Audio Watermarking Technique using Modified Discrete Cosine Transform

Aparna S
Dept. of Electronics & Communication Engineering
LBS Institute of Technology for Women
Thiruvananthapuram, India

Baiju P S
Dept. of Electronics & Communication Engineering
LBS Institute of Technology for Women
Thiruvananthapuram, India

***Abstract*— Audio watermarking is a data hiding technology widely used to increase the authentication of audio files and for copyright protection. Watermarking techniques hides copyright information, which can be a binary sequence or an image, within the audio signal in such a way that it is not perceivable by the listener. A large number of audio watermarking techniques are available. Authentication requires only a simple watermarking technique but to ensure security along with authentication, strong watermarking methods should be adopted. The proposed method uses watermarking based on Modified Discrete Fourier Transform (MDCT). Experimental analysis shows that the proposed method provides high audio quality when compared to other methods such as Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT).**

***Index Terms*— Watermarking, permutation, DCT, DWT, MDCT, robust**

I. INTRODUCTION

The rapid advancement in communication networks is providing easier ways to circulate multimedia files from one point to another. To protect the digital files from illegal copying and manipulation, watermarking techniques are used. Watermarking hides information for copyright protection such as a logo of the creator in the host signal. Digital audio watermarking [1] should satisfy the following properties.

- Imperceptibility: the embedded watermark should not affect the quality of the original audio file. I.e., there should not be a noticeable change in the original signal. Perceptual transparency is one of the main requirements of watermarking.
- Robustness: The embedded watermark should be able to withstand different types of attacks such as linear and non linear filtering, compression, analog to digital and digital to analog conversions etc.
- Security: The watermark should not be detectable for an unauthorized agent.
- Verification: The watermark should provide reliable information for copyright.

Audio watermarking is mainly of two types, time domain techniques and frequency domain techniques. In time domain techniques like LSB encoding, watermark data is added to the raw signal and hence even a low pass filter can remove it. Where as in frequency domain techniques, information is added to audio signal converted into frequency domain using suitable transformation techniques such as Discrete Fourier Transform (DFT), DCT and DWT. They provide better security.

In audio watermarking, a unique digital identifier such as an image which can be a logo of the creator or the owner of the actual audio file is embedded into it. The embedding process should be such that it is highly difficult for an intruder to remove or change the watermark. To ensure this, a secret key is used while embedding the watermark. When the audio file is copied the watermark is also carried in the copy so that it can be extracted and used as a proof of ownership when a dispute occurs. It will be difficult to identify the owner of the file if it is not containing any watermark.

This paper proposes an audio watermarking technique based on MDCT. Over the years, MDCT has emerged as an efficient audio transformation technique. MDCT reduces blocking artifacts by flexible window switching and helps cancelling time domain and frequency domain aliasing. Transformed signal can be perfectly reconstructed using overlap addition.

An image is used as the watermark information, which is encrypted before adding to the audio signal. A secret key used in the embedding section ensures security of the watermarked data. The proposed method is compared with DCT and DWT based watermarking.

The rest of the paper is organized as follows. Section II discusses the transform techniques used for audio watermarking, Section III is about the proposed method in detail, section IV gives the experimental results and analysis and finally section V concludes the work.

II. TTRANSFORM BASICS

In this section, DCT, DWT and MDCT transforms has been described.

A. Discrete Cosine Transform

The discrete cosine transform (DCT) [2] is a technique for converting a signal into elementary frequency components. It is a spectral transform which has properties of Fast Fourier Transform (FFT). For a 1-D signal of length N, DCT and inverse DCT are given as

$$C_{dct}(u) = \alpha(u) \sum_{x=0}^{N-1} f(x) \cos\left[\frac{\Pi(2x+1)u}{2N}\right], \text{for } u = 0,1,2,.....N-1 \quad (1)$$

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

where

$$\alpha(u) = \begin{cases} \sqrt{1/N} & \text{for } u = 0 \\ \sqrt{2/N} & \text{for } u \neq 0 \end{cases} \quad (3)$$

B. Discrete Wavelet Transform

The discrete Wavelet Transform (DWT) gives the time-frequency representation of a signal [3], [4].

$$W_{\varphi}(j,k) = \frac{1}{\sqrt{M}} \sum_n f(n)\varphi_{j,k}(n) \quad (4)$$

$$W_{\varphi}(j,k) = \frac{1}{\sqrt{M}} \sum_n f(n)\varphi_{j,k}(n) \quad (5)$$

Equation (4) gives the wavelet function and equation (5) gives the scalar function where Ψ is the wavelet function, and Φ is scaling function.

C. Modified Discrete Cosine Transform

The Modified Discrete Cosine Transform (MDCT) [5] can be realized practically as a block transform using a fast algorithm. It is very efficient in analyzing the channel characteristics with respect to their impulse response and frequency response. The mathematical expressions for forward and inverse MDCT are given by (6) and (7) respectively.

$$X[k] = \sum_{n=0}^{L-1} x(n)\cos\left[\frac{\pi}{2L}k\left(2n+1+\frac{L}{2}\right)(2k+1)\right], k = 0,1,2,..., \frac{L}{2}-1 \quad (6)$$

$$X[n] = \frac{1}{L}\sum_{k=0}^{L/2-1} X(k)\cos\left[\frac{\pi}{2L}k(2n+1+\frac{L}{2})(2k+1)\right], n = 0,1,2,..., L-1 \quad (7)$$

MDCT is used for audio processing in this work. The embedding and extraction process are described in section III.

III. Watermark Embedding and Extraction Process

The data to be hidden in the host audio signal is encrypted by using permutation technique. Permutation [6] is used to reduce the predictability of image components.

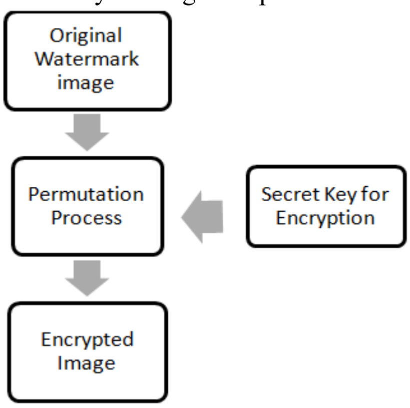
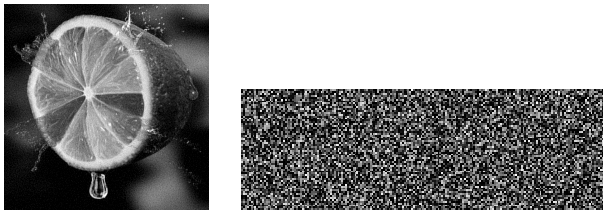


Fig. 1. Image encryption process

Usually with the given pixel information in an image one will be able to estimate its neighboring pixel values. This is possible because of the interdependency of image components. There exist different types of permutation techniques [7].

- Block permutation: In block permutation the image is divided into different sub-blocks and they are rearranged in position according to a secret key.
- Pixel permutation: In pixel permutation Pixel groups are shuffled in position.
- Bit permutation: Permutation at bit level permutes each pixel taken from the image using the chosen key.

Pixel permutation is selected in this work since it is the most effective technique among all three. The correlation between image elements will be reduced and hence it will be difficult to predict the neighboring pixel values from the given value. At the extraction stage, the image can be reconstructed using inverse permutation using the same secret key.



(a) (b)

Fig. 2. (a) Original watermark image (b) encrypted image

The audio signal is divided into different frames so that the image components are inserted into a single bit position of each audio frame. The sample frames are then converted into frequency domain using MDCT [8]. Forward and reverse MDCT are given in equations (6) and (7) respectively where x represents the input audio frames of sample length L and X represents the transformed coefficients. MDCT uses 50% overlapped samples. So it generates only L/2 coefficients for each sample frame of length L. As given in equation (8) MDCT transform can be considered as a filter bank with impulse function

$$h_k(n) = w(n)\sqrt{\frac{2}{M}}\cos\left(\frac{(2n+M+1)(2k+1)\pi}{4M}\right) \quad (8)$$

From each of the individual audio frame, the first bit position is identified and the image component is added to the audio component. This method exploits the properties of Human Auditory system (HAS) [9]. Even though HAS has a wider dynamic range compared to Human Visual System, it tend to mask [10] the slight amplitude changes in an audio signal. IMDCT is used to convert the signal back to time domain.

Step1: Divide the audio file into number of frames, according to the amount of information in the watermark data.
Step2: Apply MDCT to the frames
Step3: Locate the first bit of each audio frame.
Step4: Add one image component each to the first bit of all the frames after proper scaling.

Step5: Apply IMDCT to convert the frames back to time domain

For extracting the watermark image the corresponding bit positions where the image components were added are located and subtraction operation is performed. The extracted information from all frames are combined to get the permuted image. The original image can be obtained using inverse permutation and resizing it back to its original specifications.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

The experiments were carried out in MATLAB with a 44.1 kHz audio file in .wav format having 10 seconds duration and a grey level image as watermark, according to the proposed method. The image was resized to 64 x 64 before encryption, since large amount of embedded data may severely degrade the audio quality. After extracting the watermark image from the host signal, performance evaluation was conducted by analyzing Normalized Cross Correlation (NCR) given by

$$NCR = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} W_{original}(i,j)\, W_{reconstructed}(i,j)}{\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N} W_{original}(i,j)^2}\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N} W_{reconstructed}(i,j)^2}} \quad (9)$$

Where $W_{original}$ and $W_{reconstructed}$ stand for the coefficients of the actual watermark image and the reconstructed image respectively and N is the total number of bits in the image. A high NCR value indicates higher correlation between two images and high accuracy. NCR in the range of 0.90 to 0.95 were obtained for various audio files in MDCT. MDCT provides better performance results compared to conventional methods like DCT and DWT.

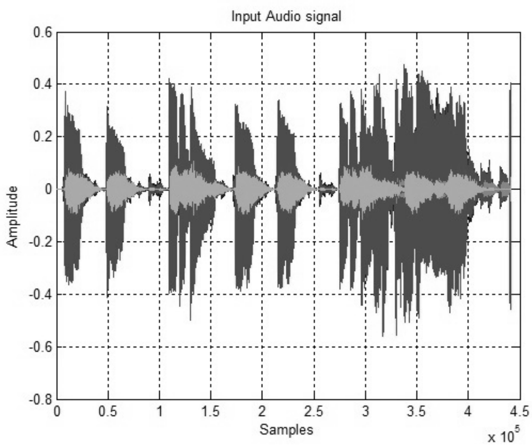


Fig. 3. Host Audio signal with sampling frequency 44.1 kHz

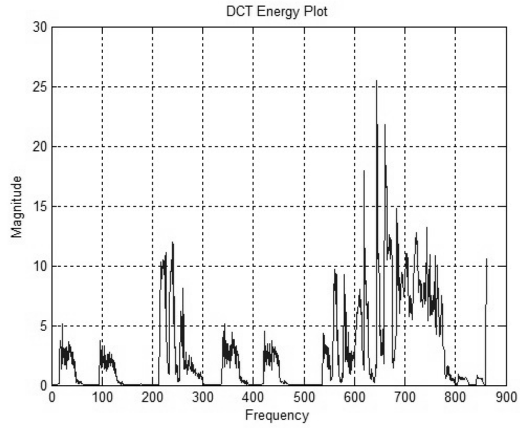


Fig. 4. Energy plot of the selected audio segment

Figure.5 shows the reconstructed image using different transform bases. Among the three methods, MDCT provides the better reconstructed image. The NCR values obtained by computing the correlation between actual watermark image and the reconstructed image is given in Table.I. Bingwei Chen and Jiying Zhao proposed a method, in which a sequence is used as watermark information instead of image. Hence it cannot be compared with the proposed method in terms of NCR.

Fig. 5. Image reconstructed using MDCT, DCT and DWT respectively

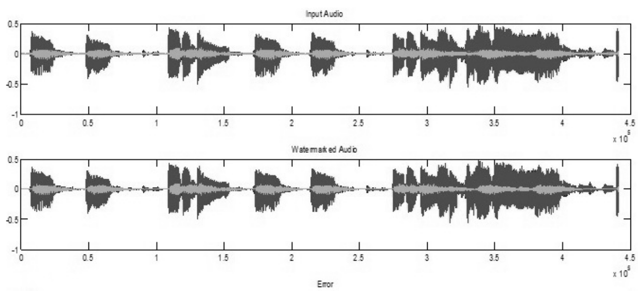


Fig. 6. Original audio signal and the watermarked audio signal

TABLE I. COMPARISON OF WATERMARKING TECHNIQUES

Method used for watermarking	MDCT	DWT	DCT
NCR value obtained	0.912	0.7354	0.8431

V. CONCLUSION

This paper presents a perceptually transparent audio watermarking method using MDCT by exploiting the properties of Human auditory system. Image components are added to a single bit of each audio frame, to maintain the hearing quality. Watermark embedding was performed using a scaling factor. Watermark extraction was performed using the same secret key used for embedding process and the correlation between actual image and reconstructed image was determined using the factor NCR. It was found that the NCR value is high for MDCT based method compared to other methods like DCT and DWT. So MDCT based watermarking algorithm is better than many of the conventional methods in terms of efficiency and perceptual quality.

REFERENCES

[1] Sanjay Pratap Singh Chauhan, S. A. M. Rizvi "A survey: Digital Audio Watermarking Techniques and Applications", International Conference on Computer and Communication Technology (ICCCT), 2013.

[2] Keqiang Ren, Huihuan Li "Large Capacity Digital Audio Watermarking Algorithm Based on DWT and DCT", International Conference on Mechatronic Science, Electric Engineering and Computer, Jilin, China August, 2011.

[3] H. L. Dai, D. He "An Efficient and Robust Zero-Watermarking Scheme for Audio Based on DWT and DCT", Microelectronics & Electronics, 2009, pp. 233-236.

[4] N.V.Lalitha, Ch.Srinivasa Rao, P.V.Y.JayaSree, "DWT-Arnold Transform Based Audio Watermarking" IEEE Asia Pacific Conference on Postgraduate Research in Microelectronics and Electronics (PrimeAsia), 2013.

[5] Bingwei Chen, Jiying Zhao, "An Adaptive and Secure Audio Watermarking Algorithm Robust to MP3 Compression", International Instrumentation and Measurement Technology Conference Singapore, May 2009.

[6] Ravi Prakash Dewangan ,Chandrashekhar Kamargaonkar "Image Encryption using Random Permutation by Different Key Size", International Journal of Science, Engineering and Technology Research (IJSETR), Volume 4, Issue 10, October 2015.

[7] Mohammad Ali Bani Younes, "Image Encryption Using Shuffling Technique", International Journal of Computer Science Research & Technology (IJCSRT) Vol. 1 Issue 3, August, 2013.

[8] Mu-Liang Wang, Hong-Xun Lin, Mn-Ta Lee "Robust audio watermarking based on MDCT coefficients", International conference on genetic and evolutionary computing,2012.

[9] Santosh Kumar Singh, Jyotsna Singh," Audio Watermarking Scheme in MDCT Domain" International Journal of Electrical and Electronics Engineering (IJEEE) , Vol-1, Iss-3, 2012.

[10] Liang Wang, Sabu Emmanuel, Mohan S. Kankanhalli, "Emd And Psychoacousticmodel Based Watermarking For Audio" IEEE International Conference on Multimedia and Expo (ICME), 2010.