EMPLOYING PSYCHOACOUSTIC MODEL FOR DIGITAL AUDIO WATERMARKING

CHAI SIEW LI

THESIS SUBMITTED IN FULFILMENT OF THE DEGREE OF COMPUTER SCIENCE

FACULTY OF COMPUTER SYSTEM AND SOFTWARE ENGINEERING

2013

UNIVERSITI MALAYSIA PAHANG

BORANG PENGESAHAN STATUS TESIS

JUDUL: …….

SESI PENGAJIAN: ……

SAYA ……..……(HURUF BESAR)

Mengaku membenarkan tesis/laporan PSM ini disimpan di Perpustakaan Universiti Malaysia Pahang dengan syarat-syarat kegunaan seperti berikut:

1. Tesis/Laporan adalah hakmilik Universiti Malaysia Pahang.
2. Perpustakaan Universiti Malaysia Pahang dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institut pengajian tinggi.
4. **Sila tandakan (√)

☐	SULIT	(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972) *
☐	TERHAD	(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan) *
☐	TIDAK TERHAD	

Disahkan Oleh

……. ……

Alamat tetap: Penyelia:

Tarikh: …… Tarikh:……..

*Sila lampirkan surat daripada pihak berkuasa/organisasi berkenaan dengan menyatakan sekali sebab dan tempoh tesis/laporan ini perlu dikelaskan sebagai SULIT atau TERHAD.

DECLARATION

I hereby declare that the work in this thesis is my own except for quotations and summaries which have been duly acknowledged.

Date: 1/05/2013

Name: Chai Siew Li
Matric No.: CB10017

SUPERVISOR DECLARATION

I hereby declare that I have read this thesis, Employing Psychoacoustic Model for Digital Audio Watermarking and in my opinion this thesis is sufficient in terms of scope and quality for the award of the degree of Bachelor of Computer Science (Software Engineering).

.....

Date: 16/05/2013 (Professor Dr. Jasni Binti Mohamad Zain)

ACKNOWLEDGEMENTS

First, I would like to thanks God for all his blessing for giving me patience and good health throughout the duration of this research. I wish to express my eternal gratitude and sincere appreciation to my research supervisor, Professor Dr. Jasni Binti Mohamad Zain. I am very fortunate to have this opportunity with the guide of Dr. Jasni throughout the research although she was very busy. Furthermore, I am very grateful to Dr. Eric Liew Siau Chuin who gave me some suggestion on working out this research. Finally, I am grateful to my family members that support me all the way when I am facing difficulty when undergo this research.

ABSTRACT

This thesis discusses about digital audio watermarking by employing psychoacoustic model to make the watermarked signal inaudible to the audience. Due to the digital media data able to distribute easily without losing of data information, thus the intellectual property of musical creators and distributor may affected by this kind of circumstance . To prevent this, we propose the usage of spread spectrum technique and psychoacoustic model for embedding process, zero-forcing equalization and detection and wiener filtering for extracting process. Three samples of audio signal have been chosen for this experiment which are categorized as quiet, moderate, and noise state signal. The findings shows that our watermarking scheme achieved the intended purposes which are to test digital audio watermarking by employing psychoacoustic model, to embed different length of messages to test on accuracy of extracted data and to study the suitability on using hash function for verification of modification attacks.

ABSTRAK

Tesis ini membincangkan tentang audio digital watermarking dengan menggunakan model psychoacoustic untuk membuat isyarat tera air didengar kepada penonton. Oleh kerana data media digital boleh mengedar mudah tanpa kehilangan maklumat data, dengan itu harta intelek pencipta muzik dan pengedar boleh dipengaruhi oleh jenis ini keadaan. Untuk mengelakkan ini, kami mencadangkan penggunaan penyebaran teknik spektrum dan model psychoacoustic untuk proses menerapkan, sifar memaksa penyamaan dan pengesanan dan sosis penapisan untuk proses mengekstrak. Tiga sampel isyarat audio telah dipilih untuk percubaan ini yang dikategorikan sebagai tenang, sederhana, dan isyarat keadaan bunyi. Hasil kajian menunjukkan bahawa skim kami watermarking mencapai tujuan yang dimaksudkan iaitu untuk menguji Mekatronik audio digital dengan menggunakan model psychoacoustic, untuk menanamkan panjang yang berbeza mesej untuk menguji ketepatan data diekstrak dan mengkaji kesesuaian menggunakan fungsi hash untuk pengesahan serangan pengubahsuaian.

TABLE OF CONTENTS

DECLARATION II

ACKNOWLEDGEMENTS IV

ABSTRACT V

LIST OF TABLES IX

LIST OF FIGURES X

LIST OF ABBREVIATIONS XI

CHAPTER 1 INTRODUCTION 1

1.1 Introduction 1

1.2 Problem Statement 2

1.3 Objectives 3

1.4 Scope of Study 4

1.5 Thesis Organization 5

CHAPTER 2 LITERATURE REVIEW 6

2.1 Watermarking 6

2.1.1 Cryptography versus Steganography 8

2.1.2 Steganography versus Digital Watermarking 9

2.2 Working Domain 9

2.3 Watermarking Technique Requirements 11

2.4 Characteristic of Audio Watermarking Techniques 12

2.5 Audio Watermarking Technique 12

2.6 Existing Research Technique 18

2.7 Analysis on Existing Research Techniques 23

2.7.1 Spread-Spectrum Watermarking 25

2.7.2 Watermarking Shaping 25

CHAPTER 3 METHODOLOGY 29

3.1 Introduction 29

3.2 Methodology 29

3.2.1 Watermark Embedding.....30

3.2.2 Watermark Extracting.....32

3.2.3 Hashing function and Watermark Evaluation 33

3.3 Hardware and Software.....34

CHAPTER 4 DESIGN AND IMPLEMENTATION 36

4.1 Introduction 36

4.2 Embedding Process.....36

4.3 Extracting Process 41

4.4 Hash Sum Function.....48

4.5 Implementation.....49

CHAPTER 5 RESULT AND DISCUSSION.....50

5.1 Introduction 50

5.2 Result Analysis 50

5.3 Research Constraints.....61

CHAPTER 6 CONCLUSION 62

6.1 Conclusion.....62

6.2 Future Direction.....63

REFERENCES.....64

LIST OF TABLES

Table Number **Page**

2.1 :Evaluation of Steganography requirements associated with cover medium 7
2.2 :Comparison of various watermarking techniques 13
2.3 :Advantages and Disadvantage of the techniques 14
2.4 :Summarized result based on robustness 20
2.5 :Analyze result for existing researches 24
5.1: Result of Watermark Embedding and Extracting by using waveform A 54
5.2: Result of Hash Watermark Embedding and Extracting by using waveform A 56
5.3: List of Legend Representation 56
5.4: Result of Watermark Embedding and Extracting by using waveform B 57
5.5: Result of Hash Watermark Embedding and Extracting by using waveform B 59
5.6: Waveform A and Waveform B 60

LIST OF FIGURES

Figure Number **Page**

2.1 :Types of watermarking techniques....6
2.2 :ETAS Model....7
2.3 :Comparison of text merged with in terms of size....8
2.4 :Magic Triangle....12
2.5 :Procedure of encoding....19
2.6 :Process of audio watermarking embedding....21
2.7 :Experiment result of audio watermark embedding simulation.....21
2.8 :Extractive process of watermark....22
2.9 :Experiment result of audio watermark extraction simulation.....22
2.10 :Embedding process of the piracy watermark....23
2.11 :Extractive process of genuine watermark....23
2.12 :Typical embedder of the spread-spectrum watermarking scheme.....25
2.13 :Temporal masking in the human auditory system (HAS)....26
3.1 :Watermark Embedding Scheme....30
3.2 :Watermark Extracting Scheme....33
4.1 :Audio signal and watermark signal....40
4.2 :A few samples of the audio, watermark, and watermarked signals....40
4.3 :Equalized audio signal and the modulated signal....44
4.4 :Power spectral density of equalized audio signal and modulated signal....44
4.5 :The wiener equalized audio signal and the wiener modulated signal....47
4.6 :The power spectral density of the wiener equalized audio signal and wiener modulated signal....47
4.7 :The Graphical User Interface (GUI) for proposed watermark scheme....49

LIST OF ABBREVIATIONS

ASCII – American Standard Code for Information Interchange
CD – Compact Disc
dB – Decibel
DCT – Discrete Cosine Transform
DFT – Discrete Fourier Transform
DSP – Digital Signal Processing
DWT – Discrete Wavelet Transform
FFT – Fast Fourier Transform
GA – Genetic Algorithm
GUI – Graphical User Interface
HAS – Human Auditory System
Hz – Hertz
IFFT – Inverse Fast Fourier Transform
IT – Information Technology
Kbps – Kilo Byte per Second
kHz – Kilo hertz
LSB – Least Significant Bit
MP3 – MPEG 1 Audio Layer III
SHA – Secure Hash Algorithm
SNR – Signal-to-Noise-Ratio
WAV – Waveform Audio Form

CHAPTER 1 INTRODUCTION

1.1 Introduction

In twentieth century, it is the rise of the digital age and there are many kind of technology had been founded in this digital world. When it comes to digital age, we can say that almost every house, company, restaurant and shop has at least one computer or laptop and everything is digitalized or in digital format. Digital format is a format system that uses binary code which is 0 and 1 only to interpret data received and data to be sent. Other than that, to fight for enrichment, most of the IT expert spends a lot of time in order to make the evolution of technology. Recently, the rapid development of the internet had influence the economy in many aspects such as music production and film production. All the multimedia data in digital format like images, audio, and video can be copied and compressed easily without decrease the fidelity of the data. Hence, with the ease of distribution and duplication, most of the people download illegal copy of the digital media products from the internet.

First of all, there are two famous information hiding techniques which frequently used by the people now, whose are Steganography and Watermarking. Steganography is the art and sciences of writing hiding information in a way that prevents people from detect the hidden messages (Krenn, 2004). Whereas, watermarking is the process of embed a message or a new signal into a host signal. Usually, steganography methods are not robust or with only limited robustness against modification or transmission of the data. On the other hand, watermarking has the additional notion of resilience against attempts to remove the hidden data and other possible attacks. Some of the people may confuse that whether watermark and digital watermark are the same. In fact, when it comes to digital watermark it means that the

watermarked data is in digital format. Digital watermarking is a technique by which copyright information is embedded into the host signal in a way that the embedded information is imperceptible, and robust against intentional and unintentional attacks. (Wang, Niu, Yang, 2009)

By now, the technology of digital audio watermarking will be discovered in this research. Digital Audio Watermarking is a technology to hide information in an audio file without the information being audible to the listener, and without affecting in any way the audio quality of the original file. In addition, to increase the quality of watermarked audio, psychoacoustic model will be employed as psychoacoustic model is a model that designed to take advantage of the masking effect in human hearing which is HAS (Naveen, Jhansi rani, 2010).

1.2 Problem Statement

In the new era of technology world, there are a lot of digital media had been revealed nowadays. There are various types of digital media such as digital video, digital audio, digital images and others. The fast growth of the Internet and the maturity of audio compression techniques enable the promising market of on-line music distribution (Wu, Su, Jay Kuo, 2000). However, with the digital technology today also allows lossless data duplication, illegal copying and distribution would be much easier than before. Consequently, the intellectual property of musical creators and distributor may affected by this kind of circumstance. The necessity of protecting copyrighted audio data has significantly increased. Thus, digital audio watermarking is recommended or promoted to be used for copyright protection and owner authentication.

Copyright is the most concern issue currently as there are a lot of people like to copy the intellectual property illegally and this kind of action is known as piracy. In Malaysia, copyright protection is governed by the Copyright Act 1987 (Perbadanan Harta Intelek Malaysia, 2012). Copyright is the legal protection and exclusive right extended to those who create an original work through sufficient skill and effort for a specific period which depends on the type of works. Apart from that, the protection subsists automatically once the work is published or stored in a material form. In

Malaysia, the works eligible for protection are literacy works, musical works, artistic works, films, sound recordings, broadcasts and derivative works (Perbadanan Harta Intelek Malaysia, 2012). If a work qualified for copyright protection, it gives the owner the right to control the use of his work as well as to prevent it from being copied and distribute to others without owner permission. In term of legal copy is the owner given the right to the person who want to use or copy his work and that person only allowed to use it personally and is not allowed to use it for trading or others thing that is related to profit. By using digital watermarking method, we are able to trace and tackle the people who make the illegal copy no matter it is make a copy into CD or distribute to the internet. Once we extract information from the illegal copy which is watermarked copy, we are able to prove whether it is an authorized copy.

Moreover, owner authentication is uses a set of data or information that can be used to identify or prove the ownership. There will be some problems may arise when people are able to change the set of data or information easily in order to get the ownership. In digital watermarking, the embedded information may not be easily removed or changed as it required the secret key and it is robust to attacks.

1.3 Objectives

Below show the objectives of this research:

i. To test digital audio watermarking by employing psychoacoustic model
ii. To embed different length of messages to test on accuracy of extracted data
iii. To study the suitability on using hash function for verification of modification attacks

1.4 Scope of Study

Below show the scope of this research:

i. Spread spectrum technique and psychoacoustic model will be used as watermark embedding scheme
ii. Zero-forcing equalization and detection, and wiener filtering method will be used as watermark extracting scheme
iv. Hash sum function will be used to hash the watermark message

In this project, psychoacoustic modeling will be used and integrated with the Spread Spectrum technique. The purpose for this integration is to enhance the Spread Spectrum technique which is generated an audible noise and presence of strong distortion for the watermarked audio. In addition, by using spread spectrum method it will increase the robustness for watermark as the presence of pseudo-random generator. In addition, zero-forcing equalization and detection will be used for watermark extracting process and wiener filtering as an enhancement method that support the zero-forcing equalization and detection method. To study the suitability on using hash function for verification of modification attacks, hash sum function will be used to hash the watermark message and embed it into the audio signal and extract it from the watermarked signal. By using those techniques, watermark can be added and detected easily. I am here to study about the transparency for the audio watermarking and to investigate the audio quality between original audio and watermarked audio. Overall, in this digital audio watermarking system, users are able to watermark or detect watermarked audio. There is no boundary in user and place. It can be used by anyone at every place.

1.5 Thesis Organization

This thesis consist total of six (6) chapters. Chapter 1 will introduces the project which includes problem of background, objectives and scope of project. It will introduce digital audio watermarking, the background issues, and overview of the project. Chapter 2 will discuss on existing research or system that done by other researchers and explain which techniques or methods more suitable to be used in this project. Chapter 3 will discuss the overall approach and model flow of research which includes the method, technique or approach to be used in this project. It will explain the psychoacoustic model, spread spectrum technique, and ASCII encoding and decoding algorithm. Chapter 4 will discuss on development for the model flow of research which includes data collection, process, and analysis. Chapter 5 will discuss about the results after implement the proposed model flow for the research and discuss on the findings includes constraints and limitation for the proposed watermarking scheme. Chapter 6 will conclude the whole research that has been done and future direction.

CHAPTER 2 LITERATURE REVIEW

2.1 Watermarking

Mohanty, S.P. (1999) had summarized types of watermarking which shown in Figure 2.1.

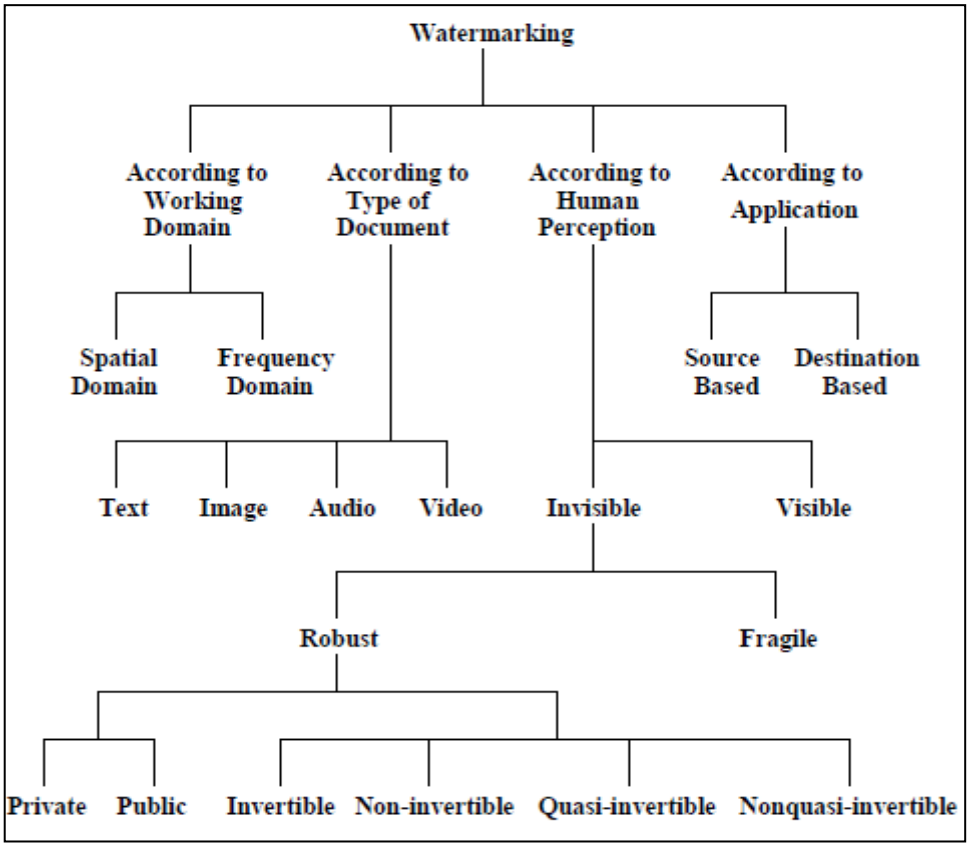

Figure 2.1 :Types of watermarking techniques

Watermarking can be applied to type of document like text, image, audio and video. Geetha and Vanitha Muthu (2010) had proposed Embedding Text in Audio

Signal (ETAS) model which is a very basic description of the audio steganography process in the sender side and receiver side. They use the LSB coding method to encode the message in audio signal. Figure 2.2 shows the ETAS model. Furthermore, table 2.1 shows the evaluation of Steganography requirements associated with cover medium that had been founded by them. Other than that, they also make a comparison between cover medium of text merged with in terms of size which shown in Figure 2.3. Anyway, they stated that ETAS model is able to ensure secrecy with less complexity at the cost of same memory space as that of encrypted text and the user is able to enjoy the benefits of cryptography and steganography combined together without any additional overhead.

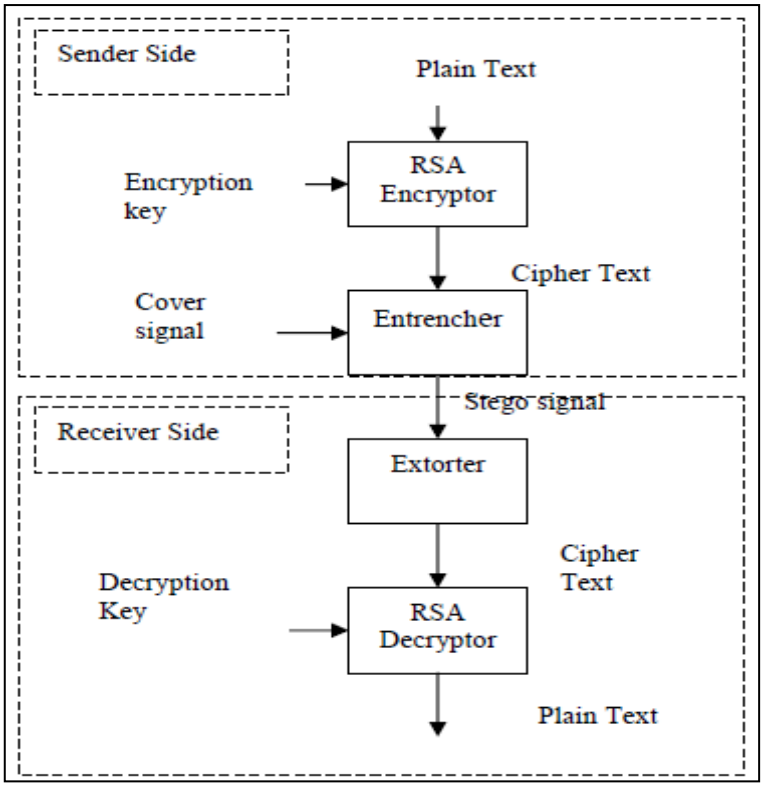

Figure 2.2 :ETAS Model

Table 2.1 :Evaluation of Steganography requirements associated with cover medium (Geetha and Vanitha Muthu, 2010)

	Plain Text	Image	Audio	Video
Invisibility	Medium	High	High	High
Payload Capacity	Low	Low	High	High
Robustness against Statistical Attacks	Low	Medium	High	High

Robustness against Text Manipulation	Low	Medium	High	High
Variation in file size	Medium	Medium	High	Medium

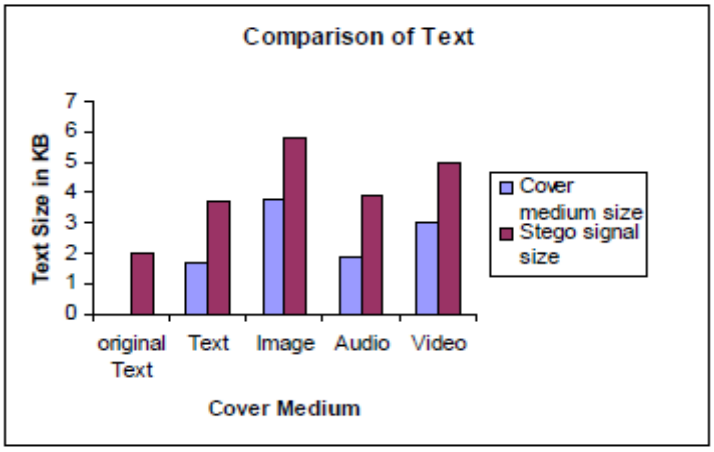

Figure 2.3 :Comparison of text merged with in terms of size (Geetha and Vanitha Muthu, 2010)

Based on the result on above, we can know that watermarking in audio is the better to embed the secret information. Hence, in this research the audio document type will be chosen for undergo watermarking process. Other than that, Agbaje, Akinwale and Njah (2011) also stated that audio signals are represented by much less samples per time interval, which indicates that the amount of information capacity that can be embedded robustly and inaudibly in audio files is much lower than the amount of information that can be embedded in visual files.

2.1.1 Cryptography versus Steganography

Amin *et al.* (2003) stated that the purpose of cryptography and steganography is to provide secret communication. However, steganography is not the same as cryptography. Cryptography hides the contents of a secret message from malicious people, whereas steganography even conceals the existence of the message. Basically, cryptography offers the ability of transmitting information between persons in a way that prevents a third party from reading it. In cryptography, the structure of a message is scrambled to make it meaningless and unintelligible unless the decryption key is

available. Whereas in steganography, it does not alter the structure of the secret message but hide it inside a cover-image so that it cannot be seen. In other words, cryptography is an encrypting process while steganography is a hiding process.

2.1.2 Steganography versus Digital Watermarking

Mohanty, S.P. (1999) said that steganography and digital watermarking primarily differ by intent of uses. A watermark can be perceived as an attribute of the carrier (cover). It may contain information such as copyright, license, tracking and authorship. While in case of steganography, the embedded message may have nothing to do with the cover. In steganography an issue of concern is bandwidth for the hidden message whereas robustness is of more concern with watermarking.

2.2 Working Domain

There are two type of working domain which is spatial and frequency domain. According to Singh (2011), spatial domain is manipulating or changing an image representing an object in space which uses statistical properties of each pixel and its immediate surrounding pixels in the host image, and also the statistical properties of the host image and that of the image to be embedded (watermark), as the pixels in the host image are replaced one by one by the pixels in the watermark image. One of the spatial domain techniques is LSB in which message is embedded in the least significant bit. While in transform domain, watermark is embedded in frequency domain of a signal such as DCT, DFT, DWT domain coefficients. Transform domain methods hide messages in significant areas of the host image which makes them more robust to attacks.

According to Cai and Chen (2011), the time domain masking effect refers to the weak voice fore-and-aft a stronger voice cannot be detected by the human's ears, that is, they would be "masking" out. The frequency domain masking effect, also known as the same time masking, indicating when two signals of nearly close frequency are concurrently working, the weak sound will be masked by strong sound and becomes unaware, and then the masking sound has some effect during the working of masking

effect, which is a much stronger masking effect. To ensure the efficient extraction of the watermarking, it can be reduce the energy of an embedded watermark as much as possible, which make it fully "submerged" in the carrier (cover) data.

Anyway, most of the audio watermarking is working in transform domain. One of the famous transform methods being used in audio watermarking is FFT. The Fourier Transform is a mathematical operation which converts time-domain signals to the frequency domain. Being able to visualize a signal in the frequency domain offers many benefits over time-domain representations. Frequency domain representations allow individual frequency components contained within a signal to be viewed including modulation sidebands, distortion effects and spurious frequency components. For example, figure 2.4 show a square wave is composed of a fundamental frequency and all it odd harmonics. The level for each harmonic decreased in amplitude with harmonic number. Viewed in the time domain representation, the square wave is gives no indication of its composition. Viewed in the frequency domain all frequencies components are displayed along with their relative amplitude (Aeroflex, 2012). Furthermore, the DFT is a version of the Fourier Transform which can be applied to sampled time-domain signals. The DFT produces a discrete frequency spectrum which is an amplitude levels at discrete frequencies. Whereas, the FFT is a development of the DFT which removes duplicated terms in the mathematical algorithm to reduce the number of mathematical operations performed. In this way, it is possible to use large numbers of samples without comprising the speed of the transformation. In other words, FFT is a fast version of DFT it transform speed is faster than the DFT.

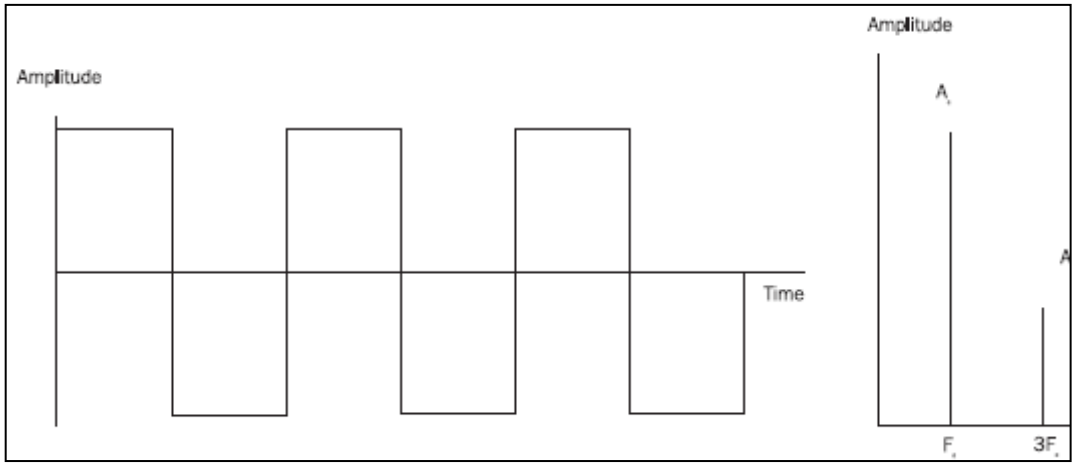

Figure 2.4 :Time and frequency domain representations of a square-wave signal

2.3 Watermarking Technique Requirements

According to Singh (2011), there are four watermarking technique requirements which are robustness, non-perceptibility, verifiability and security.

- Robustness

  Robustness refers to that the watermark embedded in data has the ability of surviving after a variety of processing operations and attacks. Then, the watermark must be robust for general signal processing operation, geometric transformation and malicious attack. The watermark for copyright protection does need strongest robustness and can resist malicious attacks, while fragile watermarking; annotation watermarking do not need resist malicious attacks.

- Non-perceptibility

  Watermark cannot be seen by human eye or not be heard by human ear, only be detected through special processing or dedicated circuits.

- Verifiability

  Watermark should be able to provide full and reliable evidence for the ownership of copyright-protected information products. It can be used to determine whether the object is to be protected and monitor the spread of the data being protected, identify the authenticity, and control illegal copying.

- Security

  Watermark information owns the unique correct sign to identify, only the authorized users can legally detect, extract and even modify the watermark, and thus be able to achieve the purpose of copyright protection.

Watermark information owns the unique correct sign to identify, only the authorized users can legally detect, extract and even modify the watermark, and thus be able to achieve the purpose of copyright protection.

2.4 Characteristic of Audio Watermarking Techniques

Audio watermarking techniques have three characteristics which are inaudibility, robustness and bit rate. Figure 2.4 shows the magic triangle which represent those characteristics (Cvejic, N.,2004). Inaudibility is presented in the upper portion of the triangle as it is the top requirement of audio watermarking process. While, the other two requirements cannot achieve together so it presented in two corners of the triangle respectively. For example, if the data rate is high then the robustness is low while if the robustness is high then the data rate is low.

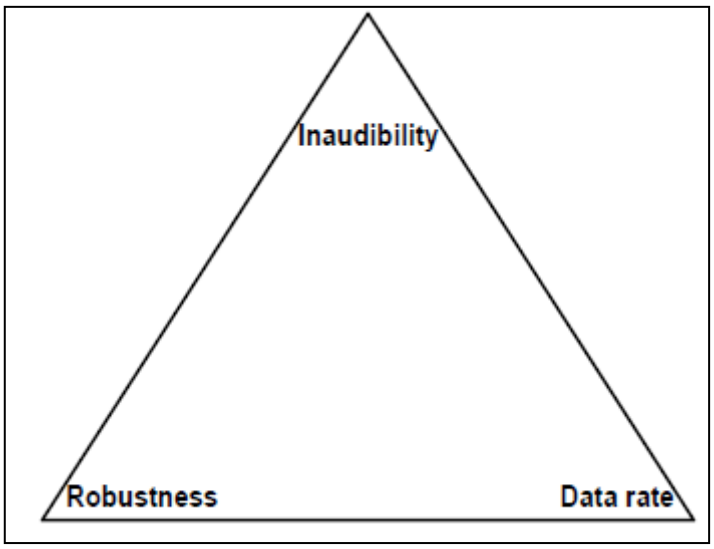

Figure 2.5 :Magic Triangle (Cvejic, N., 2004)

2.5 Audio Watermarking Technique

Kim, Choi, Seok and Hong (2004) state that there are four categories scheme for audio watermarking those are spread-spectrum scheme, two-set scheme, replica scheme and self-marking. Spread-spectrum embeds pseudo-random sequence and detects by calculating auto-correlation. While two-set scheme which is also known as patchwork scheme exploits the differences between two or more sets. Replica scheme uses the replica of the original audio clips both in embedding and detection phases. Echo hiding is a good example of replica scheme. Last one is self-marking scheme which can be used especially for synchronization or for robust watermarking, for example, against time-scale modification attack. Time-scale modification refers to the process of either compressing or expanding the time-scale of audio. The basic idea of the time-scale

modification watermarking is to change the time-scale between two extrema (successive maximum and minimum pair) of the audio signal. Such four seminal works have improved watermarking schemes remarkably. Pseudo-random sequence has statistical properties similar to those of a truly random signal, but it can be exactly regenerated with knowledge of privileged information. Good pseudo-random sequence has a good correlation property such that any two different sequences are almost mutually orthogonal. Thus, cross-correlation value between them is very low, while auto-correlation value is moderately large. In addition, Table 2.2 shows the comparison of various watermarking technique written by Agbaje, Akinwale and Njah (2011).

Table 2.2 :Comparison of various watermarking techniques (Agbaje, Akinwale and Njah, 2011)

Techniques	Advantages	Disadvantages
Spread spectrum	Easy to implement	It requires time consuming psychoacoustic shaping to reduce audible noise, susceptible to time-scale modification attacks and difficulty in synchronization
Quantization	Easy to implement and robust against noise to a particular threshold	Not robust against attacks
Two set	-	-
Replica method	Immunity to synchronization attacks	-
Echo hiding	imperceptibility	High complexity due to ceptrum or autoceptrum computation during detection and echo can be detected without prior knowledge

Moreover, according to Kiah et al. (2011), there are number of ways to hide the information or data into audio such as phase coding, spread spectrum, echo data hiding, patchwork coding, low-bit encoding and noise gate. Table 2.3 shows the advantages and disadvantages of the techniques that written by Kiah et al. (2011).

Table 2.3: Advantages and Disadvantage of the techniques

Approach	**Summary**	**Advantage and Disadvantage**
Low–bit Encoding	Low-bit encoding considered as the earliest techniques implemented in the information hiding of digital audio. It is the simplest technique to embed data into other data structures such as data of audio in image file or data of image in audio file. Low-bit encoding, can be done by replacing the LSB of each sampling point by a coded binary string (hidden data)	The major advantage of Low-bit encoding are: 1. High watermark channel bit rate 2. Low computational complexity of the algorithm compared with others techniques 3. No computationally demanding transformation of the host signal, therefore, it has very little algorithmic delay The major disadvantage is that the method are: 1. Low robustness, due to the fact that the random changes of the LSB destroy the coded watermark 2. it is unlikely that embedded watermark would survive digital to analogue and subsequent analogue to digital conversion
Phase Coding	Phase Coding watermarking works by substituting the phase of an initial audio segment	The major advantage of Phase Coding are: 1. Basic technique

	with a reference phase, this phase represents the hidden data. The phase of subsequent segments is adjusted in order to preserve the relative phase between segments	The major disadvantage is that the method are: 1. Phase coding method is a low payload because the watermark embedding can be only done on the first block. 2. The watermark is not dispersed over the entire data set available, but is implicitly localized and can thus be removed easily by the attackers
Spread Spectrum Technique	Spread spectrum (SS) is technique designed to encode any stream of information via spreading the encoded data across as much of the frequency spectrum as possible even though, there is interference on some frequencies, SS allows the signal reception.	The major advantage of Spread Spectrum are: 1. Difficult to detect and/or remove a signal 2. Provide a considerable level of robustness The major disadvantage is that the Spread spectrum are: 1. Spread spectrum technique used transform functions (e.g. DFT, DCT, or DWT) with appropriated inverse transform function, which can cause a delay. 2. Spread spectrum is not a visible solution for real time applications
Patchwork Coding	Patchwork Coding considered as one of the earliest generation for digital	The major advantage of Patchwork Coding are: 1. Patchwork based watermarking

	watermarking schemes. Patchwork Coding can be done via embedding the watermark in the audio using time domain or frequency domain. In the literature, several approaches of Patchwork Coding have been proposed on frequency domain using linear transformations, such as DWT, DFT and DCT. Frequency or time domain watermarking schemes directly tinker with sample amplitude of audio to embed the watermark	scheme has been confirmed as an valuable to those common signal processing operations, such as low-pass filtering, image/audio compression, and so on. The major disadvantage is that the Patchwork are: 1. An attack called "curve-fitting attack" has been successfully implemented for patchwork watermarking scheme. 2. Patchwork watermarking scheme is sensitive to various synchronization attacks
Echo technique	Echo technique embeds data into a host audio signal by introducing an echo; the hidden data can be adjusted by the two parameters: amplitude and offset, the two parameters represent the magnitude on time delay for the embedded echo, respectively. The embedding process uses two echoes with different offsets, one to represent the binary datum "One" and the other to represent the binary datum "Zero".	The major advantage of Echo are: 1. The main advantage of echo hiding is that the echo detection technique is easy to implement. The major disadvantage is that the echo hiding technique are: 1. More complicated computation is required for echo detection. 2. Echo hiding is also prone to inevitable mistakes, such as the echo from the host signal itself may be treated as the embedded echo. 3. If the echo added has smaller amplitude, then the cepstrum peak would be covered by the

		surrounding peaks to make the echo detection an arduous task to perform. A larger echo may increase the accuracy rate of detection but it also easily exposes the system to deliberate attacks, which then affects the sound quality
Noise Gate Technique	Noise gate technique is designed to be an alternative solution for the weakness in the previous approaches, this technique implanted in the time domain. This technique maintains a high quantity of data hidden side by side with robustness. Noise Gate Technique involve two steps approach, the first step, noise gate software logic algorithm has used to obtain a desired signal for embedding the secret message of the input host audio signal. In the second step, standard ith LSB layer embedding has been done for this desired signal by simply replaces the host audio signal bit in the ith layer with the bit from the watermark bit stream, if 16-bit per audio sample used, where (i=1,...,16).	The major advantage of Noise Gate Technique are: 1. High watermark channel bit rate 2. Low computational complexity of the algorithm compared with others techniques 3. No computationally demanding transformation of the host signal, therefore, it has very little algorithmic delay 4. Add level of complexity against Stego-Only Attack and Known Message Attack The major disadvantage is that the method are: 1. Fair robustness 2. Noise Gate technique is weak against Known Cover Attack, Known Chosen Cover or Chosen Message and Known Stego Attack

2.6 Existing Research Technique

There are a lot of researches have been done regards audio watermarking. One of the researches carried by Bassia, Pitas and Nikolaidis (2011) is about the robust audio watermarking in the time domain. They embed a watermark in the time domain of a digital audio signal by slightly modifying the amplitude of each audio sample, amplitude modification. Their watermarking scheme is statistically imperceivable and resists MPEG2 audio compression plus other common forms of signal manipulation, such as cropping, time shifting, filtering, re-sampling and re-quantization. However, their method is not robust to more sophisticated attacks like the watermark in a signal that has been subject to a change in the time scale.

Fallahpour and Megias (2012) have done a research on high capacity robust audio watermarking scheme which based on FFT and linear regression. The main idea is to divide the FFT spectrum into short frames and change the magnitudes of the selected FFT samples using linear regression and the average of the samples of each frame. Furthermore, linear regression tends to help on minimize the alterations of FFT samples which results in better transparency. The findings proved that their watermarking scheme has a high capacity (0.5 to 2.3 kbps), without significant perceptual distortion (Objective Difference Grade is about 1) and provides robustness against common audio signal processing such as echo, added noise, filtering and MP3 compression. Anyway, there are modern attacks not often considered like compression, channel fading, jitter and packet drop those particularly relevant in various networks such as GSM and the Internet.

Ketcham and Vongpradhip (2007) proposed an audio watermarking and robust algorithm using DWT - based GA. Their experimental result revealed that the concept of existing algorithm has poor performance in resist type of attacks and it can be further enhanced by the combination of quality and robust.

Zhang, Xu, and Yang (2012) had presented a paper about the robust and transparent audio watermarking based on improved spread spectrum and psychoacoustic masking. They propose an extended improved spread spectrum modulation to reduce

the host interference in the context of watermark decoding with unmatched filtering. Furthermore, they employ psychoacoustic model to obtain good transparency which suppress the embedding distortion below the quantization noise levels of the audio coding scheme. By the way, their findings also shown that the proposed scheme provides high audio fidelity and is robust against most of common attacks.

Zhao, Guo, Liu and Yan (2011) had proposed an audio watermarking scheme which using spread spectrum technique to embed watermark message and psychoacoustic auditory model to retain the perceptual quality of the audio signal which resistant to diverse removal attacks, either intentional or unintentional. The experimental results showed that their watermarking scheme is robust to common signal processing attacks. Figure 2.5 is their complete watermarking scheme.

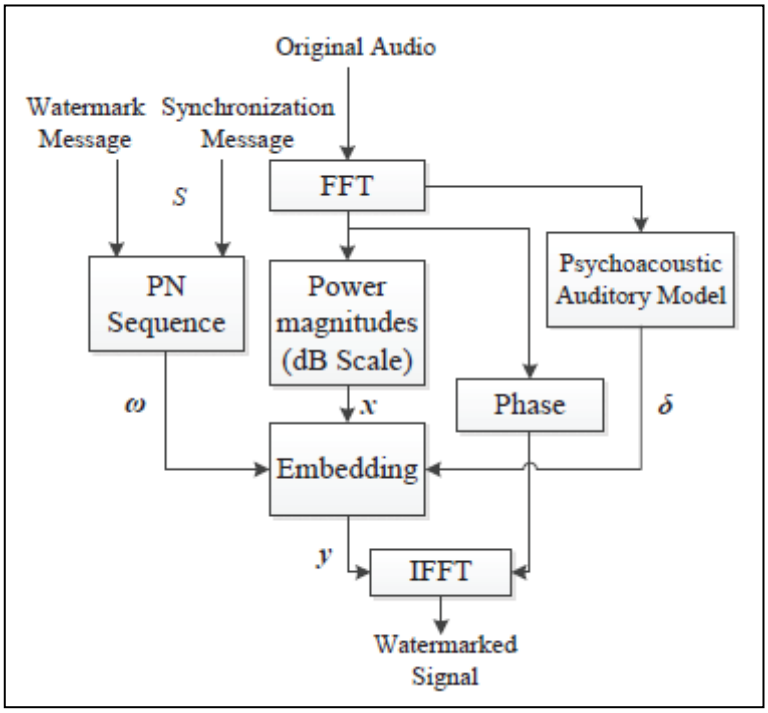

Figure 2.6 :Procedure of encoding

Baranwal and Datta (2011) proposed a robust spread spectrum based audio watermarking scheme using DWT. They use a peak detection algorithm to obtain high robustness. They use blind watermarking technique for copyright protection and content authentication for digital content. At the end, the result shows that their proposed watermarking scheme provides better performance against signal processing attacks like re-sampling, re-quantization, MP3 compression, and noise addition.

Maha, Maher and Chokri (2008) proposed the combination techniques of DWT, Neural Network, Human Psychoacoustic Model and Hamming, the frequency techniques, and the temporal Modified Least Significant Bit technique achieve better inaudibility performance then the combination techniques of DWT, Neural Network, and Hamming. The later combination techniques are less robust and the Modified LSB technique possesses the lowest robustness capability. The Hamming code permits to correct one error in 8 bits and help then to overcome the part of the corruption of the watermark. They also found that, in wavelet domain which guaranteed higher imperceptible capability by exploited the frequency masking in Human Psychoacoustic Model 1 that proposed in the paper. Table 2.4 is their summarized result based on robustness.

Table 2.4 :Summarized result based on robustness

(Maha, Maher and Chokri, 2008)

Scheme/NC	Co/Dec MP3, Silence	Noise Brown, White	Low-pass filter	Dynamic changes, Hiss reduction, Notch Filter, Conv1,Conv2
DWT, NN, HPM, Hamming	X	X		X
DWT, NN, Hamming	X			X
DCT, NN, HPM, Hamming	X	X	X	X
DCT, NN, Hamming	X	X		X
Modified LSB	X	X		

Based on their result on above, it prove that the Human Psychoacoustic Model increase the robustness of audio watermarking.

Cai and Chen (2011) proposed a WAV format audio digital watermarking algorithm based on HAS. Figure 2.7 show the process of audio watermarking embedding and Figure 2.8 show the experiment result of audio watermark embedding simulation achieved in the environment of Matlab 7.0 programming. While Figure 2.9

show the extractive process of watermark and Figure 2.10 show the experiment result of audio watermark extraction simulation achieved in the environment of Matlab 7.0 programming.

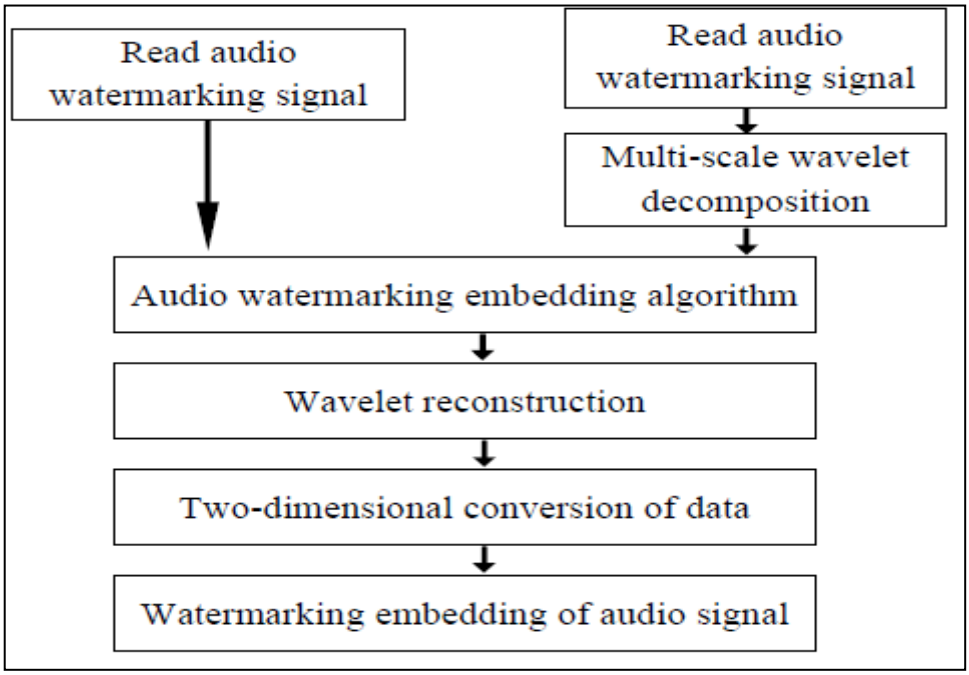

Figure 2.7 Process of audio watermarking embedding

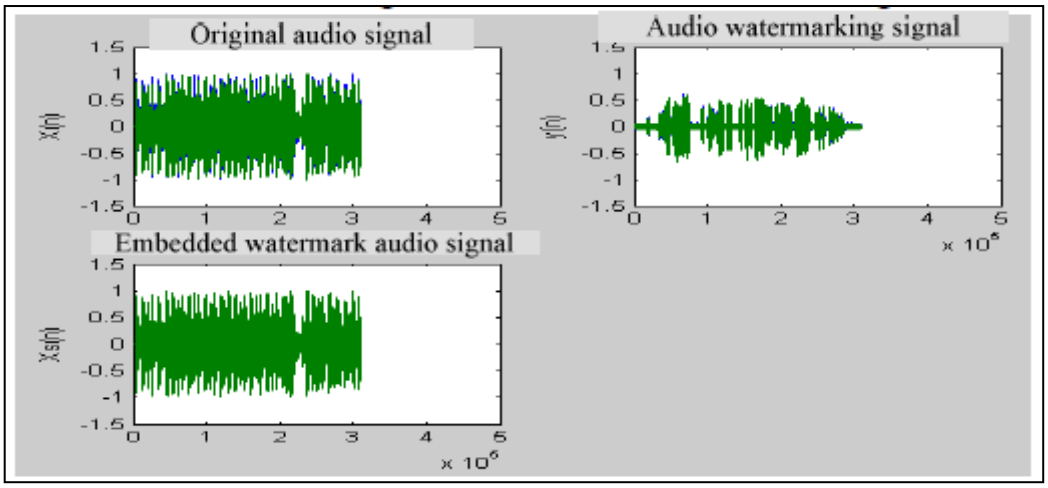

Figure 2.8 :Experiment result of audio watermark embedding simulation achieved in the environment of Matlab 7.0 programming.

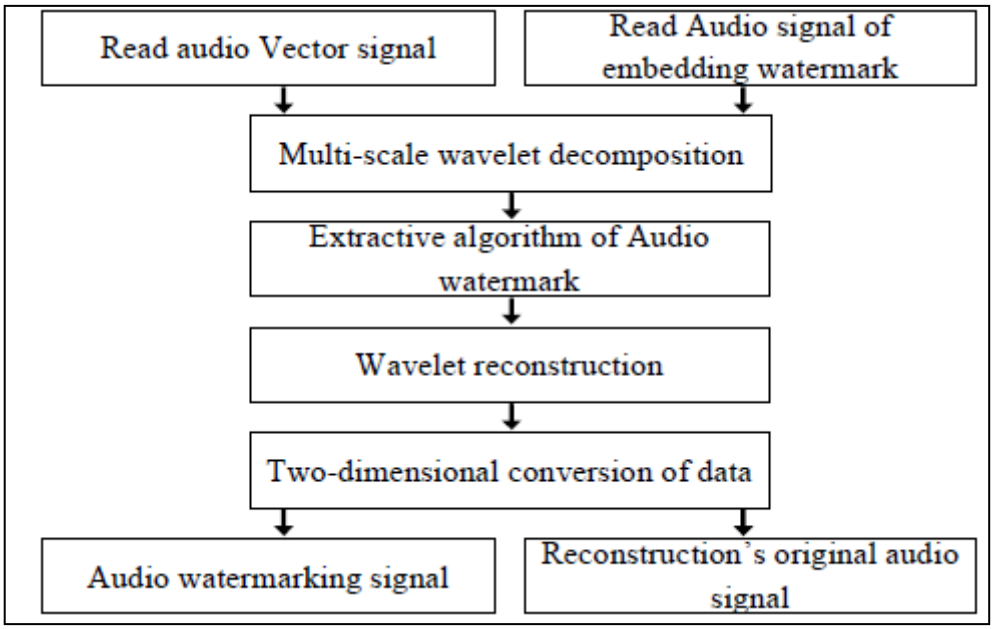

Figure 2.9 :Extractive process of watermark

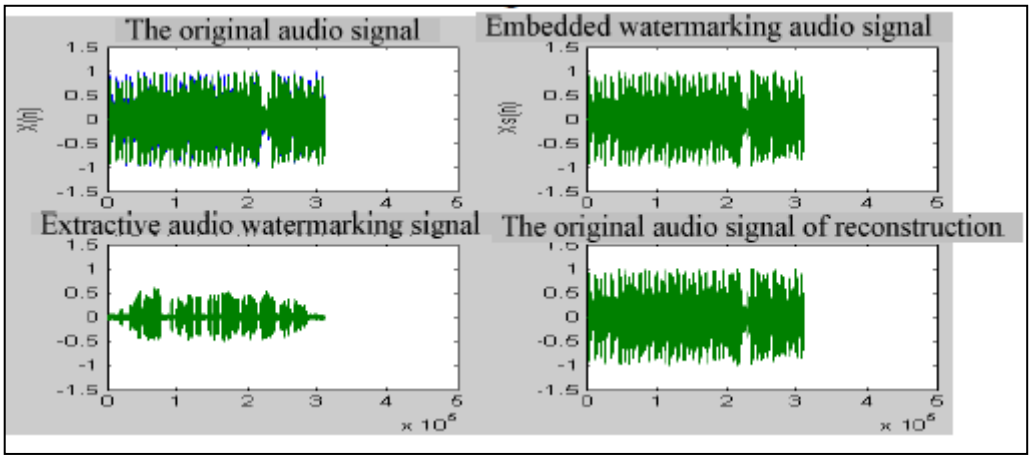

Figure 2.10 :Experiment result of audio watermark extraction simulation achieved in the environment of Matlab 7.0 programming.

They had stated that there is a possibility of attacks which is the pirates tends to embed the pirated watermark again to prove them as the owner. Hence, they make an experiment again and the experiment result is show in figure 2.11, the simulative embedding process of the piracy watermark and figure 2.12, extractive process of piracy watermark.

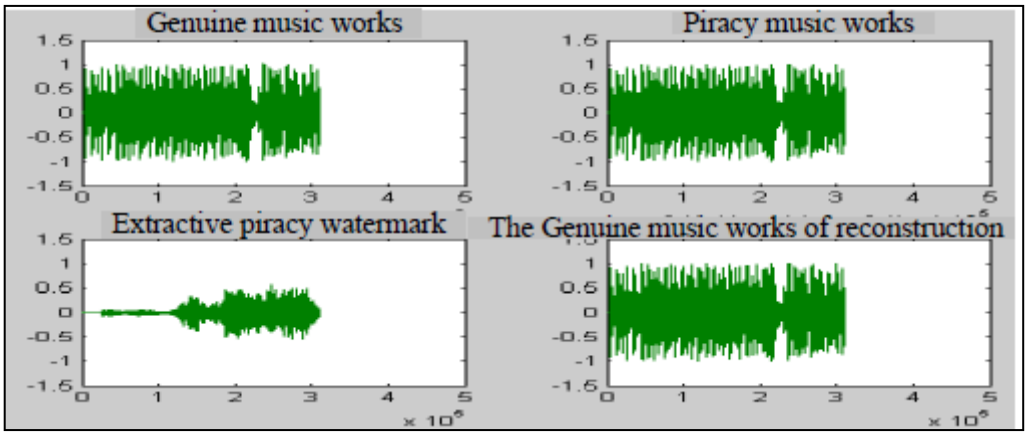

Figure 2.11 Embedding process of the piracy watermark

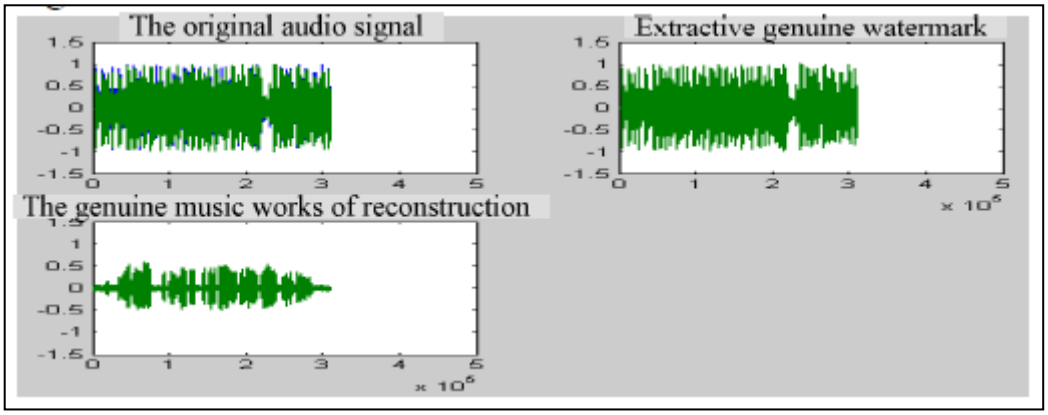

Figure 2.12 :Extractive process of genuine watermark

Experimental results show that the algorithm has good imperceptibility and robustness, with simple algorithm and easy reach ability, and has a good prospect in copyright protection of digital audio works and voice signals' covert communication.

2.7 Analysis on Existing Research Techniques

Based on the existing researches that being discusses on section 2.6 and Table 2.5 which is the analyze result for existing researches, the spread-spectrum watermarking and psychoacoustic model are the most popular among the researchers. A brief description for spread-spectrum watermarking and watermarking shaping which includes psychoacoustic model and synchronization will be described in section 2.7.1 and section 2.7.2 (Kim, Choi, Seok and Hong, 2004).

Table 2.5 :Analyze result for existing researches

Author	Technique	Results
Ketcham and Vongpradhip (2007)	DWT - based GA	Concept of existing algorithm has poor performance in resist type of attacks and it can be further enhanced by the combination of quality and robust.
Maha, Maher and Chokri (2008)	combination techniques of DWT, Neural Network, Human Psychoacoustic Model and Hamming, the frequency techniques, and the temporal Modified LSB technique	Human Psychoacoustic Model increase the robustness of audio watermarking.
Bassia, Pitas and Nikolaidis (2011)	Amplitude modification in time domain	Watermarked signal not robust to more sophisticated attacks like the watermark in a signal that has been subject to a change in the time scale.
Zhao, Guo, Liu and Yan (2011)	spread spectrum technique with psychoacoustic auditory model	Robust to common signal processing attacks
Baranwal and Datta (2011)	spread spectrum with discrete wavelet transform	Provides better performance against signal processing attacks like re-sampling, re-quantization, MP3 compression, and noise addition.
Cai and Chen (2011)	HAS	Good imperceptibility and robustness
Fallahpour and Megias (2012)	FFT and linear regression	Provide high capacity (0.5 to 2.3 kbps) and robustness against common audio signal processing such as echo, added noise, filtering and MP3 compression
Zhang, Xu, and Yang	extended improved spread spectrum modulation, and	Proposed scheme provides high audio fidelity and is robust against most of

(2012)	employ psychoacoustic model	common attacks

2.7.1 Spread-Spectrum Watermarking

Spread-spectrum watermarking scheme is an example of the correlation method which embeds pseudo-random sequence and detects watermark by calculating correlation between pseudo-random noise sequence and watermarked audio signal. This method is easy to implement, but has some serious disadvantages which is it requires time-consuming psycho-acoustic shaping to reduce audible noise and susceptible to time-scale modification attack which have difficulty of synchronization. A typical embedder of the spread-spectrum watermarking scheme is shown in figure 2.13.

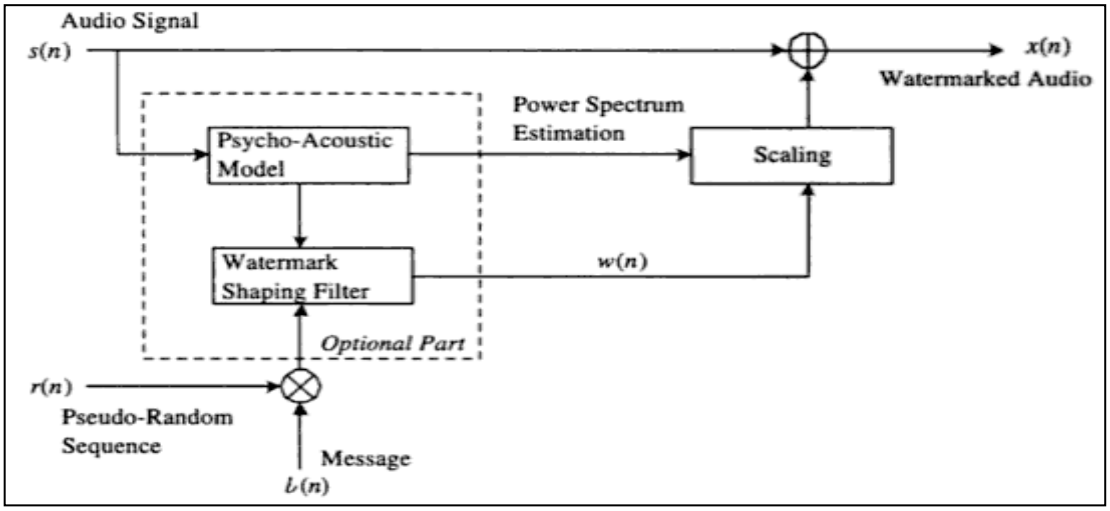

Figure 2.13 :Typical embedder of the spread-spectrum watermarking scheme.

2.7.2 Watermarking Shaping

Carelessly added pseudo-random sequence or noise to audio signal can cause unpleasant audible sound whatever watermarking schemes are used. Thus, just reducing the strength of pseudo-random sequence or noise cannot be the final solution. Because human ears are very sensitive especially when the sound energy is very low, even a very little noise with small value of pseudo-random sequence can be heard. Moreover, small pseudo-random sequence makes the spread-spectrum scheme not robust. Hence, there is

a solution proposed by Arnold and Schilz (2002), Bassia *et al.* (2001), Boney *et al.* (1996), Cvejic *et al.* (2001), and Cvejic and Seppanen (2002) which is perform watermark shaping based on the psychoacoustic model to ensure inaudibility of the watermark signal.

2.7.2.1 Human Auditory System

Before introduce the psychoacoustic model, we should have some knowledge on HAS as psychoacoustic model is a model that designed to take advantage of the masking effect in human hearing which is HAS (Naveen, Jhansi rani, 2010). Moreover, Fallahpour and Megias (2012) also said that human beings tend to be more sensitive towards frequencies in the range from 1 to 4 kHz. According to Kim, Choi, Seok and Hong (2004), there are two properties of the HAS dominantly used in watermarking algorithms are frequency (simultaneous) masking and temporal masking. Frequency masking is a frequency domain phenomenon where a low level signal, e.g. a pure tone (the maskee), can be made inaudible (masked) by a simultaneously appearing stronger signal (the masker), e.g. a narrow band noise, if the masker and maskee are close enough to each other in frequency. Whereas, temporal masking is the masking effects that appear before and after a masking signal has been switched on and off respectively. The duration of the pre-masking is significantly less than one-tenth that of the post-masking, which is in the interval of 50 to 200 milliseconds. Figure 2.14 shows the temporal masking in the human auditory system (Cvejic, 2004).

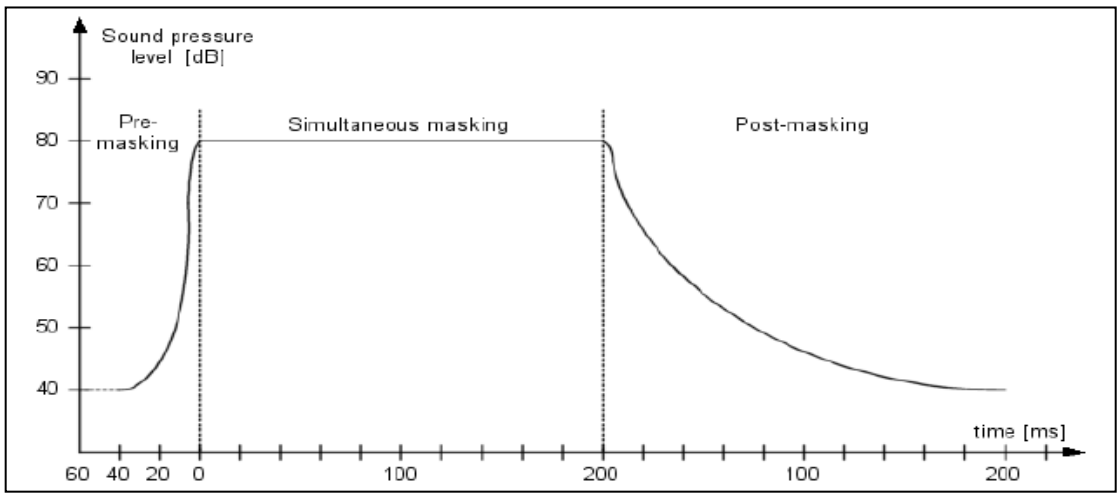

Figure 2.14 : Temporal masking in the human auditory system (HAS)

The power spectra of the received sounds are not represented on a linear frequency scale but on limited frequency bands called critical bands. (Zwicker and Fastl, 1999). Cvejic (2004) said that the auditory system is usually modelled as a band pass filter bank, consisting of strongly overlapping band pass filters with bandwidths around 100Hz for bands with a central frequency below 500 Hz and up to 5000 Hz for bands placed at high frequencies. If the highest frequency is limited to 24000 Hz, 26 critical bands have to be taken into account.

2.7.2.2 Psychoacoustic Model

Psycho-acoustic models for audio compression exploit frequency and temporal masking effects to ensure inaudibility by shaping the quantized noise according to the masking threshold. Psycho-acoustic model depicts the human auditory system as a frequency analyzer with a set of 25 band pass filters which is also known as critical bands. The required intensity of a single sound expressed in unit of decibel (dB) to be heard in the absence of another sound is known as threshold of audibility. Anyway, watermark shaping is a time-consuming task especially when we try to exploit the masking effects frame by frame in real-time because watermark shaping filter coefficients are computed based on the psycho-acoustic model. In this case, we have to use Fourier transform and inverse Fourier transform, and do frequency masking.

Maha, Maher and Chokri (2008) stated that psychoacoustic model takes advantage of the HAS's inability to hear quantization noise under conditions of auditory masking. This masking occurs whenever the presence of a strong audio signal makes a temporal or frequency neighborhood of weaker audio signals imperceptible. The psychoacoustic model analyzes the audio signal and computes the amount of noise masking available as a function of frequency. The encoder uses this information to decide how best to represent the input audio signal with its limited number of code bits.

2.7.2.3 Synchronization

Watermark detection starts by alignment of watermarked block with detector. Losing synchronization causes false detection. Time-scale or frequency-scale modification makes the detector lose synchronization. Thus, most serious and malicious attack is probably the de-synchronization. All the watermarking algorithms assume that any detector be synchronized before detection. Thus, we need fast and exact synchronization algorithms. Some watermarking schemes such echo hiding are rather robust against certain type of de-synchronization attacks. Such schemes can be used as a baseline method for coarse synchronization. For example, if an evidence of echo existence is identified, it shows that the block is near from synchronization. Synchronization code can be used to synchronize the onset of the watermarked block. Unfortunately, echo detection is considerably costly in terms of computing complexity.

CHAPTER 3 METHODOLOGY

3.1 Introduction

In this research, spread-spectrum technique has been selected for audio watermark system and it is integrated with psychoacoustic model to make the watermarked audio imperceptible to the listener. The reason for the integration is the spread spectra or spread waveform will cause the distortion of audio as the spread spectrum is some kind of white noise. Typically, the sampling frequency for audio files is 44.1 kHz which mean that the audio perform 44.1 kHz in one second. For example, an audio file with total of 10325 kHz with sampling frequency of 44.1 kHz, that means the audio file will be ended in 234 seconds. Hence, if we change the sampling frequency for audio files from 44.1 kHz to 48 kHz, then the audio file will be ended in 215 seconds. Overall approach details will be discussed on following section.

3.2 Methodology

The methodology for the digital audio watermarking system will be divided into two parts whose are watermark embedding and extracting process. In watermark embedding process, the watermark message which is the owner authentication information such as name will be embedded into the audio. While the watermark extracting process is the process to extract back the watermarked message that we embedded before.

3.2.1 Watermark Embedding

First of all, in order to watermark the audio, there should be an embedding process to embed the information into the original audio. Figure 3.1 shows the detail watermarking embedding process for this research which refer to the Baras, Moreau, and Dutoit (2009)'s watermark embedding process. The additional feature which is hash sum function is used to double up the security of the watermarked audio even thought there is a pseudo-random sequence already exists as a security feature.

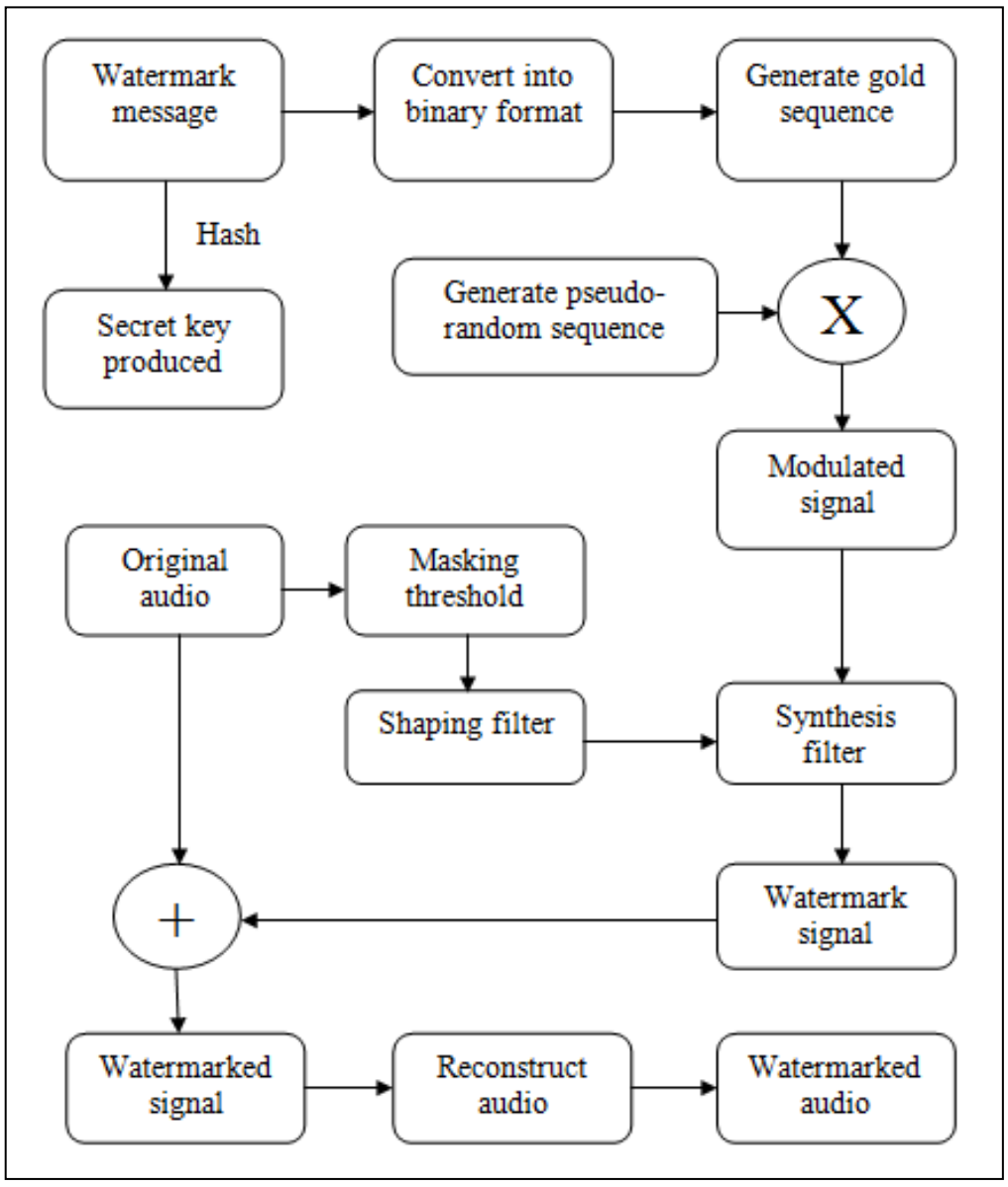

Figure 3.1: Watermark Embedding Scheme

At first, the watermark message or the data to be hidden will converted into binary vector, W_i by using the ASCII encoding information. After that, W_i will convert into gold sequence which have the rules that if $W_i = 0$, then $W_i = -1$, or if $W_i = +1$, then $W_i = 1$ by using this formula $2n - 1$. Then, pseudo-random sequence generator will generate a pseudo-random sequence or known as spread waveform, P_i which is a vector with the element of +1 and -1 only. For example, if the binary code is 0 then it will changed to -1 and while is 1 then it will changed to +1. After that, modulation of the signal will apply in order to modulate the signal thru multiplication between W_i and P_i.

On the other hand, the original audio, Q, this is a column vector that contains the information about magnitude of the audio signal. By employing psychoacoustic model, the original audio is converted into frequency domain by FFT and get all absolute value only, X: $X \in R^+$. For example if the value is -1.45 then it absolute value is 1.45. Then calculates its power magnitudes (dB scale), O_i by using this formula $O_i = 90.302 + 10 * \log_{10} X^2$ while 90.302 is the power normalize constant. Decibel is a measurement unit which used to measure the intensity of a sound. After that, analyze the power spectrum and find the tones maskers. If O_i is a local maxima and is greater than 7dB in a frequency dependent neighborhood, then it is a tone. Assume that the length for the power spectrum is 256, and define the neighborhood as, if $2 < i < 63$, it is within a 2 length neighborhood, or if $63 <= i < 127$, it is within 2 and 3 length neighborhood, or if $127 <= i < 256$, it is within 2,3,4,5, and 6 length neighborhood. For those that fall at the beginning or end of the power spectrum is not a local maxima.

After that, the next move is to find out the noise masker. At first, set all members of the neighborhood to 0 which is to eliminate the members of neighborhood. Those potential noise members that have indices of 1 are part of the noise masker. There is a process to check the masker power spectrum too. It is to check whether the value of masker power spectrum above the absolute threshold of hearing (ATH). If the masker value below ATH, then it will eliminated as it will not heard by anybody. Next is to check whether the masker value that above ATH within a critical bandwidth and the only strongest one will just remain, others will not bother anymore. At last, it is to find out the global threshold by relating the masker power to nearby frequency so that can determine whether the signal is audible over a masker. After the analyze process, a

masking threshold have produced, WE_i which is the watermark amplitude. The masking threshold is the intensity or amplitude of the sound that are audible to the listener. After getting the masking threshold, the power spectrum is converted back to the signal again by IFFT. Then filter the signal by getting the shaping filter response to get the coefficient, ai and gain, b0 of the signal via Levinson algorithm. Further details will be discussed on next chapter. Below is the Levinson equation for this research. *L(x)* is the perceptual shaping filter while *k* is the number of the coefficient.

$$L(x) = \frac{b0}{ai\ (1) + ai(2)x^{-1} + \ \ldots + ai(k)x^{-k+1}}$$

Whereas the equation for synthesis filter is show at below while *S* is the watermark signal

$$S = L(x) * (WE_i * P_i)$$

Last but not least, below is the encoding equation for this research while *Y* is the watermarked signal.

$$Y = Q + S$$

3.2.2 Watermark Extracting

In extracting process, the secret key and watermarked signal should be provided from the audio owner which is the person who hidden the message into the original audio before. At first, the watermarked audio will analyze via psychoacoustic model that same as the process mention in watermark embedding, section 3.2.1, to get the masking threshold of the watermarked audio. Next, inverse shaping filter will be performed to extract the equalized signal and hence estimate it covariance matrix.

On another hand, secret key will be used to find back the number of embed bits so that it may perform well during extracting process. Next, an arbitrary or random modulated signal is produced. After that, auto correlation the modulated signal and then pass it to the inverse synthesis filter together with the estimated covariance matrix of equalized signal. After undergo the inverse synthesis filter for the equalized signal,

watermark signal is extracted and then the watermark message is reshape and extracted. Further details will be explained on next chapter. Figure 3.2 shows the extracting scheme for this research which refer to the Baras, Moreau, and Dutoit (2009)'s watermark embedding process.

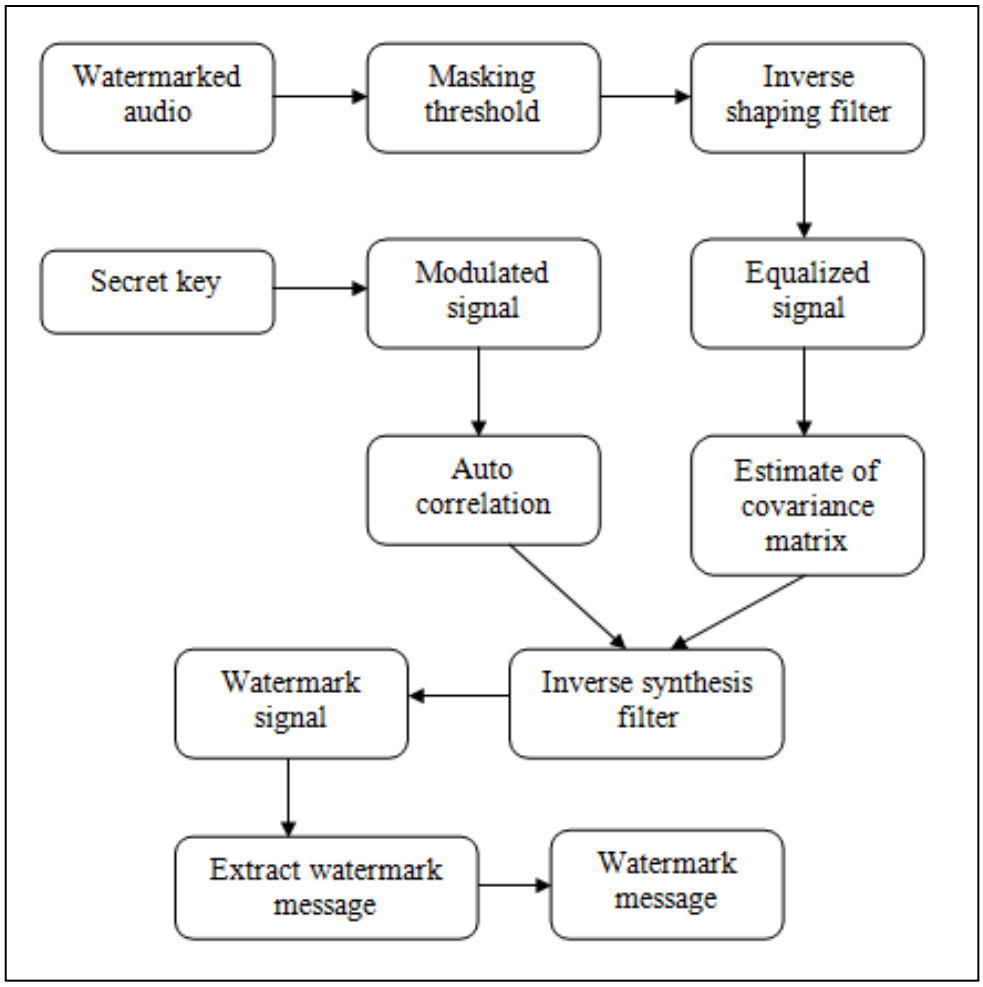

Figure 3.2: Watermark Extracting Scheme

3.2.3 Hashing function and Watermark Evaluation

In order to study the suitability of using hash function to verify the modification attacks, there should be a watermark process for the hash message. Firstly, the watermark message will be hash into 64 bytes of message by using SHA256 algorithm. After that, 21 bytes of message will embedded into the watermarked audio again at somewhere else of the watermarked audio. Next, during the extracting process for the hash watermark, the extracted hash watermark message will compare with the

embedded hash watermark message so that can check whether the hash function can be used to against modification attacks.

Besides the aspect of hash function, the performance of audio watermark also needed to be concerned throughout the research. In this research, the SNR will be used for evaluate the quality of the watermarked audio. According to Ravula (2010), SNR is a parameter that used for measure the noise with the consequence of corrupted signal. It is defined as the ratio of the signal power to the noise power. The equation for SNR is

$$SNR = \frac{Power_{signal}}{Power_{noise}} \quad or \quad SNR = 10\log\left(\frac{\sum_{a=1}^{M_t} Z^2(a)}{\sum_{a=1}^{M_t}(Z(a) - Z'(a))}\right)$$

Z is the un-watermarked audio signal and Z' is the watermarked audio signal. Both Z and Z' has M_t samples. Furthermore, BER also will be the measurement for the watermark result which is to determine how many erroneous bits found by comparing the embedded and the extracted watermark bits. The equation for BER is

$$BER = \frac{Total\ number\ of\ erroneous\ bits}{Total\ number\ of\ bits}$$

3.3 Hardware and Software

In this project, the hardware to be used is a personal laptop, an optical mouse and a printer. There is no special device or hardware needed to undergo this research during execution phase. On the other hand, the software to be used is Microsoft Word 2007, Adobe Reader X version 10.1.0, and MATLAB 7.10.0 (R2010a).

In fact, the reason for choosing MATLAB as the system development tool is because it has a lot of toolbox integrated with it. In the process of audio watermarking, we need to analyze the audio signal and there is a signal processing toolbox provided by the MATLAB. Signal processing toolbox provides industry-standard algorithms for analog and DSP. By using this toolbox, we can visualize the audio signals in time and frequency domains, computer FFT for spectral analysis and other signal processing

techniques. Other than that, it also provides the feature of estimate the power spectral density which we require it in analyze the audio frequency (The MathWorks Inc., 2012). Furthermore, in MATLAB it is more easy to plot the graph for analysis purpose and the data is easy to handle as it is in vector or matrix form. Consequently, I decide to use MATLAB as the system developing tool throughout the research project.

CHAPTER 4 DESIGN AND IMPLEMENTATION

4.1 Introduction

In this section, coding parts of the audio watermarking will explain clearly for better understanding. There will be total of four portions which is spread spectrum technique, psychoacoustic model, zero-forcing equalization and detection, wiener filtering, and hash sum function. Reference of coding is based on Baras, Moreau, and Dutoit (2009).

4.2 Embedding Process

In this session, there will be three steps for audio watermarking whose are spread spectrum technique, psychoacoustic model, and perceptual shaping filter. Below are the embedding processes of codes for audio watermarking. At first, we define what message we want to embed into the audio signal. After that, we convert the message into binary number which either 0 or 1 as a vector form. As describe in the previous chapter, there are 7 bits for one letter. Then, we will perform the direct sequence spread spectrum technique which generates gold sequence for the bit message by using this formula: 2n-1, where n is either 0 or 1 only. Next, generate random sequence for the bit message known as spread waveform and moderate the signal. Moreover, we now define which audio will be chosen for embedding the message. We should decide the sample size for watermarking and in our case the sample size is calculated by multiplying the number of bits with the number of samples.

After the signal preparation, we now perform the psychoacoustic model for the embedding process. At first, we decide the number of coefficient for the filter which

employ Levinson algorithm, and state the number of sample to be used for analysis in psychoacoustic function, 512 per frame will be used in our case. To ensure the filtering continuity from one frame to another, state vector will be used which store the final state of the filter at the end of one frame and applies it as initial conditions for the next frame. There is a looping process to embed entire watermark message into the audio signal. In every loop, psychoacoustic function/model analyzes the input signal which is audio signal with *n*th frame and it will return the power spectral density of masking threshold for the audio signal with *n*th frame. Then we will perform the shaping filter design to get the coefficient, ai and gain, b0 of the masking threshold so that we can generate the watermark signal using that information. Finally, it come to synthesis filtering stage which filter the modulate signal frame by frame and after finish looping process, there will have last synthesis filter in case there are incomplete frame of watermark signal. Figure 4.1 show the watermarked signal which combine the audio signal and watermark signal, we can see that the watermark signal is within the audio signal so it will not produce noise in the watermarked signal. While, figure 4.2 show a few samples of the audio, watermark, and watermarked signal. Based on the graph, there is no significant change between audio signal and watermarked signal.

```
embedMessage='Testing:Owner name is Chai Siew Li,Copyright@2013.';% Total
number of letter=50

binaryBits=dec2bin(embedMessage); %Convert the message into ASCII format

which is 7 bit string

binaryBits=binaryBits(:)'; % reshape to a single string which only 0 and 1

binaryBits=double(binaryBits)-48; % now convert into binary form 0 and 1

symbols=binaryBits*2-1; %Generate gold sequence, if 0 then -1 & if 1 then 1

NumberOfBits=length(binaryBits); %total number of bit=7*50=350,array[1,350] will

be produced

Fs=44100; %sampling frequency
R=100; %binary rate
NumberOfSamples=round(Fs/R); %sample per bit
```

```
spreadWaveform=2*round(rand(NumberOfsamples,1))-1; %generate random sequence known as spread waveform
modulatedSignal=zeros(NumberOfBits*NumberOfSamples,1); %modulate the signal by bit message and spread waveform
for m=0:NumberOfBits-1
modulatedSignal(m*NumberOfSamples+1:m*NumberOfSamples+NumberOfSamples)=symbols(m+1)*spreadWaveform; %if original bits is 1 -1 1 -1 -1 1 1, spread waveform is -1 1 1 -1 1 1 -1 then the modulated signal is -1 -1 1 1 -1 1 -1. To retrieve back the original bits, use the modulated signal to multiply with the spread waveform
end
audioSignal=wavread('moderate.wav'); %original audio signal
audioSignal=audioSignal(1:length(audioSignal),1); %obtain mono tone
audioSignal=audioSignal(1500000+1:1500000+NumberOfBits*NumberOfSamples); %create the sample size for the audio signal
sound(audioSignal,Fs); %hear the original audio sound
%Informed watermarking made inaudible
NumberOfCoefficients=50; %Number of coefficients for perceptual shaping filtering
NumberOfSamples_PAM=512; %number of samples for psychoacoustic model
PAM_frame=(10*NumberOfSamples_PAM+1:10*NumberOfSamples_PAM+NumberOfSamples_PAM);
figure; %11th frame will be retrieved from the audio signal sample size
plot(PAM_frame/Fs,audioSignal(PAM_frame)); %plotting the graph for audio signal with respect to the PAM frame size
 %perform perceptual shaping for whole watermark signal to process audio signal block by block
state=zeros(NumberOfCoefficients,1);
for m=0:fix(NumberOfBits*NumberOfSamples/NumberOfSamples_PAM)-1

PAM_frame=m*NumberOfSamples_PAM+1:m*NumberOfSamples_PAM+NumberOfSamples_PAM;
   maskingThreshold=psychoacousticModel(audioSignal(PAM_frame)); %Get the masking threshold for the audio signal by using the psychoacoustic function
   shapingFilterResponse=maskingThreshold;
```

```
[b0,ai]=shapingFiltering(shapingFilterResponse,NumberOfCoefficients); %computes the coefficients of an auto-regressive filter
[watermarkSignal(PAM_frame,1),state]=filter(b0,ai,modulatedSignal(PAM_frame),state); %Synthesis filter
end
PAM_frame=(m*NumberOfSamples_PAM+NumberOfSamples_PAM+1:NumberOfBits*NumberOfSamples);
watermarkSignal(PAM_frame,1)=filter(b0,ai,modulatedSignal(PAM_frame),state); % last synthesis filter in case there are incomplete frame of watermark signal
figure;
plot((0:Fs-1)/Fs,audioSignal(1:Fs));
hold on;
plot((0:Fs-1)/Fs,watermarkSignal(1:Fs),'r-');
watermarkedSignal=watermarkSignal+audioSignal;
range=(34739:34938);
subplot(3,1,1);plot(range/Fs,watermarkSignal(range),'k');
subplot(3,1,2);plot(range/Fs,audioSignal(range));
subplot(3,1,3);plot(range/Fs,watermarkedSignal(range),'r');
sound(watermarkedSignal,44100);
```

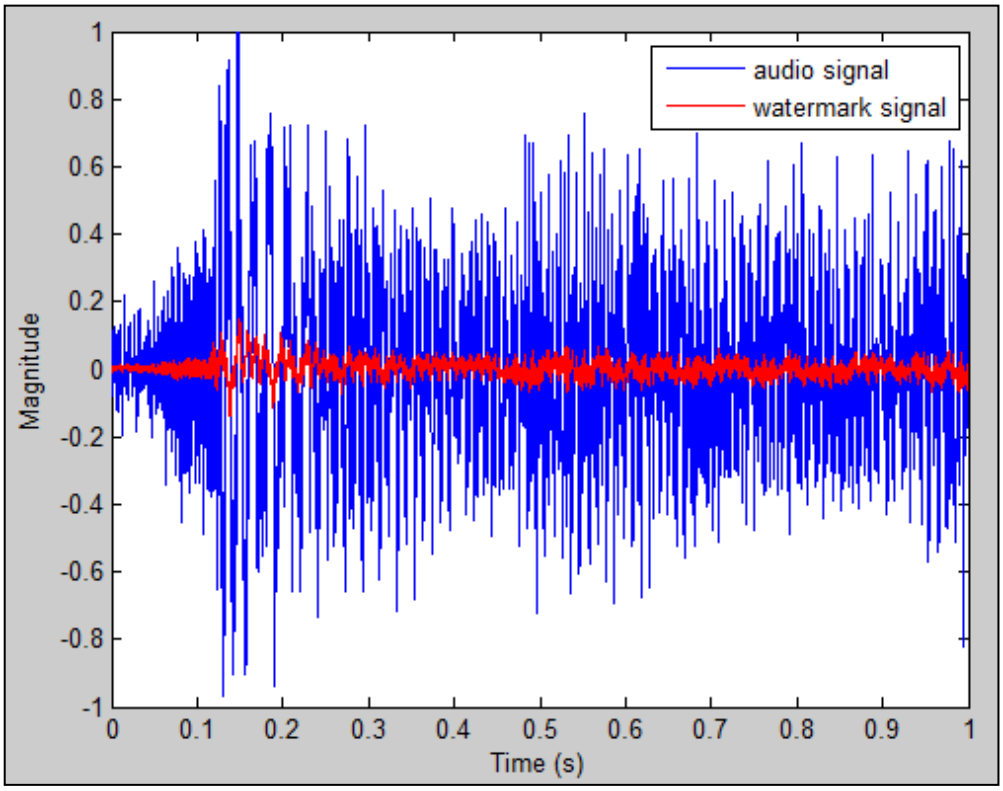

Figure 4.1: Audio signal and watermark signal

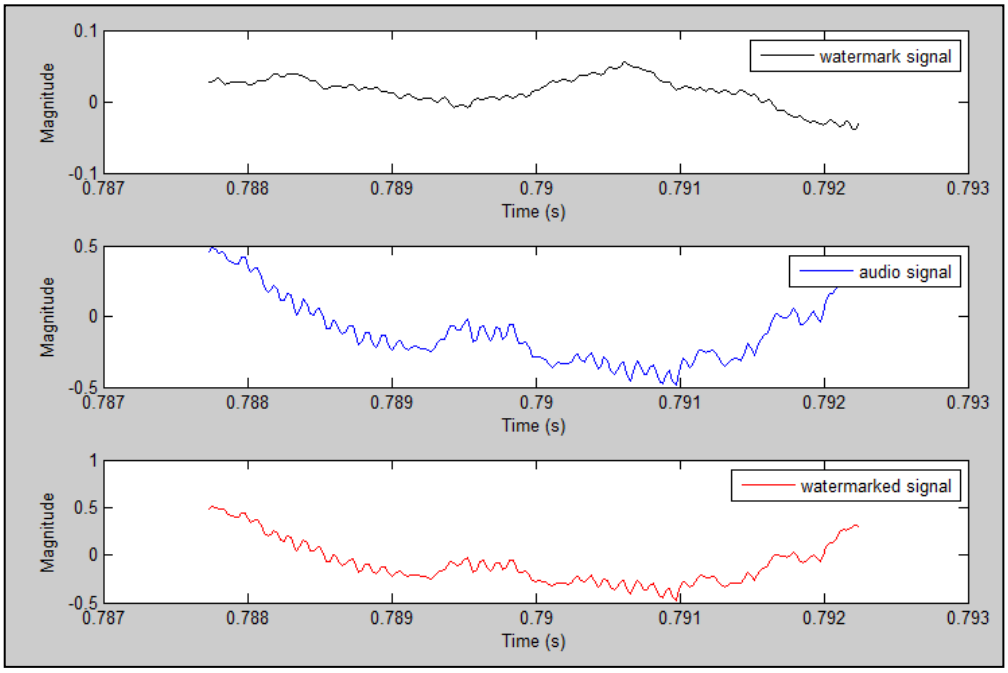

Figure 4.2: A few samples of the audio, watermark, and watermarked signals

4.3 Extracting Process

In this session, there will be two approaches used for extracting the watermark signal from the watermarked signal which are zero-forcing equalization and detection, and wiener filtering. Below are the extracting processes of codes for audio watermarking. The procedure of extract the watermark signal is more or less similar to the embedding process. The difference is only in the parts of psychoacoustic model. Due to there is no audio signal at the receiver, so that we replace the audio signal to watermarked signal. We get the power spectral density of the masking threshold from the watermarked signal and then it will later be used for perceptual shaping filter. At the filter stage, watermarked signal is filtered by the filter frequency response which is gain, b0 and coefficient, ai. Next, shaping filter response is being inversed so that it will yield the filtered received signal which denoted as equalized signal in the codes. Figure 4.3 show the equalized audio signal and the modulated signal. While figure 4.4 is the corresponding power spectral density of the equalized audio signal and the modulated signal which denoted as signal and noise in the graph respectively. After that, we extract the watermark by applying correlation demodulator denoted as spread waveform to the equalized signal. If the value is less than 0, then it will become 0. Meanwhile, if the value is more than 0, then it will become 1. Hence, the watermark bits are retrieved and then reshape it back to the characters or letters.

```
%Receiver based on zero-forcing equalization and detection
equalizedSignal=zeros(NumberOfBits*NumberOfSamples,1);
state=zeros(NumberOfCoefficients,1);
for m=0:fix(NumberOfBits*NumberOfSamples/NumberOfSamples_PAM)-1

PAM_frame=m*NumberOfSamples_PAM+1:m*NumberOfSamples_PAM+Number
OfSamples_PAM;
   [maskingThreshold g]= psychoacousticModel (watermarkedSignal(PAM_frame));
   shapingFilterResponse=g;
   [b0,ai]=shapingFiltering(shapingFilterResponse,NumberOfCoefficients);

[equalizedSignal(PAM_frame),state]=filter(ai./b0,1,watermarkedSignal(PAM_frame)
```

```
,state);
   if m==10
      pwelch(equalizedSignal(PAM_frame),[],[],[],2);
      hold on;
      [H,W]=freqz(ai./b0,1,256);
      plot(W/pi,20*log10(abs(H)),'r','linewidth',2);
   end;
end
equalizedAudioSignal=zeros(NumberOfBits*NumberOfSamples,1);
state=zeros(NumberOfCoefficients,1);
for m=0:fix(NumberOfBits*NumberOfSamples/NumberOfSamples_PAM)-1

PAM_frame=m*NumberOfSamples_PAM+1:m*NumberOfSamples_PAM+Number
OfSamples_PAM;
   [maskingThreshold g]=psychoacousticModel (watermarkedSignal(PAM_frame));
   shapingFilterResponse=g;
   [b0,ai]=shapingFiltering(shapingFilterResponse,NumberOfCoefficients);

[equalizedAudioSignal(PAM_frame),state]=filter(ai./b0,1,audioSignal(PAM_frame),s
tate);
end
range=(34739:34938);
subplot(2,1,1);plot(range/Fs,equalizedAudioSignal(range));
subplot(2,1,2);plot(range/Fs,modulatedSignal(range));

range=(34398:34938);
pwelch(equalizedAudioSignal(range),[],[],[],2);
[h,w]=pwelch(modulatedSignal(range),[],[],[],2);
hold on;
plot(w,10*log10(h),'--r');

for m=0:NumberOfBits-1
alpha(m+1)=(equalizedSignal(m*NumberOfSamples+1:m*NumberOfSamples+Num
```

```
berOfSamples)'*spreadWaveform)/NumberOfSamples;
   if alpha(m+1)<=0
      receivedExtractedBits(m+1)=0;
   else
      receivedExtractedBits(m+1)=1;
   end
end
range=(81:130);
subplot(3,1,1);
stairs(range,binaryBits(range));
subplot(3,1,2);
stairs(range,alpha(range));
subplot(3,1,3);
stairs(range,receivedExtractedBits(range));
numberOfErroneousBits=sum(binaryBits~=receivedExtractedBits);
totalNumberOfBits=NumberOfBits;
BitErrorRate=numberOfErroneousBits/totalNumberOfBits;
extractedChars=reshape(receivedExtractedBits(1:NumberOfBits),NumberOfBits/7,7);
extractedMessages=char(bin2dec(num2str(extractedChars)))'
```

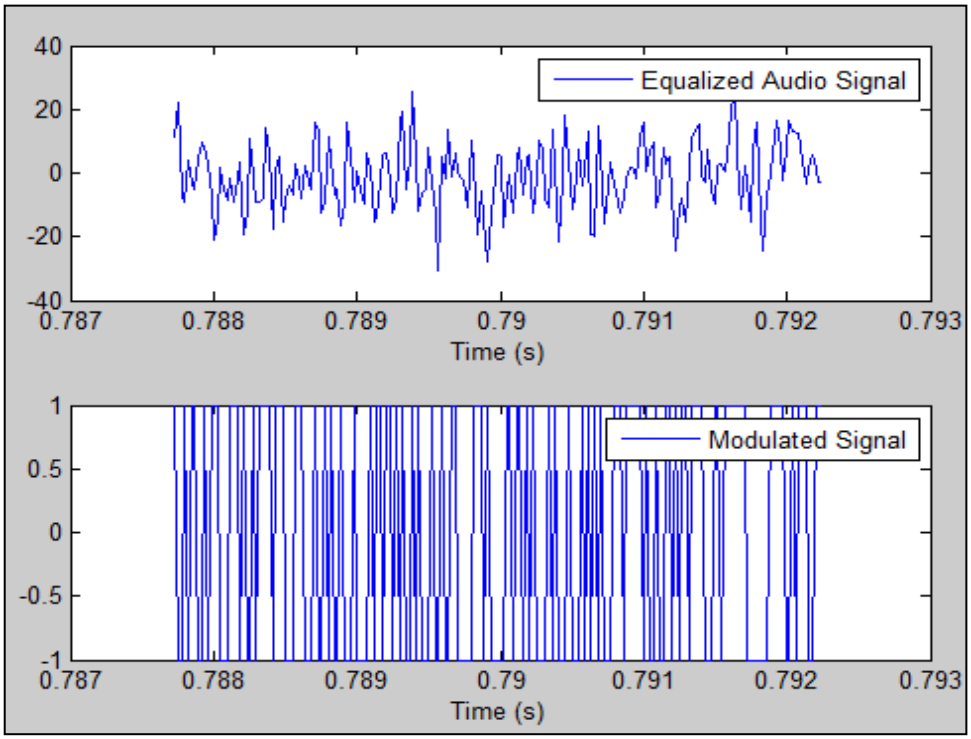

Figure 4.3: Equalized audio signal and the modulated signal

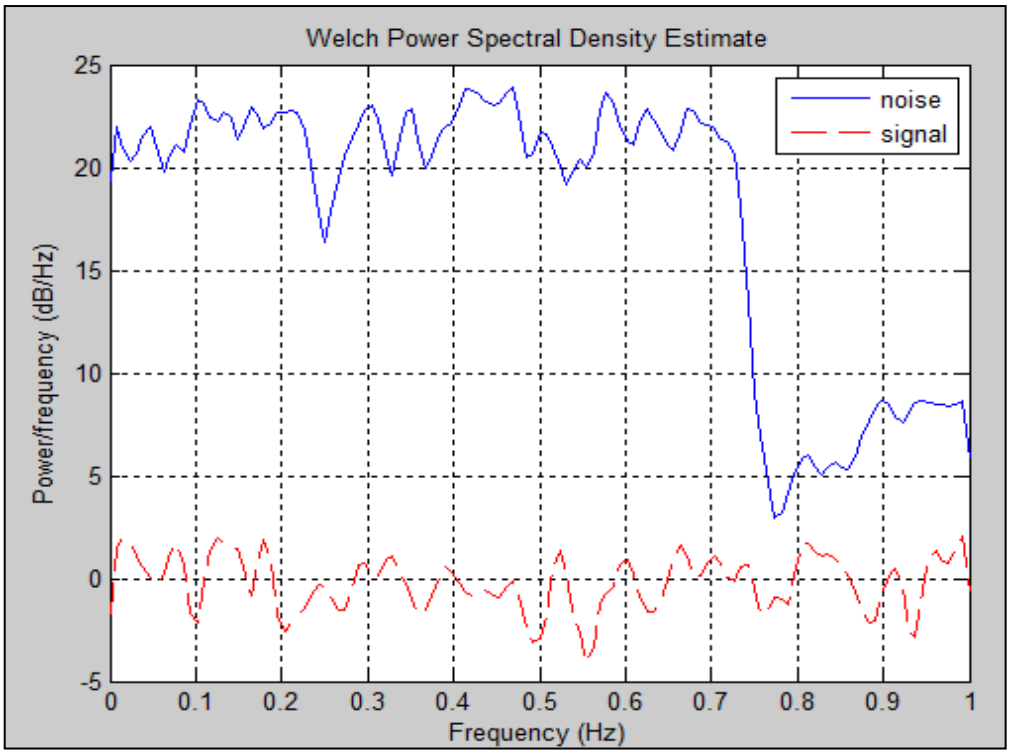

Figure 4.4: Power spectral density of equalized audio signal and modulated signal

Since the result of the extracting watermark by using this approaches is not ideally. Hence, we perform an enhancement for the watermark extracting process which is applying wiener filtering. The purpose of wiener filtering stage is to enhance the signal-to-noise ratio between the modulated signal and the equalized audio signal. Here is the Wiener-Hopf equation, hi=modulated signal auto-correlation vector/equalized signal covariance matrix. While, hi is the coefficients of the filter that will produce the output signal maximally similar to the modulated signal in the root-mean-square error (RMSE) sense. After estimate the covariance matrix of modulated signal and modulus equalized signal, we now perform the toeplitz function which will return a symmetry Toeplitz matrix formed from the covariance matrix of equalized signal since it is real number. Next, we estimate the impulse response of the wiener filter by using the wiener-hopf equations and then perform the synthesis filtering. Nevertheless, in order to get the measure of the magnitude in the wiener output signal, norm function will be used. It will return the largest singular value from the result. If the magnitude of the signal is approximate to 0, then it will obtain the root-mean-square value of the wiener output signal. Moreover, since there is number of coefficients for the non-causal part, so the result is delayed with the number of coefficients samples, in our case is 50. Hence we re-construct the output signal by adding back number of coefficients samples. At the end of extracting process, it is same like previous method. The difference is now the watermarked signal is replaced by wiener output signal.

```
%Wiener Filtering
modulated_signal_autocor_vect=xcorr(modulatedSignal,NumberOfCoefficients,'biased');
wiener_output_signal=zeros(NumberOfBits*NumberOfSamples,1);
state=zeros(2*NumberOfCoefficients,1);
for m=0:fix(NumberOfBits*NumberOfSamples/NumberOfSamples_PAM)-1

PAM_frame=(m*NumberOfSamples_PAM+1:m*NumberOfSamples_PAM+NumberOfSamples_PAM);
equalized_signal_autocor_vect=xcorr(equalizedSignal(PAM_frame),2*NumberOfCoefficients,'biased');
```

```
equalized_signal_cov_mat=toeplitz(equalized_signal_autocor_vect(2*NumberOfCoe
fficients+1:end));
   hi=equalized_signal_cov_mat\modulated_signal_autocor_vect;

[wiener_output_signal(PAM_frame,1),state]=filter(hi,1,equalizedSignal(PAM_frame)
,state);
end
wiener_output_signal=[wiener_output_signal(NumberOfCoefficients+1:end);zeros(N
umberOfCoefficients,1)];
range=(34398:34938);
wiener_output_audio=filter(hi,1,equalizedAudioSignal(range));
wiener_output_modulated=filter(hi,1,modulatedSignal(range));
range=(34739:34938);
subplot(2,1,1);plot(range/Fs,wiener_output_audio(341:540));
subplot(2,1,2);plot(range/Fs,wiener_output_modulated(341:540));
pwelch(wiener_output_audio,[],[],[],2);
[h,w]=pwelch(wiener_output_modulated,[],[],[],2);
hold on;
plot(w,10*log10(h),'--r','linewidth',2);
for m=0:NumberOfBits-1
alpha(m+1)=(wiener_output_signal(m*NumberOfSamples+1:m*NumberOfSamples+
NumberOfSamples)'*spreadWaveform)/NumberOfSamples;
   if alpha(m+1)<=0
      receivedExtractedBits(m+1)=0;
   else
      receivedExtractedBits(m+1)=1;
   end
end
numberOfErroneousBits=sum(binaryBits~=receivedExtractedBits);
totalNumberOfBits=NumberOfBits;
BitErrorRate=numberOfErroneousBits/totalNumberOfBits;
extractedChars=reshape(receivedExtractedBits(1:NumberOfBits),NumberOfBits/7,7);
extractedMessage=char(bin2dec(num2str(extractedChars)))'
```

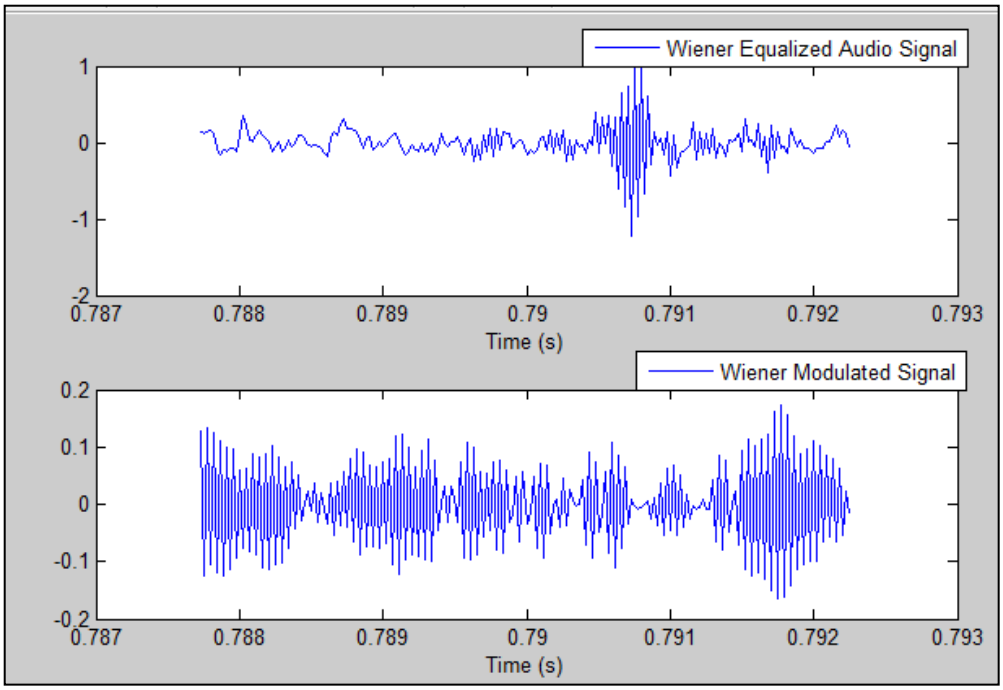

Figure 4.5: The wiener equalized audio signal and the wiener modulated signal

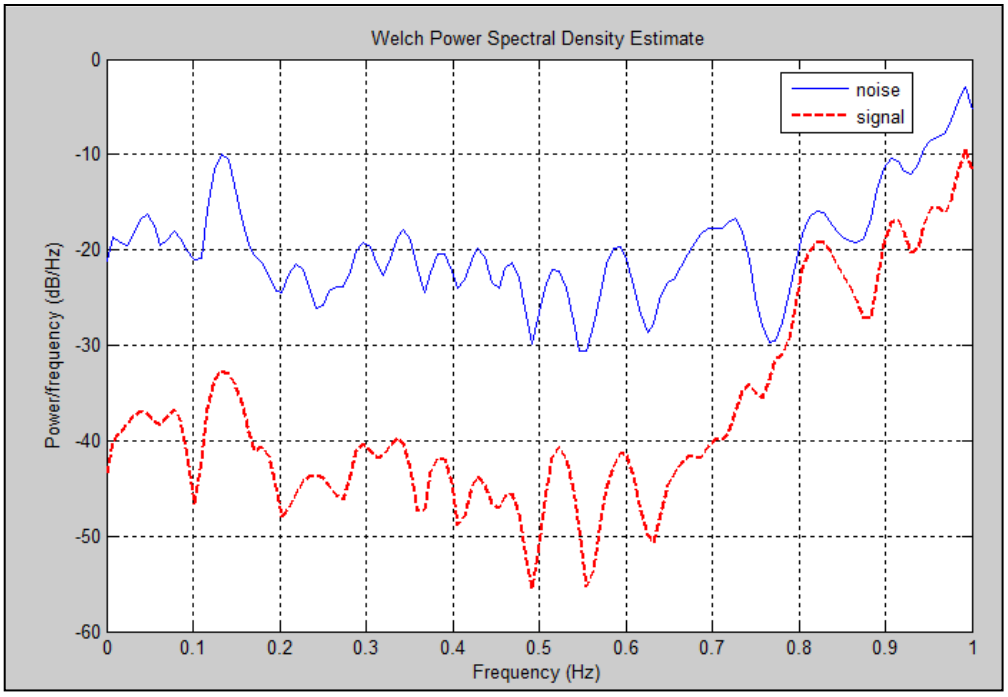

Figure 4.6: The power spectral density of the wiener equalized audio signal and wiener modulated signal

4.4 Hash Sum Function

After the completion of audio watermarking, hash function will be used for verify the modification attacked by third parties. At first, the half watermark message will be encrypted by using SHA256 function. Then the encrypted message will convert into another watermark signal and embed into the audio secretly that the owner do not know about it existence. Below are the modification processes of codes for audio watermarking. The watermark signal is watermarked at the start frequency which is 1500000 while the hash watermark signal is watermarked at the start frequency which is 3500000. At first, we hash 21 bytes of the watermark message and then embed it into the audio signal. So now, the whole watermarked signal is included the owner message and the hash message. After that, we extract the owner message and then perform hashing again. Next, compare it with the extracted hash message. If both the message is similar, then the watermarked signal did not modified by third parties. Once both the message is not similar, then the watermarked signal did modified by third parties.

```
embedMessage='Testing:Owner name is Chai Siew Li,Copyright@2013.';
audioSignal=wavread('moderate.wav');
audioSignal=audioSignal(1:length(audioSignal),1);
watermarkedSignal=watermark_embed(audioSignal,1500000,embedMessage);
extractedMessage=watermark_extract(watermarkedSignal);
substringMessage=embedMessage(1:length(embedMessage)/4);
hash_value=hash(substringMessage);%case-sensitive
watermarkedHashSignal=watermark_embed(audioSignal,3500000,hash_value);
extractedHashMessages=watermark_extract(watermarkedHashSignal);
substringExtractMessage=extractedMessage(1:length(extractedMessage)/4);
hash_value=hash(substringExtractMessage);%case-sensitive
if(hash_value== extractedHashMessages)
    disp('Hash value is the same,it do not modified by third parties');
else
    disp('Hash value is not same,it do modified by third parties');
end;
```

4.5 Implementation

By converting the abstract concept for the proposed embedding and extracting scheme into concrete concept, a simple graphical user interface which developed via MATLAB 7.10.0 (R2010a) have been introduced as shown in figure 4.7.

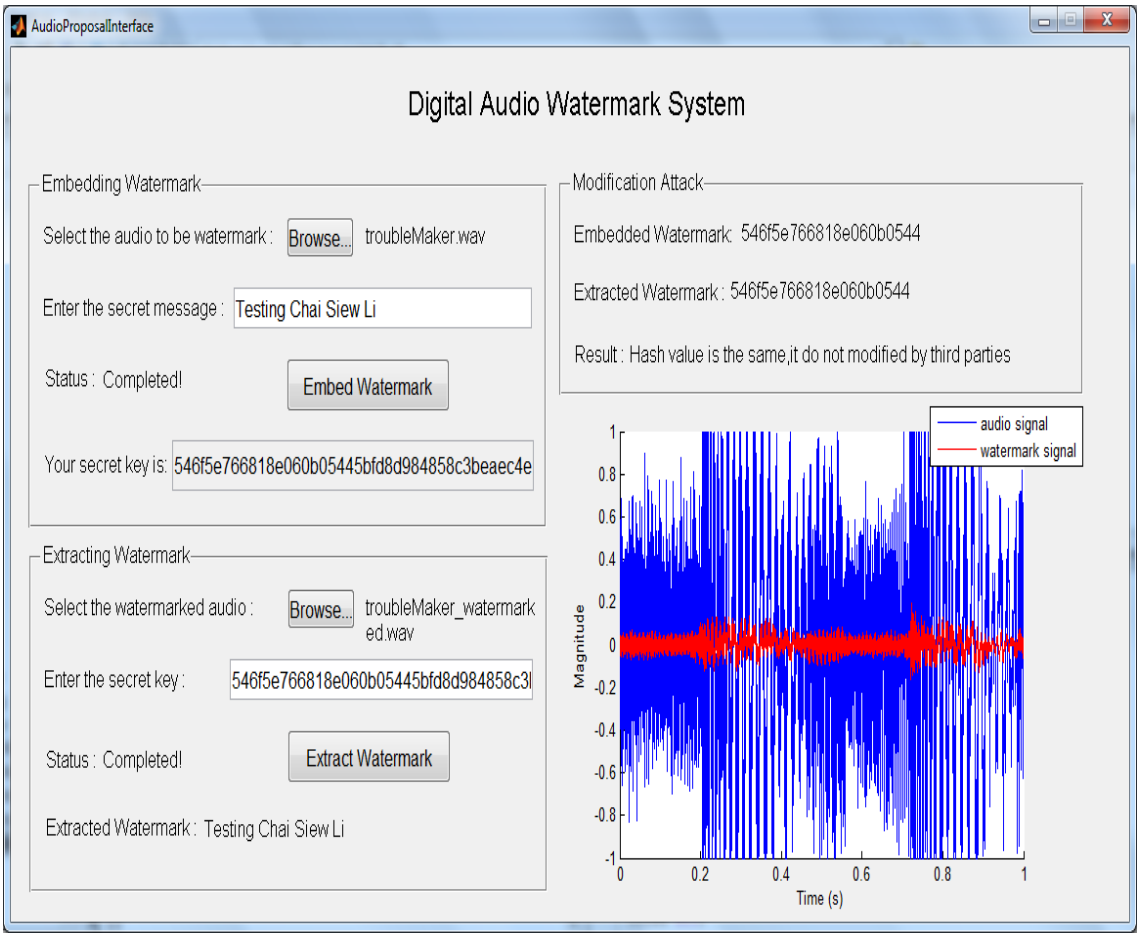

Figure 4.7: The Graphical User Interface (GUI) for proposed watermark scheme

CHAPTER 5 RESULT AND DISCUSSION

5.1 Introduction

In this section, we will analyze and discuss the result of audio watermarking and the research constraints. As discuss in previous chapter, 3 samples of audio signal have been chosen those categorized as quiet state, moderate state, and noise state signal. Figure 5.1, 5.2 and 5.3 shows the magnitude of the quiet state, moderate state, and noise state signal. On the other hand, each sample of audio signal will be tested with different length of the watermark message so that we can analyze the watermark results. There are two type of measurement in order to evaluate the result which is signal-to-noise ratio and the bit error rate.

5.2 Result Analysis

Results for both watermark process and hash watermark process by using waveform A is shown in table below. The results of watermark embedding and extracting by using waveform A is shown in Table 5.1. The results of hash watermark embedding and extracting by using waveform A is shown in Table 5.2. The list of legend representation is shown in Table 5.3. Based on the results in table 5.1, we can observe that the BER for the quiet.wav, moderate.wav and noisy.wav is within the range of 0.0019 - 0.0071, 0.0000 – 0.0032 and 0.0000 – 0.0032. The average ranges are 0.00328, 0.00116 and 0.00128 respectively. The moderate.wav achieved better results if compared with quiet.wav and noisy.wav. Besides, the accuracy of the extracted watermark message depends on the length of the message as we can see that when the total number of bits is increased, the number of erroneous bits also increased. We can say that, the number of erroneous bits is directly proportional to the total number of

embedded bits. Furthermore, the SNR for the quiet.wav, moderate.wav and noisy.wav are in the range of 23.9848 - 24.4196, 22.8721 - 23.0840, and 22.4289 - 22.5278 accordingly. We can see that, the range is slightly decreased when the audio signal change from quiet state to moderate state and then to noisy state. As in the quiet state signal, the watermark signal apparently looks like a noise signal while in the noisy state signal the watermark signal apparently not treated as a noise signal. According to Ketcham and Vongpradhip (2007), they claimed that the International Federation of the Phonographic Industry (IFPI) which is the organization that represents the worldwide recording industry, the SNR of the watermarked audio signal should be greater than 20 dB. Consequently, the results that uses spread waveform A satisfies the requirement that stated by IFPI.

Due to the reason of not getting the good results, there is another reason need to be concerned is the spread spectrum waveform. Hence, there is another result for watermark process and hash watermark process by using different spread spectrum waveform which is waveform B. Table 5.4 show the results of watermark embedding and extracting by using waveform B, Table 5.5 show the results of hash watermark embedding and extracting by using waveform B and Table 5.6 show the type of waveform pattern. Based on the result in table 5.4, we can observe that the BER for the quiet.wav, moderate.wav and noisy.wav is in the range of 0.0000 - 0.0043, 0.0000 – 0.0032 and 0.0000 – 0.0036, and it average range of 0.00138, 0.00102 and 0.00132 respectively. For quiet.wav and moderate.wav, there is an improvement when using waveform B and only noisy.wav do not have significant changes on the results. Based on that it can be said that different waveform will produce different results and affect the effectiveness of the watermark process. Anyway, its SNR is quite low if compared to the results of using waveform A which is in the range of 20.7874 – 21.0217, 19.7907 – 19.8917, and 19.8279 – 20.0291 respectively.

On the other hand, based on the table 5.2, we can see that BER for the quiet.wav, moderate.wav and noisy.wav are all 0 respectively. Due to the embedded position for the watermark message and the hash message are different, so the results for both are different. Besides, the SNR for the quiet.wav, moderate.wav and noisy.wav are in the range of 23.4333 – 23.5025, 22.9085 – 22.9728, and 21.9175 – 22.1399 respectively. In

the three music samples, the accuracy achieved is 100 %. Therefore, it shows the contrast between the result of watermark embedding and extracting. In table 5.5, the BER for the quiet.wav, moderate.wav and noisy.wav are in the range of 0.0000 – 0.0068, 0.0000 – 0.0000 and 0.0000 – 0.0000 respectively. Meanwhile, it SNR for the quiet.wav, moderate.wav and noisy.wav are in the range of 20.4932 – 20.5403, 19.6902 – 19.7020, and 20.0883 – 20.4373 respectively. Based on the results in table 5.5, it had shown that only moderate.wav and noisy.wav achieved 100% of accuracy. It do not same as the results in table 5.2 which all the three music samples achieved 100% of accuracy.

However, based on the analyzed results it is found that there are some factors that causes the accuracy of extracted data which is the length of message or the total number of bits embed and the spread waveform. When the length of message is longer, the possibility to extract the watermarked message is lower. The possible reason is due to some inaccurate data produced during filtering stage because of the usage of correlation method to retrieve the watermark bit where if the value is less than 0, the value will become 0. Meanwhile, if the value is more than 0, then the value will become 1.It is possible that the actual bit value is 1 but it has been treated as 0 due to some dependent factor such as masking threshold. As the masking threshold for the audio signal and the watermarked signal may not be same, it will also affect the results of the extracted watermark.

Overall, it can be said that the proposed watermark scheme shows consistencies for the results as the bit rate accuracy for the extracted watermark achieved at least 99.99 % or even reached the maximum 100% for certain time and conditions. To get the better results, it is encouraged to embed the watermark message into the moderate state signal and not more than 532 of bits by using waveform A. It is not encouraged to embed the watermark message into quiet state signal because it shows the highest BER among another two music samples.

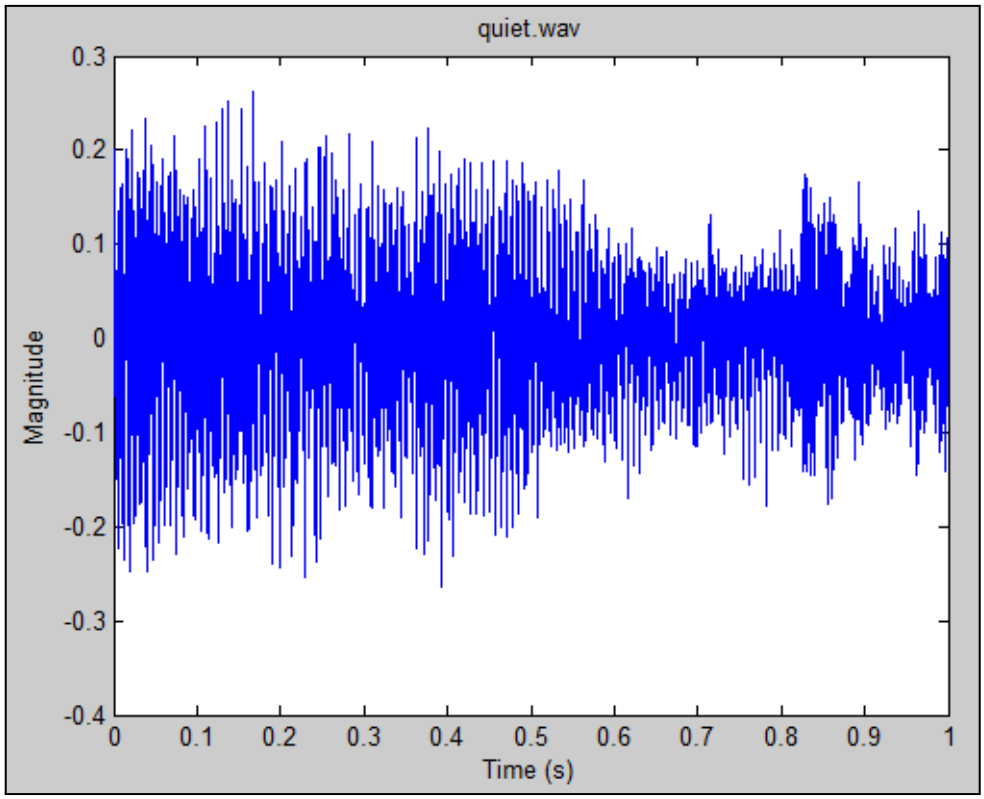

Figure 5.1: Magnitude of quiet state signal over time

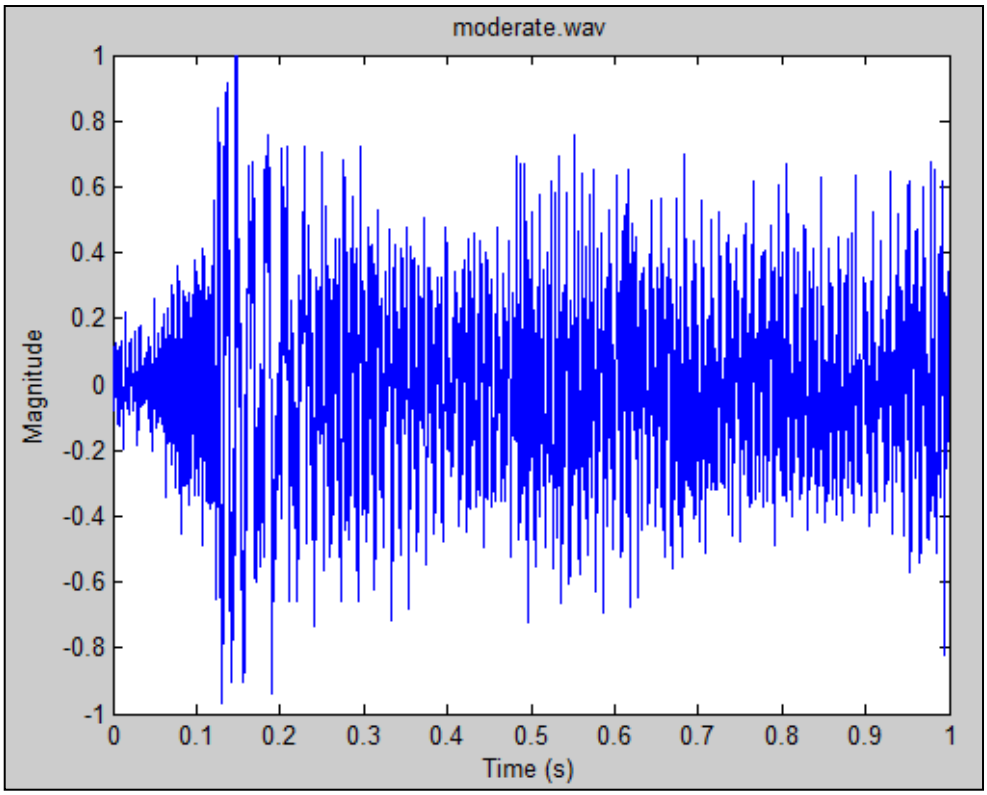

Figure 5.2: Magnitude of moderate signal over time

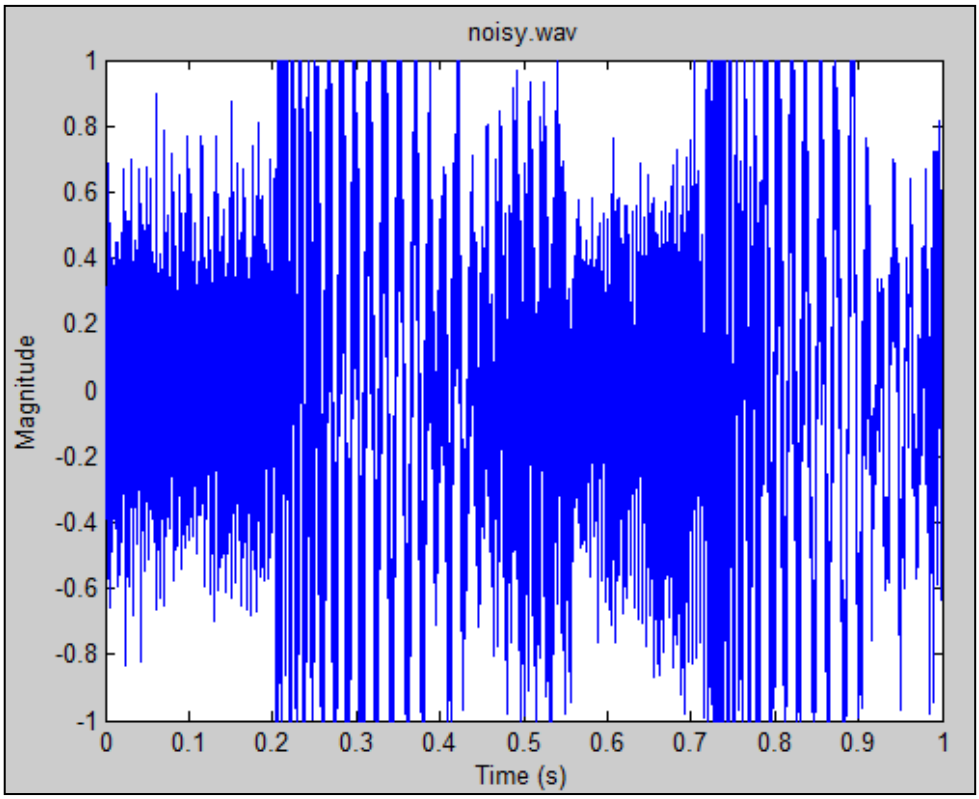

Figure 5.3: Magnitude of noisy state signal over time

Table 5.1: Result of Watermark Embedding and Extracting by using waveform A

Music Sample	Embedded Message	Extracted Message	Total Number of Bits	Number of Erroneous Bits	Bit Error Rate (BER)	Signal-to-Noise Ratio (SNR)
quiet .wav	M1	Te{ting Chai Siew Li	140	1	0.0071	23.9848
	M2	Testing Chai Siew Li, audio Watermarking	280	1	0.0036	24.1218
	M3	Testing Chai Siew Li, Audio Watermarking and will retrieved back$the message	532	1	0.0019	24.4196
	M4	This message will embedded and rmtrieved into and from the audio skgnal. Now want to test the length of messaGe	784	3	0.0038	24.0794

	M5	Testing Chai Siew Li, this message will embedded and retrieved`into and from the audio sIgnal. Now wqnt to test the length of messagd	938	4	0.0043	24.1089
Moderate.wav	M1	Testing Chai Siew Li	140	0	0.0000	23.0260
	M2	Testing Chai Siew Li, Audio Watermarking	280	0	0.0000	22.8721
	M3	Testing Chai Siew Li, Audio Watermarking and will retrieved back the message	532	0	0.0000	22.9653
	M4	This message wi|l embedded and retrieved into and from the audio signal. Now want to test the length of mersage	784	2	0.0026	23.0499
	M5	Testing Chai Siew Li, this message will embedded and retrieved into and from the audio signal. Now want to teut the length of messagd	938	3	0.0032	23.0840
noisy.wav	M1	Testing Chai Siew Li	140	0	0	22.4289
	M2	Testing Chai Siew Li, Audio Watermarking	280	0	0	22.5238
	M3	Testing Chai Siew Li, Audio Waterm`rking and will retrieved back the message	532	1	0.0019	22.5278
	M4	This message will embedded and retrieved ijto and from the audio signal. Now want to test the length of message	784	1	0.0013	22.5031
	M5	Testing Chai Siew Li, this messcge will embedded and retrieved into and from the audio sagnal. Now want to test the length of messagd	938	3	0.0032	22.4976

Table 5.2: Result of Hash Watermark Embedding and Extracting by using waveform A

Music Sample	Embedded Message	Extracted Message	Total Number of Bits	Number of Erroneous Bits	Bit Error Rate (BER)	Signal-to-Noise Ratio (SNR)
quiet.wav	H1	546f4e766818e060b0544	147	0	0	23.5025
	H2	fb0f24a8ad41fa46d7215	147	0	0	23.4656
	H3	d1025c07d8fec81a16fea	147	0	0	23.4487
	H4	0b697f53892f3e1c88ace	147	0	0	23.4452
	H5	a247fa889d1c959e80e77	147	0	0	23.4333
moderate.wav	H1	546f5e766818e060b0544	147	0	0	22.9085
	H2	fb0f24a8ad41fa46d7215	147	0	0	22.9728
	H3	d1025c07d8fec81a16fea	147	0	0	22.9492
	H4	0b697f53892f3e1c88ace	147	0	0	22.9227
	H5	a247fa889d1c959e80e77	147	0	0	22.9341
noisy.wav	H1	546f5e766818e060b0544	147	0	0	22.0970
	H2	fb0f24a8ad41fa46d7215	147	0	0	21.9175
	H3	d1025c07d8fec81a16fea	147	0	0	22.0417
	H4	0b697f53892f3e1c88ace	147	0	0	22.1399
	H5	a247fa889d1c959e80e77	147	0	0	22.0648

Table 5.3: List of Legend Representation

M1	Testing Chai Siew Li
M2	Testing Chai Siew Li, Audio Watermarking
M3	Testing Chai Siew Li, Audio Watermarking and will retrieved back the message
M4	This message will embedded and retrieved into and from the audio signal. Now want to test the length of message
M5	Testing Chai Siew Li, this message will embedded and retrieved into and from the audio signal. Now want to test the length of message

H1	546f5e766818e060b0544
H2	fb0f24a8ad41fa46d7215
H3	d1025c07d8fec81a16fea
H4	0b697f53892f3e1c88ace
H5	a247fa889d1c959e80e77

Table 5.4: Result of Watermark Embedding and Extracting by using waveform B

Music Sample	Embedded Message	Extracted Message	Total Number of Bits	Number of Erroneous Bits	Bit Error Rate (BER)	Signal-to-Noise Ratio (SNR)
quiet.wav	M1	Testing Chai Siew Li	140	0	0.0000	20.7874
	M2	Testing Chai Siew Li, Audio Watermarking	280	0	0.0000	20.8688
	M3	Testing Chai Siew Li, Audio Watermarking and will retrieved back the message	532	0	0.0000	21.0217
	M4	This message will embedded and retrieved into and from the audIo signal. Now want to test the length of messaGe	784	2	0.0026	20.8274
	M5	Testing Chai Siew Li, this messege will Embedded and retrieved into and from the audio sIgnal. Now want to test the length of messagd	938	4	0.0043	20.9351

Moderate.wav	M1	Testing Chai Siew Li	140	1	0.0000	19.7907
	M2	Testing Chai Siew Li, Audio Watermarking	280	0	0.0000	19.8316
	M3	Testing Chaa Siew Li, Audio Watermarking and will retrieved back the message	532	1	0.0019	19.8833
	M4	This message will embedded and retrieved into and from the audio signal. Now want to test the length of message	784	0	0.0000	19.8541
	M5	Testhng Chai Siew Li, this message will embedded and retrieved into and frmm the audio signal. Now want to test the length of messagd	938	3	0.0032	19.8917
noisy.wav	M1	Testing Chai Siew Li	140	0	0.0000	19.8279
	M2	Testing Chai Siew Ni, Audio Watermarking	280	1	0.0036	19.8544
	M3	Testing Chai Siew Li, Audio Watermarking and will retrieved back txe message	532	1	0.0019	20.0077
	M4	This message will embedded and retrieved into and from the audio signal. Now want to test the	784	0	0.0000	19.9782

		length of message				
	M5	Testing Chai Siew Li, this message will embedded and retrieved into and from the audio signal. Now want to test the length of messagd	938	1	0.0011	20.0291

Table 5.5: Result of Hash Watermark Embedding and Extracting by using waveform B

Music Sample	Embedded Message	Extracted Message	Total Number of Bits	Number of Erroneous Bits	Bit Error Rate (BER)	Signal-to-Noise Ratio (SNR)
quiet.wav	H1	546f5e766818e060b0544	147	0	0.0000	20.5265
	H2	fb0f24a8ed41fa46d7215	147	1	0.0068	20.5403
	H3	d1025c07d8fec81a16fea	147	0	0.0000	20.5091
	H4	0b697f53<92f3e1c88ace	147	1	0.0068	20.5117
	H5	a247fa88=d1c959e80e77	147	1	0.0068	20.4932
moderate.wav	H1	546f5e766818e060b0544	147	0	0.0000	19.6904
	H2	fb0f24a8ad41fa46d7215	147	0	0.0000	19.6999
	H3	d1025c07d8fec81a16fea	147	0	0.0000	19.6947
	H4	0b697f53892f3e1c88ace	147	0	0.0000	19.7020
	H5	a247fa889d1c959e80e77	147	0	0.0000	19.6902
noisy.wav	H1	546f5e766818e060b0544	147	0	0.0000	20.3222
	H2	fb0f24a8ad41fa46d7215	147	0	0.0000	20.0883
	H3	d1025c07d8fec81a16fea	147	0	0.0000	20.2109
	H4	0b697f53892f3e1c88ace	147	0	0.0000	20.3229
	H5	a247fa889d1c959e80e77	147	0	0.0000	20.4373

Table 5.6: Waveform A and Waveform B

Waveform	Elements (441 of Elements)
A	1 -1 1 -1 -1 1 -1 -1 1 1 1 1 -1 1 -1 1 -1 -1 1 1 -1 -1 1 -1 -1 1 1 1 1 1 1 1 1 1 -1 1 -1 -1 -1 1 -1 -1 1 1 -1 -1 -1 -1 1 1 1 1 -1 1 -1 1 1 -1 -1 -1 1 -1 -1 -1 -1 -1 1 -1 1 -1 1 -1 -1 1 1 1 -1 1 -1 -1 -1 -1 -1 -1 -1 -1 -1 1 1 -1 1 1 1 -1 -1 1 1 -1 -1 1 1 1 -1 1 1 -1 -1 -1 1 -1 -1 1 1 -1 -1 1 -1 1 -1 1 1 1 -1 1 1 1 -1 1 1 -1 1 -1 1 -1 -1 -1 1 -1 -1 1 1 -1 1 1 -1 1 -1 1 -1 -1 1 -1 1 1 -1 -1 1 -1 1 -1 -1 -1 1 -1 -1 1 1 1 -1 -1 1 -1 -1 1 1 -1 1 -1 1 -1 -1 1 -1 -1 1 -1 -1 1 -1 -1 -1 -1 -1 -1 1 1 -1 -1 1 1 1 1 -1 1 -1 - 1 -1 1 -1 1 1 -1 1 1 1 -1 -1 1 1 -1 1 -1 1 -1 1 1 -1 -1 -1 1 1 1 1 -1 1 1 -1 1 -1 -1 -1 1 1 1 -1 -1 -1 -1 1 -1 -1 1 -1 -1 1 1 1 1 1 1 -1 -1 1 1 1 1 -1 -1 -1 -1 1 -1 1 1 1 -1 1 -1 -1 1 1 -1 1 -1 1 -1 -1 -1 1 1 1 1 1 -1 1 -1 1 -1 1 1 -1 1 -1 -1 1 1 1 -1 -1 -1 -1 1 1 -1 1 1 1 1 1 -1 -1 1 -1 1 1 -1 1 -1 -1 1 -1 -1 -1 -1 1 1 -1 1 1 1 1 -1 1 1 1 1 1 1 -1 -1 -1 -1 -1 1 1 1 1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 1 -1 1 -1 -1 1 1 1 1 1 1 1 -1 -1 1 -1 1 -1 -1 -1 1 1 1 1 1 1 1 -1 -1 -1 1 1 -1 1 -1 -1 1 1 1 1 1 1 -1 -1 1 -1 -1 -1 1 1 1 1 1 1 -1 1 -1 -1 -1 1 1 1 -1 1 1 -1 -1 1 1 -1 -1 -1 1 -1
B	-1 1 1 -1 1 1 -1 1 -1 1 -1 -1 -1 -1 1 -1 1 -1 -1 -1 1 -1 - 1 -1 1 -1 1 -1 1 -1 1 -1 -1 -1 1 1 1 -1 1 1 1 1 1 -1 1 -1 -1 1 -1 -1 1 1 -1 1 -1 1 1 -1 1 1 1 1 -1 -1 -1 1 1 -1 1 1 1 1 1 1 1 -1 -1 -1 -1 -1 -1 -1 1 1 1 1 1 -1 -1 1 -1 1 1 1 -1 -1 1 1 1 -1 -1 1 -1 -1 1 -1 -1 1 -1 1 -1 1 1 -1 -1 -1 -1 -1 -1 1 -1 1 -1 1 -1 1 1 -1 -1 1 1 -1 1 -1 1 1 -1 1 1 1 1 1 -1 -1 -1 1 1 1 -1 1 -1 -1 -1 1 -1 -1 -1 -1 -1 1 1 1 1 1 1 1 1 -1 1 1 1 1 -1 1 -1 -1 1 1 1 1 1 -1 1 - 1 1 1 1 1 1 1 1 1 1 -1 -1 1 -1 1 -1 1 1 -1 1 -1 -1 -1 -1 -1 -1 -1 -1 -1 1 -1 1 1 -1 1 -1 -1 -1 1 1 -1 -1 1 1 -1 - 1 1 -1 -1 -1 -1 1 1 -1 1 1 1 -1 1 1 1 1 1 1 -1 -1 -1 -1 1 -1 -1 -1 -1 1 -1 1 1 1 -1 -1 -1 -1 1 -1 1 -1 1 1 1 1 -1 -1 -1 1 1 -1 -1 -1 -1 1 -1 -1 -1 1 1 -1 1 -1 1 1 1 -1 1 -1 -1 -1 1 -1 -1 -1 -1 -1 1 -1 -1 -1 -1 -1 1 -1 -1 -1 -1 -1 1 1 1 -1 -1 1 1 1 -1 1 -1 -1 1 -1 -1 1 1 1 1 1 1 -1 1 1 -1 -1 1 -1 1 1 -1 -1 1 -1 1 1 1 -1 -1 -1 1 1 1 -1 1 -1 -1 1 -1 1 1 1 1 -1 1 1 1 -1 -1 -1 1 1 1 -1 -1 -1 1 -1 -1 1 -1 1 -1 -1 1 -1 -1 -1 1 -1 -1 -1 -1 1 -1 1 1 -1 1 1 -1 1 -1 1 -1 -1 1 -1 1 -1 -1 -1 1 -1 -1 1 -1 -1 1 1 -1 -1 -1 -1 -1 1 1 1 -1 1 -1 -1 -1 -1

5.3 Research Constraints

Throughout the research, there are some constraints need to be discovered. First is regarding the watermark message, it should contain at least 15 words as the sampling rate is 44100 Hz per second which mean that the watermark message should embedded into the audio at least one second. In addition, another constraint is the length of watermark message. In order to get the better accuracy of the watermark message, the watermark should embed into a moderate signal by not more than 280 of bits message. Furthermore, the watermark message can only embed into WAV format of audio file. Lastly, due to the time limit, there is only two different of waveform to be tested in the experiment. It will be further carry out in future.

CHAPTER 6 CONCLUSION

6.1 Conclusion

Overall, the proposed scheme for digital audio watermarking by employing psychoacoustic model is tested and it shows that the watermark signal is imperceptible to the audience based on evaluated results, SNR. By using waveform A, the SNR is above 20dB which satisfies the requirement of IFPI. Furthermore, the experimental results revealed that by embedding different length of messages into the audio signals, the results has shown consistency in watermark embedding and extracting process which achieved at least 99.99% or even up to 100% of bit rate accuracy. As discussed in section 5.2, the number of erroneous bits is directly proportional to the total number of embedded bits, hence the accuracy of extracted data depends on the length of the messages. In addition, due to the results of extracted hash message as ideally as expected or achieved 100% of accuracy for the extracted message by using waveform A. Consequently the hash function is suitable to be used to verify the modification attacks by using the proposed scheme.

At the end of the research, the purposes of this research had been achieved which are to test digital audio watermarking by employing psychoacoustic model, to embed different length of messages to test on accuracy of extracted data, and to study the suitability on using hash function for verification of modification attacks. In conclusion, digital audio watermark is a good approach to hide owner information as it is imperceptible and invisible from the watermarked audio. Although there is still weakness that can be found in the proposed methods, but it still can be improved in future.

6.2 Future Direction

The possible future direction or works for the proposed watermark approach may aim at allow the watermark message against modification attacks. It is also aim at allow the watermark message to embed into MP3 format of audio file and against other attacks such as MP3 compression and any other attacks that available. Moreover, it is possible to improve the spread spectrum technique as the spread waveform may affect the accuracy of extracted data and try to achieve the BER at 0.00% and the bit rate of accuracy at 100%.

REFERENCES

Aeroflex Corporation (2012). Official Portal of Aeroflex Corporation. Retrieved from: *http://www.aeroflex.com/ats/products/prodfiles/appnotes/192/888a.pdf* (10 Dec 2012)

Agbaje, M.O., Akinwale, A.T. & Njah, A.N. (2011). *Audio Watermarking: ACritical Review. International Journal of Scientific & Engineering Research, 2(11)*, ISSN2229-5518

Amin, M.M., Salleh, M., Ibrahim, S., Katmin, M.R. & Shamsuddin, M.Z.I. (2003). *Information Hiding using Steganography.* Paper presented at 4th National Conference on Telecommunication Technology Proceedings, Shah Alam, Malaysia. Retrieved from October 19, 2012, IEEE database.

Baranwal, N. & Datta, K. (2011). *Peak Detection based Spread Spectrum Audio Watermarking using Discrete Wavelet Transform. International Journal of Computer Applications, 24(1)*, pp. 0975-8887

Baras, C., Moreau, N. & Dutoit, T. (2009). *How could music contain hidden information. Applied Signal Processing*. DOI: 10.1007/978-0-387-74535-0_7

Bassia, P., Pitas, I., & Nikolaidis, N. (2001). *Robust Audio Watermarking in the Time Domain. IEEE Transactions On Multimedia, 3(2)*, pp. 232 – 241

Cai, Q. & Chen, Y. (2011). *A WAV Format Audio Digital Watermarking Algorithm Based on HAS. International Conference on Control, Automation and Systems Engineering (CASE).* DOI: 10.1109/ICCASE.2011.5997686

Cvejic, N. (2004). Algorithms for Audio Watermarking and Steganography (Master's thesis, Department of Electrical and Information Engineering, Information Processing Laboratory, University of Oulu). Retrieved from http://herkules.oulu.fi/isbn9514273842/isbn9514273842.pdf

Fallahpour, M. & Megias, D. (2012). *High Capacity Robust Audio Watermarking Scheme based on FFT and Linear Regression. International Journal of Innovative Computing, Information and Control, 8(4)*, pp. 2477-2489

Geetha, K. & Vanitha Muthu, P. (2010). *Implementation of ETAS (Embedding Text in Audio Signal) Model to Ensure Secrecy. International Journal on Computer Science and Engineering, 2(4)*, pp 1308-1313

Ketcham, M. & Vongpradhip, S. (2007). *Intelligent Audio Watermarking Using Genetic Algorithm in DWT Domain. International Journal of Electrical and Computer Engineering, 2(6)*, pp. 429-433

Kim, H.J., Choi, Y.H., Seok, J.W & Hong, J.W. (2004). Audio Watermarking Techniques. In J.S. Pan, H.C. Huang & L.C. Jain (Eds.), *Intelligent Watermarking Techniques, Part 4* (pp. 185-217). Singapore: World Scientific Press.

Krenn, R. *Steganography and steganalysis*, An Article, January 2004.

Maha, C., Maher, E. & Chokri, B.A. (2008). *A blind audio watermarking scheme based on Neural Network and Psychoacoustic Model with Error correcting code in Wavelet Domain. The 3rd International Symposium on Communications, Control and Signal Processing (ISCCSP2008)*, Malta, pp.1138-1143

Mat Kiah, M.L., Zaidan, B.B., Zaidan, A.A., Mohammed Ahmed, A., & Hasan Al-bakri, S. (2011). *A review of audio based steganography and digital watermarking. International Journal of the Physical Sciences, 6(16)*, pp.3837-3850. doi: 10.5897/IJPS11.577

Mohanty, S.P. (1999). Digital Watermarking: A Tutorial Review. Reported at Indian Institute of Science, Bangalore. Retrieved from:

http://informatika.stei.itb.ac.id/~rinaldi.munir/Kriptografi/WMSurvey199

Mohanty.pdf

Naveen, D., & Dr Jhansi rani, A. 2010. Implementation of Psychoacoustic model in Audio Compression using Munich and Gammachirp Wavelets. *International Journal of Engineering Science and Technology*. 2(5): 1066-1072

Perbadanan Harta Intelek Malaysia (2012). Official Portal of Intellectual Property Corporation of Malaysia. Retrieved from: *http://www.myipo.gov.my/hakcipta* (12 Oct 2012)

Ravula, R. (2010). *Audio Watermarking Using Transformation Techniques.* Master Thesis. Louisiana State Univeristy and Agricultural and Mechanical College. Retrieved December 10, 2012, from Louisiana State University Electronic Theses & Dissertation Collection.

Singh, S. (2011). *Digital Watermarking Trends. International Journal of Research in Computer Science, 1(1).* pp 55-61

The MathWorks, Inc. (2012). *Documentation Center – Signal Processing Toolbox.* Retrieved November 26th, 2012, from

http://www.mathworks.com/help/signal/gs/product-description.html

Wang, X., Niu, P., & Yang, H. 2009. A robust digital audio watermarking based on statistics charecteristics. *Elsevier Ltd., Pattern Recognition*. 42 pp: 3057-3064

Wu, C.P., Su, P.C., & Jay Kuo, C.C. 2000. Robust and Efficient Digital Audio Watermarking Using Audio Content Analysis. *Department of Electrical Engineering-Systems.* 3971(2000): 0277-786X

Zhang, P., Xu, S.Z. & Yang, H.Z. (2012, March 23-25). *Robust and Transparent Audio Watermarking Based On Imporved Spread Spectrum and Psychoacoustic Masking.* Paper presented at the 2012 IEEE International Conference on Information Science and Technology, Wuhan, Hubei, China. Retrieved September 20, 2012, from IEEE database.

Zhao, X.M., Guo, Y.H., Liu, J. & Yan, Y.H. (2011). *A Spread Spectrum Audio Watermarking System with High Perceptual Quality.* Paper presented at the 2011 Third International Conference on Communications and Mobile Computing, ThinkIt Speech Lab, Institute of Acoustics, Chinese Academy of Sciencies 21 Beisihuan XiLu, Beijing, China. Retrieved from October 3, 2012, from IEEE database.

Zwicker, E. & Fastl, H. (1999). Psychoacoustics: Facts and Models. Germany : Springer Verlag, Berlin