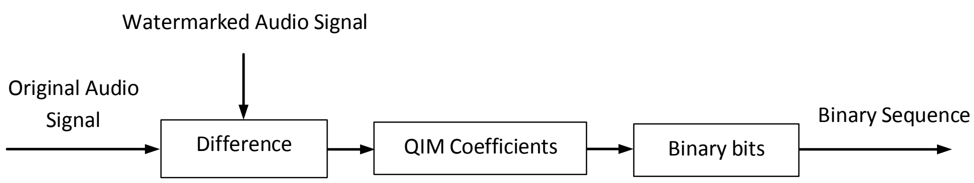

Fig.4. Watermark Extraction

DOI 10.17148/IJARCCE.2015.4679

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 6, June 2015

III.PERFORMANCE ANALYSIS

The image is compressed using the energy compaction property of the DCT (Discrete Cosine Transform). The compression quality is checked by using the PSNR (Peak Signal to Noise Ratio) between the original image and the compressed image. The value of PSNR more than 30 is considered good.

To calculate the PSNR first we have to calculate the mean squared error given as

$$MSE = \frac{\sum_{M,N}[U(m,n) - V(m,n)]^2}{M \times N} \quad (5)$$

where M and N are the row and column length of the image respectively and U is the input image and V is the compressed image. Now the PSNR is calculated as

$$PSNR = 10 \log_{10} \frac{R^2}{MSE} \quad (6)$$

where R is the relative data redundancy. The PSNR value changes for different compression ratio and we should select the compression ratio according to the PSNR we needed. There is a tradeoff between PSNR and compression ratio as the compression ratio decreases PSNR value increases and the quality of the image increase, also it will increase the number of bits used for watermarking.

The performance of our method is evaluated in terms of SNR (Signal to noise Ratio) between the original audio signal and the watermarked audio signal, Bit error rate (BER), Normalized Cross-Correlation and data payload.

As per the IFPI (International Federation of the Photographic Industry) rules, a watermark should maintain more than 20 dB SNR.

Bit error rate (BER) can be used to evaluate the watermark detection accuracy after attacks.

$$BER = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} U(i,j) \oplus V(i,j)}{M \times N} \quad (7)$$

where U is the original watermark and V is the recovered watermark. M and N are the row and column length of the watermark respectively.

The similarity between the original watermark and the extracted one can be found out using the NC measure explained as:

$$NC = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} U(i,j)V(i,j)}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N} U^2(i,j)}\ \sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N} V^2(i,j)}} \quad (8)$$

Payload refers to the amount of data hidden in the host audio signal. The number of bits which is embedded into the host audio signal within a unit of time is referred as payload and its unit is bits per second (b/s).

IV. RESULTS AND DISCUSSION

The performance of our experiment is evaluated by taking the host audio signal sampled at 44.1 kHz. The embedded watermark is 'onion.png' which is a coloured image. For embedding, the coloured image is compressed and converted into binary sequence. The audio signal is divided into frames of size 256 and the embedding strength S is taken as 0.05. Each frame is then decomposed by EMD as shown in the Fig. 6.

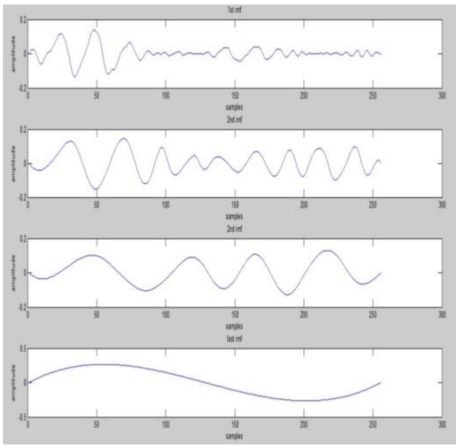

Fig.6. EMD decomposition

The image is compressed using DCT and for different compression ratio values is calculated and is as shown in Table I. The SNR of audio before and after watermark is calculated to find out the imperceptibility quality of audio watermark. If the value of SNR is greater, then the imperceptibility is assured. For different values of embedding strength different values of SNR are obtained. Table II shows the SNR values for the watermarked audio with the original audio for different embedding strengths. To check the robustness of the watermark under various attacks we perform the robustness test. The watermarked audio signal is exposed to different types of attacks such as AWGN (Additive White Gaussian Noise), MP3 compression. White Gaussian Noise is added to the watermarked signal with various intensities and watermark is extracted.

TABLE I : COMPARISON OF PSNR FOR DIFFERENT COMPRESSION RATIO

Compression Ratio	PSNR	Compressed image
1	n.a	
2	35.2010	
4	33.0787	

 DOI 10.17148/IJARCCE.2015.4679

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 6, June 2015

5.33	31.7368	
8	29.7003	
10.66	28.7835	

TABLE II: SNR VALUES FOR DIFFERENT EMBEDDING STRENGTH

Embedding Strength (S)	**SNR**
0.005	55.0481
0.05	35.0555
0.5	15.0568
1	9.0509

Performed MP3 compression on the audio watermarked signal and after decompression the watermark is extracted and quality is checked. Table III reports the results of the tests for robustness.

The robustness is mainly due to the fact that EMD decomposition adapts to audio signal even if its perceptual parameters vary.

TABLE III: RESULTS OF ROBUSTNESS TESTS

Attack type	BER %	NC	Watermark
No attack	0	1	
AWGN (25dB)	0.2749	0.8026	
AWGN (28dB)	0.0247	0.9743	
AWGN (30dB)	0	1	
MP3 compression	0.0079	.9930	

The payload obtained in the proposed method is greater while compared to previous methods. The payload obtained while considering compression is 84.83 bits per second and without considering compression is 61.197 bits per second.

V. CONCLUSION

In this paper audio watermarking using EMD and DCT is proposed. The data driven decomposition property of EMD and energy compaction property of DCT is utilized. This method offers consistent performance for all types of audio as the EMD decomposition is not pre deterministic and only based on the data values. The performance is evaluated based on imperceptibility, robustness and payload. The proposed method significantly improved the payload in comparison to that obtained by some existing method. The watermark is robust under attacks such as AWGN, MP3 compression. As the SNR is significantly high between the watermarked and original audio, the audio is not audible to the listener.

ACKNOWLEDGMENT

The authors acknowledge the support of the Mar Baselios College of Engineering And Technology in this Research.

REFERENCES

[1] Yong Xiang, Iynkaran Natgunanathan, Dezhong Peng, Member, IEEE, Wanlei Zhou, Senior Member, IEEE, and Shui Yu, Member, IEEE, *"A Dual-Channel Time-Spread Echo Method for Audio Watermarking", IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, April 2012.g,

[2] O.T.-C. Chen and W.-C. Wu, *"Highly robust, secure, and perceptual quality echo hiding scheme," IEEE Trans. Audio, Speech, Language Process.*, vol. 16, no. 3, pp. 629–638.

[3] 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.

[4] Kais Khaldi and Abdel-Ouahab Boudraa, *" Audio Watermarking Via EMD",in IEEE Transactions on audio, speech, and language processing*, vol. 21, no. 3, March 2013,pp. 675-680..

[5] Hemam A. Alshammas, *"Robust audio watermarking based on dynamic DWT with error correction," IEEE Conference. ITU Kaleidoscope:Building Sustainable Communities(K-2013)*, vol. 3, no. 2, pp. 1 - 6, 2013.

[6] Ai Hong Yan, Liu Quan, and Jiang Xue-mei, *"Synchronization Audio Watermarking Algorithm based on DCT and DWT," in Conference Anthology,IEEE*,2013, pp. 1-4.

[7] Ali Al-Haj, Christina Twal, and Ahmad Mohammad, *"Hybrid DWT-SVD audio watermarking," IEEE Conference. Digital Information Management(ICDIM)*, pp.525–529, 2010.

[8] Ken Cabeen and Peter Gent,*"Image Compression and the Discrete Cosine Transform"*, Math 45,College of the Redwoods.

[9] Mr. Ramin Samadani, "*Characterizing and Estimating Block DCT Image Compression Quantization Parameters*", *Asilomar Conference on Signals, Systems, and Computers*, 30 October - 2 November 2005, Pacific Grove, CA, USA © Copyright 2005 IEEE.

[10] 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.

[11] 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