# A Digital Audio Watermarking Scheme using DCT and Synchronization Code


**Tribhuwan K. Tewari , Vikas Saxena**
**CS& IT Dept.**
**JIIT Noida,**
**Noida,,India**
**tribhuwan.tewari@jiit.ac.in**, vikas.saxena@jiit.ac.in

**J. P. Gupta**
**Vice Chancellor.**
**Sharda University**
**Greater Noida, ,India**
jaipgupta@gmail.com



***Abstract*— In this paper, a blind digital audio watermarking technique is proposed which uses a Discrete Cosine transform (DCT) and synchronization code. In our method the watermark is embedded into the selected coefficient of the DCT. The selected coefficients are middle band DCT coefficients which are modified and quantized through the average energy of the selected sub frame for watermark embedding. The mid band frequency components are selected through experiments such that the effect of mp3 compression and the common signal processing operation has a minimum effect on these coefficients . Original audio will not be required for extraction of the watermark. To adapt to the mp3 attack a preprocessing step is embedded which did an mp3 conversion of the audio and convert again the audio to the original format prior to embedding. For dealing with the synchronization attacks a 12 bit Barker code is used. Experiments show that our scheme produces imperceptible audios and the watermarked audio is robust against common signal processing attacks.**



**Keywords—** ***Audio watermarking, Intellectual property, Barker code, Robustness, Discrete Cosine Transform (DCT).***


## I INTRODUCTION

As the digital data can be replicated and distributed over the fast working networks in no time, Digital revolution which was a boon for the society has become a curse for the multimedia creator now. To counter this and for protecting the rights of the media creators, a technique called Digital Watermarking is introduced & evolved for copyright protection. Thus Digital Audio watermarking has also established itself as a need and necessity of copyright protection for audio industries. Digital Watermarking is a process of embedding an author specific data (watermark) on to the digital media. When any issue comes regarding the ownership the watermark is retrieved and the ownership issue is resolved. There are practical cases also where audio watermarking was used as a legal aid to punish the people involved in piracy of the audio or audio video content [1]. As far as audio watermarking is concerned it is the latest among image, text and video watermarking. Audio watermarking methods exploit the deficiencies in the human auditory systems. They try to embed the owner specific information to the audio regions where it produces minimum distortion. The mandatory properties which must be followed by any watermarking system are:

1) **Imperceptibility** – which essentially says the after the watermark has been added to the cover audio the resultant audio should be same as the original. According to the International Federation of Phonographic Industry (IFPI) the minimum of 20dB Signal to Noise Ratio (SNR) should be maintained by the watermarked audio.
2) **Robustness –** It is the ability to resist to the attacks which tries to remove the watermark or corrupt it so that during extraction no useful information is extracted.
3) **Payload-** It is the least talked about property which is defined as the amount of watermarking data which is embedded on to the cover audio. It is measured in bits per second i.e. bps.

But the contradiction to these requirements is always there as all the three cannot be achieved simultaneously. With increase in capacity robustness can be increased but will definitely reduce the imperceptibility property. These three properties can be viewed as the vertices of a triangle which cannot be visited along a single line direction. The maximum you can visit are the two vertices along a single edge.

Audio Watermarking methods have much application which is as follows:

1) **Ownership protection**: the objective is to embed the information that identifies the owner of the digital media in order to prevent the other parties claim the copyright. The utmost requirement is a high level of robustness & certainty of the watermark to still resolve rightful ownership if other parties embed additional watermarks.
2) **Fingerprinting**: the objective is to convey information about the legal recipient rather than a source of digital media, in order to identify single distributed copy of the digital media.
3) **Authentication & Tamper Detection**: The objective is to detect modification of data .This is achieved through **fragile** watermark which are having low robustness to certain modifications.

4) **Copy Protection**: The objective is to find a mechanism to disallow unauthorized copy of the digital media. This can only be efficiently implemented in a closed system.
5) **Broadcast Monitoring:** A watermark is embedded into the multimedia on behalf of the service provider. All the information relevant to the embedded watermark is stored in a central database.
6) **Position finding of pirates with in a movie hall :** One of the recent application of audio watermarking is in finding the position i.e. the seat number and thereafter his/her name and address who is charged of making a pirated copy from a movie hall**.[2]**
7) **Stationarising audio:** these is also one of the latest application where watermarking method is used to stationarise an audio signal specially when a video conferencing is done and the audio is required to be made as clear as possible [3]

In audio watermarking, the embedding of the watermark can be done in temporal as well as transform domain. LSB (Least Significant Bit) modification and substitution are common temporal domain techniques in which the embedding is directly done on the LSB or higher bit layer or uses time domain statistics or masking principle, echo hiding for embedding of the watermark. [4][5][6][7].In transform domain techniques first the audio is transformed and then the embedding of the watermark is done. The temporal domain watermarking methods exhibits high payload but they lack in robustness as through simple signal processing attacks the watermark can be removed. Research is going on to make these methods robust by shifting the embedding bit layer from LSBs towards higher MSB's [8].The transform domain techniques are robust as compared to temporal domain techniques.

In this paper, the watermark embedding is done in DCT domain by modifying the mid band frequency DCT coefficients. The mid band frequency components are identified through the experiments which are least effected through mp3 compression. The audio is first segmented into blocks and only those blocks are selected which are carrying minimum threshold energy as against the traditional DCT based approach where all the blocks are considered for watermark embedding. The energy threshold is decided through experiments and is adaptive to different audio signals. In addition the block size has been made as dynamic so that the intruder cannot identify the block of embedding. The results shows that the watermarked audio is having high SNR and also it is robust to common signal processing attacks along with the intentional attacks.The rest of the paper is organized as follows. Section II gives a brief of the related audio watermarking approaches which uses DCT for embedding. Section III gives our proposed embedding and extraction strategy. Section IV gives the discussion on the result mainly the imperceptibility and the robustness of the watermarked audio followed by the conclusion in Section V.

## II RELATED RESEARCH

In the past years a number of audio watermarking methods were evolved for copyright protection. The DCT based approaches are said to be simple and effective as the watermark embedding is done in transform domain. The DCT based approach is useful as it tries to accumulate the energy in some of the coefficients and also it minimizes the relationship between adjacent samples with in an audio. Z Zhou et al. (2007) [9] proposed robust DCT based scheme where the watermark bits are embedded by quantization of the DCT coefficients. W. Youngqi et al. (2008) [10] proposed DCT based audio watermarking scheme using a synchronic signal embedded in the low frequency components. J Wei's et al. (2010) [ 11] scheme uses DCT coefficients for embedding but the DCT coefficients are selected in a non uniform manner for enhancing security .Xia Zhang et al. (2011) [11] uses the double DCT method for the transmission of the audio through air channel. Di Chang et al. (2011) [12] proposed DCT domain technique which modifies the low frequency DCT coefficients for watermark embedding. He also uses Barker code for synchronization. K Ren et al. (2011) [13] proposed DCT and DWT based algorithm which uses a color image as a watermark. He claims the scheme to have a high payload as compared to the contemporary DCT based technique. For scrambling the watermarking bits he applies Arnold transform. The watermark is embedded in the low frequency DWT coefficients. Lalitha et al. (2012) [14] uses a DCT and SVD based approach for embedding and extended it to DWT and SVD. He compared the two approaches and claim DCT based method to be more robust. Q. Gou et al. (2012) [15][16] also gives a DCT based scheme which he claims to be robust especially against analog to digital and digital to analog conversion for air channel transmission applications. The proposed methods in the literature don't consider the energy of the individual blocks for watermark embedding. This is a problem since low energy block are always susceptible to being removed altogether through an attack without disturbing the cover audio. Also less work was being done in analyzing the effect of common signal processing operation including mp3 conversion onto the different DCT coefficients and the DCT blocks. In the next section we propose a method in which the selective DCT blocks are used for watermark embedding and only the mid frequency DCT Coefficients are manipulated for this purpose. The mid frequency DCT coefficients are selected through exhaustive experiments by comparing the normalized correlation of the extracted watermark and the original watermark at different mid frequency range selected for watermark embedding.

## III PROPOSED WATERMARKING METHOD

In this section we describe the embedding and extraction strategy of watermarking. The grey image which is the watermark information is converted into a bit of stream. If the size of the image is M x N with M & N as the number of rows and columns respectively then it is converted into a one dimensional bit stream of MxN.

Let $W_i$ be the binary watermark thus produced

$W_i = \{[0, 1],\}$

With Len =Length(W)=MxN

### III.1 WATERMARK EMBEDDING METHOD

**Step 1**. The preprocessing operation is done on the original cover audio .The original wav audio file is subjected to mp3 conversion at 128kbps and converted back to the same original format.

**Step 2**. The average energy of the audio segment is calculated. For this step silence portion of the audio are not considered.

$$E=\sum_{i=1}^{N} \left|\left(A_i * A_i\right)\right|$$

**Step 3**. Audio signal is sampled at a rate of 44.1 kHz. Then partition the entire samples into segments & segments into sub segments of equal length say L consisting N frames. Each frame is composed of 25 ms samples. The summation of *N* frames makes up the overall sampled audio sub segment signal as illustrated in the following equation:

$$\mathbf{F} = \sum_{i=1}^{N} A_i$$

**Step 4**. Let $F^k$ denotes k $^{th}$ sub segment , $F^k$ is cut into two **s**ections $F^{k1}$ and $F^{k2}$ with length as $L_1$ and $L_2$ samples respectively. Synchronization code and watermark are embedded into $F^{k1}$ and $F^{k2}$ respectively.

**Step 5**. The frames of $F^k$ for which frame energy is greater than 2/3 of the overall energy of the audio are selected for watermark embedding.

**Step 6**. For watermark embedding on the kth selected sub segment $F^{k2}$ DCT is performed.

$D^{k2} = DCT(F^{k2})$

**Step 7**. Select the mid frequency components with the index as [132:232] for every selected frame.

**Step 8**. Embed the watermark bit by quantizing the energy of the selected coefficients and redistributing them as follows

$E^{k2} \longrightarrow$ Energy $(D^{k2})$

$E \longrightarrow E^{k2}$

```
n = floor( E/del)
if (hide(k)= =1)
Do
nm=n+d-mod(n,(2*d))
else
nm= n+d-mod(n+d,(2*d))
End

xm=del+nm+del/2
for i=1: Len
D k2(i) = D k2(i) * xm/E
end
```

floor is the floor function which returns highest integer value lesser than the argument passed

mod is the modulus function which returns the remainder of the division of two arguments passed.

D and del are the key used for embedding.

**Step 9** Repeat steps 4 -6 to embed multiple copies of the watermark on each sub segments.

### III.2 WATERMARK EXTRACTION METHOD

**Step 1**. Step no. 2 –3 of the watermark embedding process are repeated.

**Step 2**. The signal is aligned with respect to the synchronization code.

**Step 3**. Steps 4- 7 of the watermark embedding process are repeated

**Step 4.** $E^{k2} \longrightarrow$ Energy $(D^{k2})$

$E \longrightarrow E^{k2}$

```
n = floor( E/del)

for i =1 : size(D k2)
if mod(n,2) == 0
hide(i) = 0
else
hide(i)=1
end
```

**Step 5** Repeat steps 1 to 4 to extract multiple copies.

## IV. RESULTS

The performance analysis of the proposed method is done by finding the signal to noise ratio, and the normalized correlation of the embedded and the extracted watermark.

For ready reference they are all defined below.

1. **Peak Signal to Noise Ratio (PSNR):** The signal to noise ratio is defined as the ratio of the signal power to the noise power and is given by

$$\sum_{i=1}^{M} \max(X(i)) \Big/ \sqrt{\sum_{i=1}^{M}(X(i) * X(i) - N(i) * N(i))}$$

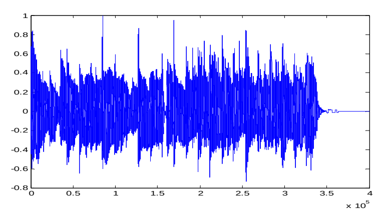
2. **Normalized Correlation Coefficient (NC)**: It is defined as the similarity between the extracted watermark and the embedded watermark and is given by

$$\sum_{i=1}^{M} W(i) * W'(i) \Big/ \sqrt{\sqrt{\sum_{i=1}^{M} W(i) * W(i)}\ \sqrt{\sum_{i=1}^{M} W'(i) * W'(i)}}$$

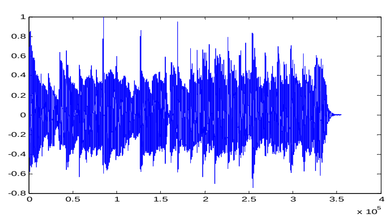
Where W & W' are the the original and the extracted watermark respectively.

3. **Bit Error Rate (BER)** : It is the ratio of the no. of watermarking bits which are having interpreted falsely to the total no watermarking bits.

The following figures shows the spectra of original and watermarked guitar audio. The spectra of both audios Doesn't show any considerable change.



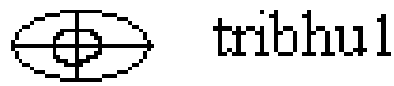
a )Original audio



b)Watermarked audio

Fig. 1. Watermarked and Original Audio

Two watermarks are used as copyright information for watermarking the audios which are shown in the figure.




a) Image b) text image

Fig. 2. Watermarks used

This sub-section in the form of tables introduces the test results of watermarking algorithm on different type of audio set. Table 1 gives the result of the PSNR of the watermarked audio against the original with the BER of the extracted watermark. Table 2 gives the NC and BER of the extracted watermark when the watermarked audio is subjected to different attacks. From the result we analyze that our algorithm works good for different content categories and attacks like the synchronization & the compression attacks. Also the watermarked audio exhibits a good PSNR.

**TABLE 1: BER(%) & PSNR value of proposed Algorithm**

| S. No | Audio Name | Type | PSNR | BER % |
|---|---|---|---|---|
| **1** | Bell.wav | Mono | 33.93 | 2 |
| **2** | Noise.wav | Mono | 31.23 | 2 |
| **3** | Beat.wav | Stereo | 33.13 | 0 |
| **4** | Guitar.wav | Mono | 34.8 | 0 |
| **5** | Metal.wav | Stereo | 34.25 | 1 |
| **6** | Loop.wav | Stereo | 34.23 | 2 |
| **7** | Sports.wav | Stereo | 33.5 | 2 |
| **8** | Female.wav | Stereo | 33.7 | 3 |
| **9** | Piano.wav | Mono | 34.20 | 1 |
| **10** | Bird.wav | Mono | 33.3 | 0 |

Some of the standard attack results to check the robustness of the proposed algorithm are presented next.

**TABLE 2: %BER, NC and SNR values of after introduced to different attacks**

| S. No | Attack | Guitar.wav (35 sec) | | Piano.wav (25 sec) | | Instrumental.wav (28 sec) | | Voice.wav (15 sec) | |
|---|---|---|---|---|---|---|---|---|---|
| | | **BER** | **NC** | **BER** | **NC** | **BER** | **NC** | **BER** | **NC** |
| **1** | Amplify | .03 | .95 | .1 | .98 | .02 | 1 | .07 | .91 |
| **2** | Add Sinus | .06 | .94 | .03 | 1 | .07 | .92 | .02 | .95 |
| **3** | Add Noise | .07 | .95 | .09 | .94 | .05 | .95 | .04 | .92 |
| **4** | Smoothening | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| **5** | Over Sampling | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| **6.** | Sub Sampling | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| **7.** | Half Resolution | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| **8.** | Mp2 compression | .03 | .92 | .04 | .91 | .07 | .91 | .05 | .89 |
| **9.** | Mp3 compression | .12 | .68 | .11 | .71 | .72 | .7 | .11 | .67 |

## V. CONCLUSION

In this paper, we propose a digital audio watermarking algorithm based on DCT domain. To improve the robustness of audio watermark, the proposed algorithm uses a preprocessing step. For dealing with synchronization attacks robust Barker code is used as synchronization code, which is embedded in the mean value of several sub frames and embedding of watermark is done on the selected mid frequency DCT coefficients through energy quantization. Experimental results have illustrated the robust nature of our synchronization embedding scheme and inaudible nature of our watermarking scheme. In addition, the watermark can be extracted without the help from original audio signal and procedure for embedding and extraction can be easily implemented. Despite the success of the proposed method, it also has drawbacks. Further research will focus on over-coming the drawbacks.

## REFERENCES


[1]. http://if.com.au/2012/07/09/article/PKUQMLJDOK.html.

[2]. Y. Nakashima, R. Tachibana, N. Babaguchi, "Watermarked Movie Soundtrack Finds the Position of the Camcorder in a Theater", IEEE Transactions on Multimedia, 2009, Volume: 11 , Issue: 3, Page(s): 443 – 454.

[3]. S.D Larbi, M. JaidaneSaidane "Audio watermarking: a way to stationnarize audio signals", IEEE Transactions on Signal Processing, 2005 ,Volume: 53 , Issue: 2 , Part: 2, Page(s): 816 – 823.

[4]. M. Swanson, B. Zhu, A. Tewfik, and L. Boney, "Robust audio watermarking using perceptual coding", Signal Processing, vol. 66, no. 3, pp. 337-356 , April 1998.

[5]. P. Bassia, L. Pitas, and N. Nikolaidis, "Robust audio watermarking in the time-domain", IEEE Transactions on Multimedia., vol. 3, no. 2, pp. 232-242 , July 2001.

[6]. A. Lemma, J. Aprea, and L. Kherkhof, "A temporal-domain audio watermarking technique", IEEE Transactions on Signal Processing, vol. 51, no. 4, pp. 1088-1097 , April 2003.

[7]. M Zamani, A B Manaf, "Magalingam, P.; A Novel Approach for Audio Watermarking", Proc. of Intl. Conf. on Information Assurance and Security (IAS) '09. Volume 2, Page(s): 83 – 86.

[8]. W. Cao, "Bit Replacement Audio Watermarking using Stereo Signals ",Proc. of Intl. Conf. on New Trends in Information and Service Science, NISS '09,Page(s): 603-606.

[9]. Z. Zhou et. Al," A Novel Algorithm for Robust Audio Watermarking Based on Quantification DCT Domain", Proc. of Intl. conf. on Intelligent Information Hiding and Multimedia Signal Processing IIHMSP 2007 , Volume :1, Page(s): 441-444.

[10]. W. Youngqi et al.," Asynchronous Audio Watermarking Algorithm Based on Chaotic Encryption in DCT Domain", Proc. of Intl. Conf. on Information Science and Engineering, ISISE 08, Volume :2, Page(s): 371-374.

[11]. X. Zhang et.al. “An Audio watermarking algorithm transmitted via air channel in double DCT domain”, Intl. Conf. on Multimedia Technology ,ICMT 2011,Page(s): 2926: 2936.

[12]. X. Zhang et. al. ,”A robust digital watermarking algorithm in DCT domain for air channel transmission”, Proc. of Intl. Conf. on Multimedia Technology ICMT 2011,Page(s): 2896:2899.

[13]. K. Ren,”Large capacity digital audio watermarking algorithm based on DWT and DCT”, Proc. of Intl. Conf. on Mechatronic Science, Electric Engg. And Computer, MEC 2011, Page(s): 1765-1768.

[14]. Lalitha et. al. “ An efficient and simple Audio Watermarking using DCT-SVD”, Proc. of Intl. Conf. on Dvices,Circuits and Systems ,ICDCS 2012,Page(s): 177-181.

[15]. Q. Guo et. al. “An air channel digital audio watermarking algorithm based on DCT domain”, Proc. of Intl. Conf. on Consumer Electronics ,Communications and Networks ,CECNet 2012, Page(s): 853-856.

[16]. Q. Guo et. al,” An audio digital watermarking algorithm against A/D and D/A conversions based on DCT domain”, Proc. of Intl. Conf. on Consumer Electronics ,Communications and Networks CECNet 2012 ,Page(s) : 871-876.