Audio Watermarking with Encryption

John Britto
Student, SPIT
Mumbai, India

M.Mani Roja
Professor, TSEC
Mumbai, India

Abstract—Watermarking is an effective means of hiding data, thereby protecting the data from unauthorized or unwanted viewing. Watermarks have been proposed previously as a way for protecting the data. This digital signature could discourage copyright violation, and may help determine the authenticity and ownership of an image. In this paper, audio watermarking is done with the help of Discrete Fourier Transform (DFT). To enhance the security of the watermark, pseudo noise (PN) based encryption is used.

Keywords—Water Marking; DFT; Encryption; PN sequence; MSE

I. INTRODUCTION

Watermarking is “the practice of imperceptibly altering a Work to embed a message about that Work” [1]. Watermarking is becoming increasingly popular, especially for insertion of undetectable identifying marks, such as author or copyright information to the host signal. Watermarking may probably be best used in conjunction with another data-hiding method such as steganography, cryptography etc. Such data hiding schemes, when coupled with watermarking, can be a part of an extensive layered security approach.

Audio watermarking is defined as secure communication of data related to the host audio signal, which includes embedding into, and extraction from, the host audio signal” [2]. Watermarking, being ideally imperceptible, can be essentially used to mask the very existence of the secret message [3]. In this manner, watermarking is used to create a covert channel to transmit confidential information [4]. Digital audio watermarking involves the concealment of data within a discrete audio file.

II. PROPOSED BLOCK DIAGRAM

Audio watermarking will be implemented using the following steps

- Read the secret data to be hidden.
- Encrypt the secret data using PN sequence
- Read the audio signal in which data is to be hidden.
- Take DFT of the audio signal
- Hide the data in the audio at the High frequency DFT coefficients.
- Take IDFT

The proposed block diagram for audio watermarking is shown in fig 1.

The block diagram for watermark extraction is shown in fig 2.

A. Audio Input

MATLAB® audio support provides the ability to:

- Read and write audio files in common formats such as WAV, AVI, FLAC, MP3, and MPEG-4 AAC
- Playback and record audio files using the PC sound card

In this research, an audio file with .wav extension is considered as cover data and sampled at a rate of 44.1 kHz with 16 bit analog to digital converter.

B. Discrete Fourier Transform [5,6]

The Fourier transform is simply a method of expressing a function in terms of the sum of its projections onto a set of basis functions. The DFT is a specific kind of discrete transform, used in Fourier analysis. It transforms one function into another, which is called the frequency domain representation, or simply the DFT, of the original function (which is often a function in the time domain). The DFT requires an input function that is discrete.

The DFT transforms N complex numbers into another sequence of complex numbers, which is defined by F(k) where k=1,1 …N-1
The inverse Fourier Transform can be calculated using the following equation.

\[ F(k) = \sum_{n=0}^{N-1} f(n)e^{-j\frac{2\pi kn}{N}} \]  

The input, i.e., spectrum of the original audio signal from the song “Maa Tujhe Salam” has been considered as cover audio. The “cameraman image from MATLAB directory is considered as watermark. The cover audio undergoes Discrete Fourier Transform and got converted back into frequency domain. The maximum energy components lie near the origin and low energy components lie at higher frequencies. Hence it has been decided to embed the watermark in the high frequency components of cover audio.

### A. Need for Normalization

Since the cover signal is the audio signal, the amplitude variations are within ±1V. Watermark is a grey level image with 256 quantization levels where the pixels values lies within the range [0, 255]. Hence it is necessary to normalize the grey values within ±1 V so that without affecting the magnitude spectrum, the stuffing of grey values can be implemented.

### B. Z- Score (ZS) Normalization[11]

Z-score normalization calculates the normalized scores using arithmetic mean and standard deviation of the given data. This method transforms the score ‘s’ to a distribution with mean ‘0’ and a deviation of 1. Let mean (s) denote the arithmetic mean of score range and std (s) denote the standard deviation operators, respectively. Then the normalized score (N) is calculated as

\[ N = \frac{s - \text{mean}(s)}{\text{std}(s)} \]

After normalization, the high frequency components of cover audio are replaced with normalized values of watermark pixel grey values. After completing the process, IDFT is taken for the combined data so that the signal comes back to time domain. This signal almost resembles the original signal as shown in fig.3.

For the extraction of the watermark, the DFT of the received audio has been taken and from the high frequency components, the watermark gets extracted as shown in fig.3

### C. PN sequence based Encryption[13]

PN sequences look like random noise but they are not purely random in nature. When we multiply the PN sequence with an image, the resulting image resembles noise. However, during decryption, this noise-like signal can be used to exactly reconstruct the original image by multiplying it by the same pseudorandom sequence. Using this scheme, the initial seed state which is nothing but the key is only needed to generate exactly the same sequence of length. We have decided the following algorithm to generate the seed state for the PN sequence generator.

- Decide the password.
- Generate the seed state from the password

Since the seed state is known only to the authorized user, an intruder trying to access the database will not be able to decrypt the templates.
The encryption scheme using PN sequence is done as given below [13]:

- Let the watermark be image K
- Generate the Pseudo image R using the seed state. The size of R image should be same as image K
- Do the bit Ex-OR operations on both the images and find out the resultant image C as

$$C = \text{bit XOR}(R, K)$$  \hspace{1cm} (4)

The decryption scheme using XOR(R, K) is done with the help of following steps [13]:

- Generate the Pseudo image R using the seed state. The size of R image should be same as extracted image C
- Do the bit Ex-OR operations on both the images R and C
- The watermark is extracted as

$$W = \text{bit XOR}(R, C)$$  \hspace{1cm} (5)

The watermarked image is encrypted using PN sequence and the resultant encrypted image is transmitted to the receiver along with cover audio. At the receiver side, first the encrypted watermark is extracted from the cover audio and then decryption using PN sequence is done on the extracted image as shown in fig.4

IV. RESULT ANALYSIS

For all practical purposes, it is preferable to have a quantitative measurement to provide an objective judgment of the extracting fidelity. This is done by calculating the following parameters in each case. Results, thus obtained, have been tabulated. For performance evaluation, an audio clipping with one minute duration is considered as cover data.

A. Mean Square Error

Mean Square Error (MSE) [14], first introduced by C. F. Gauss, serves as an important parameter in gauging the performance of the watermarking system. MSE is essentially a signal fidelity measure [15, 16]. The goal of a signal fidelity measure is to compare two signals by providing a quantitative score that describes the degree of similarity/fidelity. Suppose that

$$x = \{x_i \mid i = 1, 2, N\}$$ \hspace{1cm} and

$$y = \{y_i \mid i = 1, 2, N\}$$ are two finite-lengths, discrete signals, The MSE between the signals are given by the following formula:

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (X_i - Y_i)^2$$  \hspace{1cm} (6)

where, N is the number of signal samples, $x_i$ is the value of the i\textsuperscript{th} sample in x. $y_i$ is the value of the i\textsuperscript{th} sample in y.

B. Entropy

The entropy H of a discrete random variable X is given as

$$H(x) = E(I(x))$$  \hspace{1cm} (7)

Here ‘E’ is the expected value, and ‘I’ is the information content of X. If ‘p’ denotes the probability mass function of X then the entropy can be explicitly written as:

$$H(x) = \sum_{i=1}^{n} p(x_i) \log_2 p(x_i)$$  \hspace{1cm} (8)
C. Moment

In image processing, computer vision and related fields, an image moment is a certain particular weighted average (moment) of the image pixels intensities, or a function of such moments, usually chosen to have some attractive property or interpretation. For a gray scale image with pixel intensities I(x, y), raw image moments are calculated by:

\[ M_{ij} = \sum_x \sum_y x^i y^j I(x, y) \]  

(9)

Table I. Performance Evaluation (1minute audio)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Simple Watermark</th>
<th>Encrypted watermark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entropy</td>
<td>(7)</td>
<td>(7)</td>
</tr>
<tr>
<td>Moment</td>
<td>(1.17 \times 10^9)</td>
<td>(1.17 \times 10^9)</td>
</tr>
<tr>
<td>MSE</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Embedding Time</td>
<td>1.16s</td>
<td>1.49s</td>
</tr>
<tr>
<td>Extraction Time</td>
<td>0.18s</td>
<td>0.77s</td>
</tr>
<tr>
<td>MSE in Audio</td>
<td>(2.2 \times 10^4)</td>
<td>(5.11 \times 10^4)</td>
</tr>
</tbody>
</table>

V. CONCLUSION

In this paper, an audio watermarking scheme based on Discrete Fourier Transform has been implemented. To enhance the security of the system, a PN sequence based encryption has also been considered. The proposed algorithms were implemented using MATLAB 2010a on an Intel core i3 based platform. For a cover audio of size one minute, the average time taken for extraction simple watermarking with an image of size \(256 \times 256\) was 0.18 seconds and 0.77 seconds using the PN sequence scheme. In terms of performance, the calculated values of MSE, entropy and moment are found to be the same for both the schemes. Since the MSE between original audio and watermarked audio was very high, the clarity of the audio at the receiver decreased. As a future scope, this MSE can be reduced with the help of some other watermarking techniques.

REFERENCES


