

Audio watermarking: A Novel approach using hybridization of DWT+DCT

Mrunali J. Patel¹, Dhavalsinh V. Solanki², Dr. K. R. Borisagar³

Post Graduate Fellow, Atmiya Institute of Technology & Science, ECD, Rajkot, Gujarat, India^{1,2}

Guide & Associate Professor, Atmiya Institute of Technology & Science, ECD, Rajkot, Gujarat, India³

Abstract: First of all, we are in the era of digital multimedia. Because Images, Videos & Audios are the global information sources now. Intellectual property protection & copyright protection in digital multimedia is the core driving force behind this research area. Also watermark can be some verification or hidden messages. Exceptionally successful watermarking innovations have been executed for images and videos. Though, very few algorithms have been proposed for audio behind audio watermarking. This is because of the way that, the human audio framework is a great deal more unpredictable and touchy than visual framework. In this paper we have designed novel and efficient approach based on hybridization of two existing techniques DCT & DWT. Critical performance evaluation with different attacks & different types of audios has been done to demonstrate superiority & robustness of the technique. Also we have compared it with traditional techniques DCT and DWT.

Keywords: DCT, DWT, Robustness, Watermarking

I. INTRODUCTION

Due contemporary technology, there are broadly available tools in order to reproduce and retransmit multimedia data. The rapid growth of the internet has greatly lead to the unauthorized distribution and hacking of digital media. The hacking of the digital systems is easier due to the availability of several processing platforms. As a result, the music industry suffers a multibillion dollar annual revenue loss due to piracy [1]. Thus a technique is needed for the security and protection of the digital data. Watermarking is a descendent of a technique known as steganography, which has been in existence since 1490. Steganography is a strategy which is utilized for hid communication. As opposed to cryptography where the substance of the message is a secret, in steganography the very presence of the message is a secret and just parties included in the communication know its vicinity. Steganography is a technique where a secret message is hidden within another unrelated message and then communicated to the other party [2]. The secret communication between the transmitter and the receiver plays a major role in militaries, radar and wireless communication [3].

Watermarking provides secrecy to the communication between them. Watermarking can be considered as an uncommon method of steganography where one message is implanted in an alternate and the two messages are identified with one another somehow.

Watermarking is characterized as the procedure of attaching the identity of the owner of the copyright of an imaginative work (host media) in such a way that it should be imperceptible (inaudible if there should be an occurrence of audio and imperceptible in the event of videos and images) Watermark should statically be irremovable yet effectively removed by owner.

Watermarking should be strong enough to subsidiary and calculated attacks. To ensure the power of audio watermarking there must be exchange off between measure of watermark embedding and imperceptibility that means inaudibility. Research in audio watermarking is not as much as in image and video watermarking. The fact is human auditory system is much more sensitive than human visual system that is why it is difficult to achieve inaudibility than invisibility so less amount of watermark is embedded in audio than image or video to ensure inaudibility. Evaluation of imperceptibility is usually based on objective measure of quality, called peak signal-to noise ratio (PSNR) [4].

Audio watermarking techniques reported in literature can be grouped into two types; spatial-domain techniques and frequency-transform domain techniques. The two areas have dissimilar attributes, and accordingly performance of their systems may change as robustness and imperceptibility (inaudibility) necessities of audio watermarking. Inaudibility refers to the condition that the embedded watermark should not produce audible distortion to the sound quality of the original audio, in such a way that the watermarked marked version of the file is indistinguishable from the original one. Robustness determines the resistance of the watermark against removal or degradation. The watermark should survive malicious attacks such as random cropping and noise adding, and its removal should be impossible without perceptible signal alterations [5].

Spatial-domain techniques include the Least Significant Bit substitution (LSB) and echo hiding techniques, among many others [6]. But in general, spatial-domain audio watermarking is relatively easy to implement, and requires few computing resources, however, it is weak against signal processing attacks such as compression and filtering.

Frequency domain audio watermarking techniques employ human perceptual properties and frequency masking characteristics of the human auditory system for effective watermarking. In these techniques, the phase and amplitude of the transform domain coefficients are modified in a certain way to carry the desired watermark information. Popular transforms include the Discrete Fourier Transform (DFT), the Discrete Cosine Transform (DCT), and the Discrete Wavelets Transform (DWT). In, the Fourier transform magnitude coefficients over the frequency range from 2.4 KHz to 6.4 KHz are replaced with the watermark sequence since human sensitivity declines compared to its peak around 1 KHz . Moreover, human ears are relatively insensitive to phase distortion, and especially lack the ability to perceive the absolute phase value, therefore in, the watermark is represented by the relative phase between selected coefficients and their neighbours. The problem with these watermarking schemes that they are less robust to signal processing and malicious attacks, such as audio compression. so that the DFT is not used.

Lalitha et. al. [7] evaluated the performance of the Discrete Cosine Transform (DCT) and SVD based audio watermarking and compared it with the performance of DWT-SVD based audio watermarking. The comparison led to the conclusion that the DWT-SVD had shown better outcomes than the DCT-SVD techniques. Al-Haj et. al. [5] proposed a DWT based technique for audio signals watermarking which has the least effect on the original signal.

Inspiring from the literature mentioned above, a novel audio watermarking algorithm based on DWT (Digital Wavelet Transform) and DCT (Discrete Cosine Transform) for audio signal is proposed in this paper. The rest of paper is organized as follows: Section 2 introduces the DWT and DCT. Section 3 discusses our proposed embedding and extraction algorithm in detail. Section 4 gives the simulated experimental results and Section 5 concludes the paper.

II. TRADITIONAL ALGORITHMS

a. DISCRETE WAVELET TRANSFORM

The wavelet transform is a technique for analysing signals [8]. Signal can be analysed both in time and frequency at the same time by using Wavelet transforms [9]. The basic idea of DWT is to separate frequency detail, which is multi-resolution decomposition. One time of decomposition can divide main image to four sub graph as the size of a quarter. They are a low frequency approximate sub graph and three horizontal vertical and diagonal direction high frequency details sub graphs. Figure1 as shown below is a schematic diagram of four-level DWT decomposition. In the wavelet transform domain, high frequency parts represent detailed information of image's edge, contour and texture and so on. Embedding watermarking in these places cannot be easily detected as people are not easily conscious of it. But after processing or attacking, it has poor stability. Most

energy of audio is centralized in low frequency. Low frequency coefficients are nearly unchanged to common attack so that watermarking information embedded in low frequency coefficients has better robustness [10].

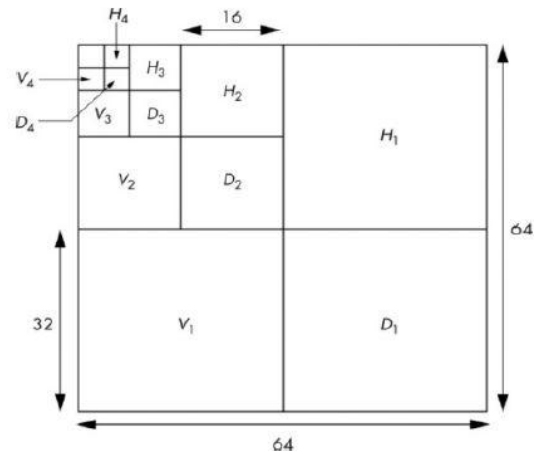


Fig 1 Four level DWT

The DWT is defined by the following equation

$$W(j, k) = \sum_j \sum_k x(k) 2^{-\frac{j}{2}} \Psi(2^{-j} n - k) \quad (1)$$

Where $\Psi(t)$ is a time function with finite energy and fast Decay called the mother wavelet [5].

b. DISCRETE COSINE TRANSFORM

The discrete Cosine Transform is a technique for Converting a signal into elementary frequency components. The most common DCT definition of a 1-D sequence of Length N is

$$C(u) = \alpha(u) \sum_{x=0}^{N-1} f(x) \cos \left[\frac{\pi(2x+1)u}{2N} \right] \quad (2)$$

For $u=0,1,2, \dots, N-1$. Similarly, the inverse transform is defined as,

$$f(x) = \sum_{u=0}^{N-1} \alpha(u) c(u) \cos \left[\frac{\pi(2x+1)u}{2N} \right] \quad (3)$$

For $x=0, 1, 2 \dots N-1$. In both equations (2) and (3) $\alpha(u)$ is defined as,

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{N}} & \text{for } u = 0 \\ \sqrt{\frac{2}{N}} & \text{for } u \neq 0 \end{cases}$$

It is clear from eq. (2) that for $u=0$,

$$C(0) = \sqrt{\frac{1}{N}} \sum_{x=0}^{N-1} f(x)$$

The first transform coefficient is, referred to as DC Coefficient, the average value of the sample sequence. The other transform coefficients are called the AC Coefficients [11].

In particular, a DCT is a Fourier-related transform similar to the Discrete Fourier Transform (DFT), but using only real numbers. DCTs are equivalent to DFTs of roughly twice the length, operating on real data with even symmetry (since the Fourier transform of a real and even function is real and even), where in some variants the input and/or output data are shifted by half a sample [7].

III. FLOW TOWARDS WATERMARKING

a. DCT:

Embedding Procedure:

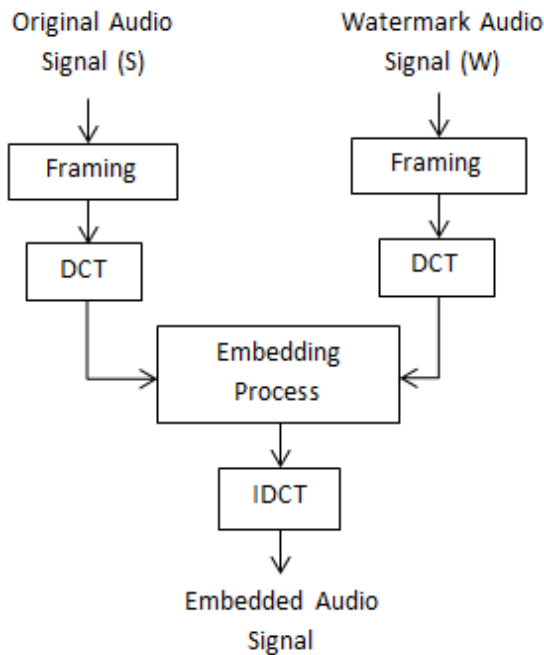


Fig.2 Embedding Procedure of DCT

Step 1: First, do partition the input audio file into frames. Repeat Step 1 for second (Watermark) audio also.

Step 2: Perform DCT transformation on original audio Signal. Repeat Step 2 for second (Watermark) audio also.

Step 3: Perform Embedding Process.
 $S_{em} = S + k \cdot W$; S_{em} = Embedded matrix

Step 4: Apply the inverse DCT operation to obtain each Watermarked audio frame. The overall Watermarked audio signal is obtained by Summing all Watermarked frames.

Extraction Procedure:

Step 1: First, do partition the watermarked audio file into Frames.

Step 2: Perform DCT transformation on watermarked Audio signal.

Step 3: Perform Extracting Process.
 $S_{ex} = (S_{em} - W)/k$;
 S_{em} = Embedded matrix,
 S_{ex} = Extracted matrix

Step 4: Apply the inverse DCT operation to obtain each Watermark audio frame.

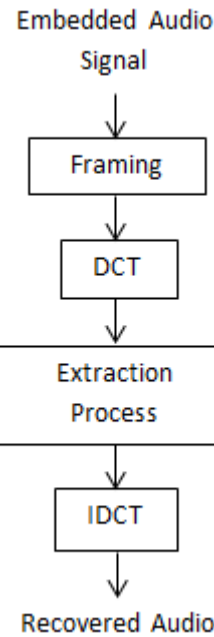


Fig.3 Extracting Procedure of DCT

b. DWT:

Embedding Procedure

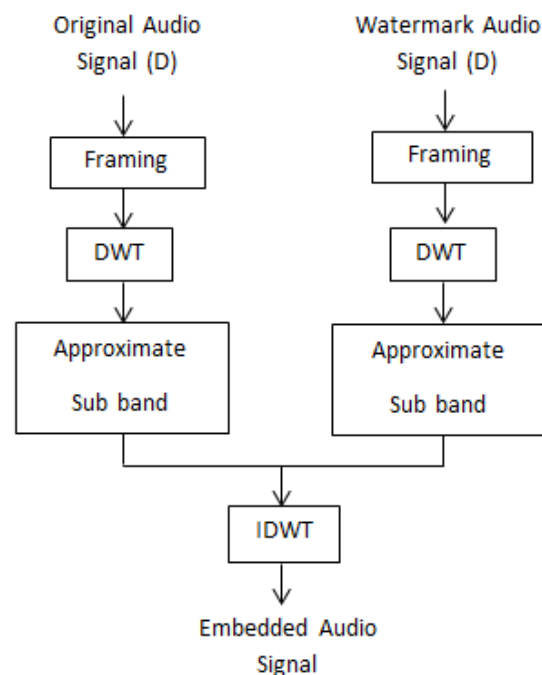


Fig.4 Embedding Procedure of DWT

Step 1: First, do partition the input audio file into frames. Repeat Step 1 for second (Watermark) audio also.

Step 2: Perform DWT transformation on original audio signal. Repeat Step 2 for second (Watermark) audio also.

Step 3: Performing DWT will produce two sub-bands: A, D. the D represents the Details sub-band; And A represents the approximation sub- band.

Step 4: Perform IDWT on approximation component of both the input signal. This will give the embedded audio of given signals.

Extraction Procedure

Step 1: First, do partition the watermarked audio file into frames.

Step 2: Perform DWT transformation on watermarked audio signal.

Step 3: Performing DWT will produces two sub-bands: Details sub-band and approximation sub-band.

Step 4: Perform IDWT on approximation component of the watermarked audio signal. This will give the extracted audio signal back.

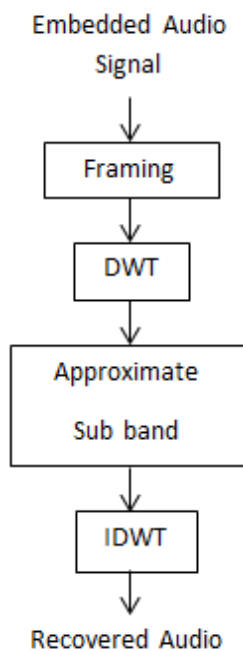


Fig.5 Extracting Procedure of DWT

c. PROPOSED ALGORITHM (DCT+DWT):

Embedding Procedure

Step 1: First, do partition the input audio file into frames. Repeat Step 1 for second (Watermark) audio also.

Step 2: Perform DWT transformation on original audio Signal. Repeat Step 2 for second (Watermark) audio also.

Step 3: Performing DWT will produces two sub-bands: Details sub-band and approximation sub-band.

Step 4: Perform DCT transformation on approximation Sub-band of both audio signal.

Step 5: Perform Embedding Process.
 $Sem = S + k*W$; Sem = Embedded matrix

Step 6: Apply the inverse DCT operation to Embedded matrix.

Step 7: Perform IDWT on given IDCT components. This will give the embedded audio of given signals.

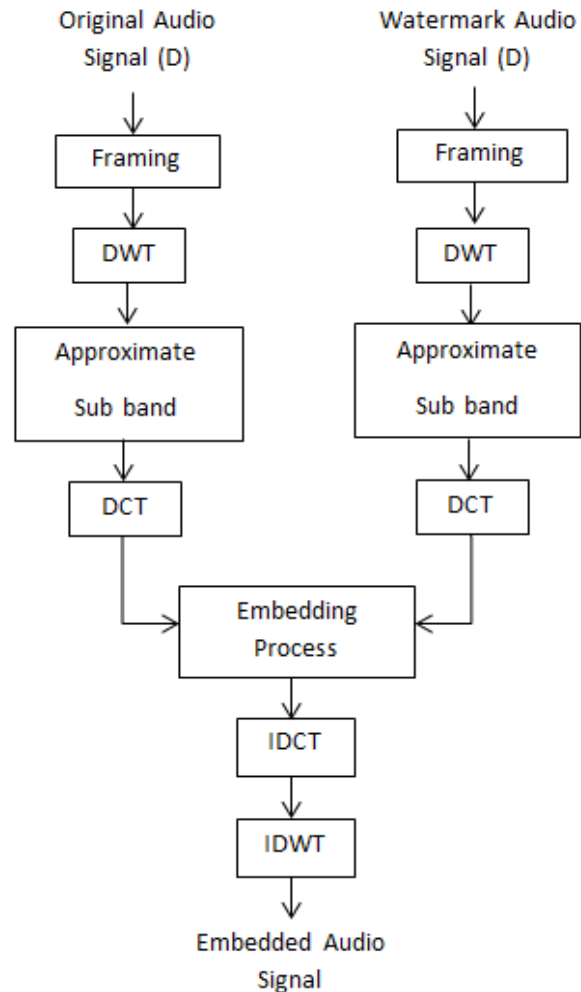


Fig.6 Embedding Procedure of Proposed Algorithm

Extraction Procedure

Step 1: First, do partition the watermarked audio file into frames.

Step 2: Perform DWT transformation on watermarked audio signal.

Step 3: Performing DWT will produces two sub-bands: Details sub-band and approximation sub-band.

Step 4: Perform DCT transformation on approximation sub-band of watermarked audio signal.

Step 5: Perform Extracting Process.
 $Sex = (Sem - W)/k$;
 Sem = Embedded matrix, Sex = Extracted matrix.

Step 6: Apply the inverse DCT operation to Extracted matrix.

Step 7: Perform IDWT on given DCT components. This will give the extracted audio of given signals.

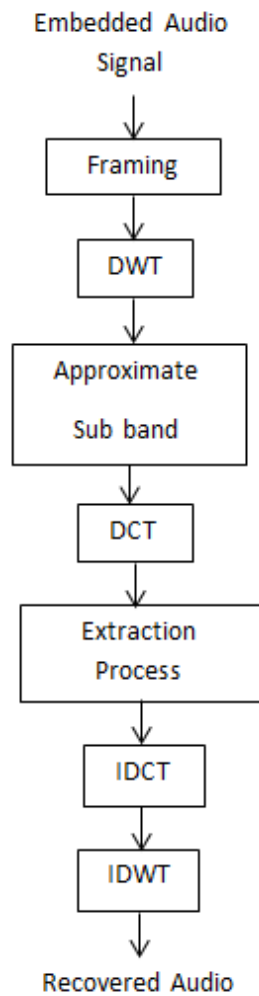


Fig.7 Extracting Procedure of Proposed Algorithm

IV. EXPERIMENTAL RESULTS AND EVALUATION

To evaluate the proposed algorithm, Experiments were conducted on five different types of audio signals such as Pop music, Country music, Folk music, Blue music and Classical music because they are much different in their spectral properties. Also we have introduced five different attacks like Random Noise, Additive White Gaussian Noise (AWGN), Cropping, MP3 and Re Sampling. They are here to examine the robustness of proposed algorithm. To measure imperceptibility, we use Peak Signal-to-Noise Ratio (PSNR) as an objective qualitative measure of a recovered audio signal.

PSNR CALCULATION:

PSNR [12] is given by the following equation:

$$PSNR = 10 * \log_{10} \left(\frac{MAX^2}{MSE} \right)$$

Where, MSE = Mean Square Error, which is given by,

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2$$

Where I (i, j) = Original Audio,
K (i, j) = Extracted Audio

Class of Audio Signal	Attack Type	DCT + DWT	DWT	DCT
POP	Without Attack	92.8639	92.8639	86.8235
	Random Noise	65.0980	65.1256	62.1347
	AWGN	79.2450	77.7976	76.4170
	Cropping	69.6111	74.9403	63.6248
	MP3	86.1413	67.7809	83.9064
	Re-Sampling	82.5876	67.7351	82.4942
COUNTRY	Without Attack	89.6808	89.6808	82.3896
	Random Noise	65.1076	65.0977	62.1020
	AWGN	77.6218	77.5976	74.8670
	Cropping	68.4479	72.3511	61.8463
	MP3	83.5364	63.3476	79.9093
	Re-sampling	77.6606	63.2935	76.0142
FLOK	Without Attack	76.0564	76.0564	79.9728
	Random Noise	64.7617	63.8327	62.1173
	AWGN	73.2866	73.3348	73.4845
	Cropping	65.8133	67.5517	65.6980
	MP3	74.8498	61.1358	77.7049
	Re-sampling	68.1457	60.9370	67.9201
BLUE	Without Attack	95.2765	95.2765	77.7918
	Random Noise	65.1356	64.0049	62.1259
	AWGN	73.9081	73.9117	70.9183
	Cropping	64.5936	63.7599	63.8930
	MP3	76.8586	58.7028	74.8980
	Re-sampling	74.2944	58.7056	72.2344
CLASSICAL	Without Attack	102.5566	102.5567	94.7836
	Random Noise	65.1333	65.1051	62.1152
	AWGN	83.1883	83.1634	80.2190
	Cropping	73.0006	82.4742	72.9974
	MP3	91.1440	75.6600	89.4309
	Re-sampling	87.8567	75.6582	86.1991

Table 1: Comparative PSNR Evaluation

V. CONCLUSIONS

A simple but Novel algorithm is discussed above. Also from comparative analysis with traditional techniques (as per table 1) proposed algorithm is far more robust & efficient against them. Above calculation (table 1) says that in 70% of total results proposed algorithm proves its robustness. But also in there are some attacks DWT & DCT are still gives better results than hybrid method.

DWT is still better in attack like cropping, while for other attack hybridization of DCT & DWT mostly proves its efficiency over this existing techniques.

REFERENCES

- [1] Michael Arnold, "Audio watermarking: features, applications and algorithms" and Expo, ICME 2000 IEEE International Conference vol. 51, 2000.
- [2] Paraskevi Bassia, Ioannis Pitas, Senior Member, IEEE, and Nikos Nikolaidis, Associate Member, "Robust audio watermarking in time domain" Can.J. Elect. Comput. Eng Vol. 33, No. 3/4, Summer/Fall 2008.
- [3] Seethal Paul, Sreelakshmi T.G, "Performance Analysis and Study of Audio Watermarking Algorithms", International Journal Of Engineering And Computer Science, ISSN:2319-7242, Volume - 3 Issue -8 August, 2014 Page No. 7540-7547.
- [4] Achintya Singhal, Anurag Narayan Chaubey, Chandra Prakash, "Audio Watermarking Using Combination of Multilevel Wavelet Decomposition, DCT and SVD", ETNCC, 2011, IEEE, PP: 239 – 243.
- [5] Ali Al-Haj, Ahmad Mohammad and Lama Bata, "DWT-Based Audio Watermarking", The International Arab Journal of Information Technology, Vol. 8, No. 3, July 2011.
- [6] Bender W., Gruhl D., Morimoto N., and Lu A., "Techniques for Data Hiding," IBM System Journal, vol. 35, pp. 313-336, 1996.
- [7] N. Lalitha, G. Suresh, V. Sailaja, "Improved Audio Watermarking Using DWT-SVD", International Journal of Scientific & Engineering Research, Vol. 2, Issue 6, June, 2011,ISSN 2229-5518.
- [8] Wei ,Zeng "A Novel Audio Watermarking Algorithm based on Chirp Signal and Discrete Wavelet Transform", WICOM, 2012,IEEE,PP:1-4.
- [9] Kakkirala, Krishna Rao,Chalamala, Srinivasa Rao Rao, G Bala Mallikarjuna, "DWT-SVD Based Blind Audio "Watermarking Scheme For Copyright Protection", Audio, Language and Image Processing (ICALIP), 2014 International Conference on IEEE Publications, Page(s): 180 – 183.
- [10] Rakhi Dubolia , Roop Singh, Sarita Singh Bhadoria , Rekha Gupta, "digital image watermarking by using discrete wavelet transform and discrete cosine transform and comparison based on psnr," 2011 International Conference on Communication Systems and Network Technologies, IEEE.
- [11] The Discrete Cosine Transform(DCT): Theory and Application, Syed Ali Kbayam, Department of Electrical & Computer Engineering, Michigan State University, March 10th, 2003.
- [12] Liu Ping Feng, Liang Bin Zheng, Peng Cao, "A DWT-DCT Based Blind Watermarking Algorithm for Copyright Protection", IEEE Transactions, 2010.
- [13] T. Thielde, et al., "PEAQ-The ITU standard for objective measurement of perceived audio quality," Journal of Audio Engineering, vol. 48, no. 1/2/3, pp. 3-29, January 2000.
- [14] Dataset : <http://www.audiomicro.com/royalty-free-music> Standard ITU Audio dataset : <http://www.itu.int/net/itu-t/sigdb/menu.aspx>