AUDIO WATERMARKING USING DWT OF SECOND LEVEL DECOMPOSITION OF LOW FREQUENCY AND USING RMS ON APPROXIMATION COEFFICIENTS

HEYAM ESSAM JIBRAEL MARAHA

AUGUST 2014

AUDIO WATERMARKING USING DWT OF SECOND LEVEL DECOMPOSITION OF LOW FREQUENCY AND USING RMS ON APPROXIMATION COEFFICIENTS

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF NATURAL AND APPLIED

SCIENCES OF

ÇANKAYA UNIVERSITY

BY

HEYAM ESSAM JIBRAEL MARAHA

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF

MASTER OF SCIENCE

IN

THE DEPARTMENT OF

MATHEMATICS AND COMPUTER SCIENCE

AUGUST 2014

Title of the Thesis : **Audio Watermarking Using DWT of Second Level Decomposition of Low Frequency and Using RMS on Approximation Coefficients.**

Submitted by **Heyam Essam Jibrael MARAHA**

Approval of the Graduate School of Natural and Applied Sciences, Çankaya University.

Prof. Dr. Taner ALTUNOK
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Billur KAYMAKÇALAN
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Assist. Prof. Dr. Reza HASSANPOUR
Supervisor

Examination Date: 06.08.2014

Examination Committee Members:

Assist. Prof. Dr. Reza HASSANPOUR (Çankaya Univ.)
Assist. Prof. Dr. Abdül Kadir GÖRÜR (Çankaya Univ.)
Assoc. Prof. Dr. Fahd JARAD (THK Univ.)

STATEMENT OF NON PLAGIARISM PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct, I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name : Heyam MARAHA
Signature :
Date : 06.08.2014

ABSTRACT

AUDIO WATERMARKING USING DWT OF SECOND LEVEL DECOMPOSITION OF LOW FREQUENCY AND USING RMS ON APPROXIMATION COEFFICIENTS

MARAHA, Heyam Essam Jibrael

M.Sc., Department of Mathematics and Computer Science

Supervisor: Assist. Prof. Dr. Reza HASSANPOUR

August 2014, 75 pages

Digital media are distributed through the internet system in a magnificent manner due to the efficiency and inexpensiveness of such a system. Audio clips and other digital signals are transmitted, shared and used easily. This causes a real security problem. Solving this problem leads to finding authentication techniques to provide security for these digital media.

Digital watermarking is a suggested solution technique to offer such security for these distributed digital signals. Digital watermarking is the art of embedding information (a watermark) in the original signal, taking into consideration preserving the quality of the host digital signal.

There are different types of watermarking media such as image, video and audio; our thesis focuses on audio watermarking techniques.

Also, there are a variety of techniques using the watermarking algorithm that depends on the type of algorithm that is used to embed the watermark in the original signal. In our thesis, we have used DWT (Discrete Wavelet Transform) of second level decomposition of the approximation coefficients of each frame. These frames are resulted from segmenting the original audio clips into frames, each one equal to

10 seconds. Also, we have applied the RMS (Root Mean Square) to these approximation coefficients which represent the low frequency of each frame. The aim of using the RMS is to find the higher magnitude approximation coefficients among the segmented frames; then to embed the watermark in a quarter of the selected frames according to the magnitude order.

A watermark of 256 x 256 pixels was embedded in two tested audio clips (Pop and Classic) according to the former algorithm. We have applied 17 attacks that try to remove or modify the watermark that is hidden in the original audio clips. Also variety of models such as subjective and objective models were applied to evaluate the efficiency of the proposed algorithm results. The results of the proposed watermarking technique show an effective robustness against most of the applied attacks, especially in MP3 Compression, Quantization, Additive Noise, Resample and Low Pass Filter; however, it shows fragile robustness in Time Stretch and Pitch Shift attacks.

Keywords: Digital Watermarking, Discrete Wavelet Transform (DWT), Root Mean Square (RMS), Robustness.

ÖZ

DÜŞÜK FREKANSI İLE İLGİLİ DWT İKİNCİ DÜZEYİ AYRIŞTIRMA VE TAHMİNİ KATSAYILARI İLE İLGİLİ RMS KULLANIMI İLE SESLİ DAMGALAMA

MARAHA, Heyam Essam Jibrael

Yüksek Lisans, Matematik-Bilgisayar Anabilim Dalı

Tez Danışmanı: Yrd. Doç. Dr. Reza HASSANPOUR

Ağustos 2014, 75 Sayfa

Dijital ortamlar, sistemin etkinliği ve düşük maliyetli olmasından dolayı internet sistemi üzerinde çok büyük miktarda paylaşılmaktadır. Sesli klipler ve diğer dijital sinyaller iletilmekte, paylaşılmakta ve kolaylıkla kullanılmaktadır. Bu da büyük bir güvenlik problemine neden olmaktadır. Bu problemin çözülmesi, söz konusu dijital ortamlar için güvenliğin sağlanması amacıyla doğrulama tekniklerinin keşfedilmesini gerektirir.

Dijital damgalama, dağıtılan orijinal dijital sinyallerin güvenliğini sağlamak için önerilen bir çözüm tekniğidir. Dijital damgalama, alıcı dijital sinyalin kalitesinin korunmasını dikkate alarak bilgileri gömme (filigran) sanatıdır.

Medya ile ilgili görüntü, video ve sesli damgalama gibi farklı damgalama yöntemleri kullanılmaktadır; araştırmamız sesli damgalama tekniklerine odaklıdır.

Ayrıca algoritmaların damgalanması için orijinal sinyalin içine damganın gömülmesi için kullanılan algoritmanın tipine bağlı olarak birkaç teknik kullanılmaktadır; tezimizde her biri 10 saniye ile eşit olan orijinal sesli klip parçalarından ibaret her bir

çerçevenin tahmini katsayılarının ikinci düzeydeki ayrıştırmaları DWT (kesikli dalgacık dönüşümü) kullanılmıştır ve her çerçevenin düşük frekansını temsil eden bu tahmini katsayılarına RMS (ortalama kare kökü) uygulanmıştır. RMS kullanma amacı parçalanmış çerçeveler arasındaki en yüksek tahmini katsayısını bularak damganın büyüklük düzenine göre seçilmiş çerçevelerin dörtte birine gömmek idi.

İki adet test edilmiş sesli klip içine 256 x 256 piksel değerinde damga gömülmüştür (Pop ve Klasik). Birinci algoritmaya göre orijinal sesli klipte gömülü olan 17 adet saldırı ile damga çıkarılmaya veya değiştirilmeye çalışılmıştır; aynı zamanda önerilen algoritma sonuçlarının etkinliğini değerlendirmek üzere farklı çeşitteki sübjektif ve objektif modelleri uygulanmıştır. Önerilen damgalama teknik sonuçlarına göre özellikle MP3 sıkıştırma, niceleme, toplanır gürültü, yeniden örnekleme ve alçak geçirgen filtre açısından uygulanan saldırılara karşı sağlam bir direniş sergilemekte ancak Zaman Uzatımı ve Yükseklik Değişimi konusunda hassasiyet sergilemektedir.

Anahtar Kelimeler: Dijital Damgalama, Kesikli Dalgacık Dönüşümü (DWT), Ortalama Kare Kökü (RMS), Sağlamlık.

ACKNOWLEDGEMENT

First of all, I want to thank GOD for all the graces that he has given me, especially in carrying out this thesis.

I would like to express my deepest gratitude to my thesis Advisor Assist.Prof. Dr. Reza HASSANPOUR for his guidance, support and encouragement throughout the research.

My special thanks to my family, especially to my parents, brothers and sisters for their endless advice, bestowals and support for me through all my life. I would like also to thank my husband for all his patience and support. Finally, I want to thank my sweet kids, Farah and Yousef, GOD bless them, for being the motivation in my life.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURES

FIGURES

LIST OF TABLES

TABLES

LIST OF ABBREVATIONS

DWT	Discrete Wavelet Transform
DCT	Discrete Cosine Transform
FFT	Fast Fourier Transform
HAS	Human Auditory System
HVS	Human Visual System
IFPI	International Federation of the Phonographic Industry
LPF	Low Pass Filter
MOS	Mean Opinion Score
MSE	Mean Square Error
PSNR	Peak Signal to Noise Ratio
RMS	Root Mean Square
SDG	Subjective Difference Grade
SR	Similarity Ratio
SDMI	Secure Digital Music Initiative
TSM	Time Scale Modification
AWGN	Additive White Gaussian Noise
BER	Bit Rate Error
ECC	Error Correct Code
PRN	Pseude Random Number
SNR	Signal to Noise Ratio
LDPC	Low Density Parity Check
AOAA	Average Of Absolute Amplitude
GOS	Group Of Samples
ANN	Artificial Neural Network
ASDAW	Adaptive Single Dependent Audio Watermarking

MAE	Mean Absolute Error
SVD	Singular Value Decomposition
LSB	Least Significant Bit
NCC	Normalized Cross Correlation
ODG	Objective Difference Grade
KFCM	Kernel Fuzzy C_Mean
SVM	Support Vector Machine
LCM	Log Coordinate Mapping
VQ	Vector Quantization
NC	Normalized Correlation
NDCT	Non-uniform DCT
SFN	Selected Frame Number

CHAPTER 1

INTRODUCTION

1.1. The Problem and Motivation

With the development of the internet and the great usage of multimedia techniques through such an environment, the protection for these multimedia techniques is a real problem that has emerged which must be solved effectively.

The necessary method to solve such a problem, the digital watermarking technique which is the process of hiding information in digital multimedia, occupies a great amount of attention nowadays, especially for multimedia that is distributed via internet.

The embedded watermark or the information should be imperceptible and cannot be easily distorted. For this reason an effective watermarking algorithm should be improved.

1.2. Scope of Thesis

Digital watermarking techniques are classified according to the type of multimedia into four types, image, audio, video and text. Our thesis will examine the digital audio watermarking technique with the DWT (Discrete Wavelet Transform) of second level decomposition for the low frequency coefficients.

1.3. Summary of Proposed Method

The proposed watermarking method is an embedded watermark of 256×256 bits in an audio. Two level of DWT decomposition is applied to the two tested audio clips, Pop and Classic.Our algorithm suggests to segment the audios into frames, each frame being about 10 seconds. The second level of DWT decomposition for approximation coefficients is applied on these frames; then the RMS (Root Mean Square) model is applied on the second level of approximation coefficients to find the higher magnitude approximation coefficients among these frames. The watermark is embedded in the low frequency of the quarter of the higher magnitudes frames, the extraction algorithm is a blind algorithm that needs the original audio in the extraction method.

1.4. Summary of Results

The proposed audio watermarking algorithm is tested against seventeen malicious attacks that try to remove or to alter the embedded watermark; we have tested the qualification of our proposed watermarking algorithm according to the watermark requirements by using the objective modules PSNR (Peak Signal -To- Noise Ratio) and SR (Similarity Ratio), and also by applying the subjective modules MOS (Mean Opinion Score) and SDG (Subjective Difference Grade). All of these modules confirm that the proposed method is robust against tested attacks, especially for MP3 compression of different parameters, Resample of different parameters, Additive Noise attacks (Gaussian, NSR_11,NSR_15,Standard Deviation) and Quantization, but it shows a fragility against Time Scale and Time Shift attacks.

CHAPTER 2

STEGANOGRAPHY AND WATERMARKING

2.1. Steganography

Steganography, which is a Greek word consists of two parts: ‘stegano’ and ‘graphy’.The first has the meaning of ‘covered’ while the second part means ‘to write’, and by combining these two words we will get ‘covered writing’[1] i.e. “to hide in plain sight” [2].

Steganography is a science-art that is performed to communicate with hidden messages that are not easily detected by unauthorized persons, and the hidden messages will only be discovered by a specific group, i.e. “to avoid drawing suspicion to the transmission of a secret message” [3], any attempt to detect the Steganography will cause an attack on the secret message.

Steganography and Encryption are used to provide data particularity, but in Encryption everybody can see that the communication performed between two parties was secret, unlike in Steganography with which nobody can notice that communication is done secretly between two parties. Steganography can’t be removed without modifying the embedding data, unless the attacker can find a way to remove it privately [2]. Steganography was used for many hundred years. The need to transmit digital files leads to the invention of new techniques suitable for these needs [2].

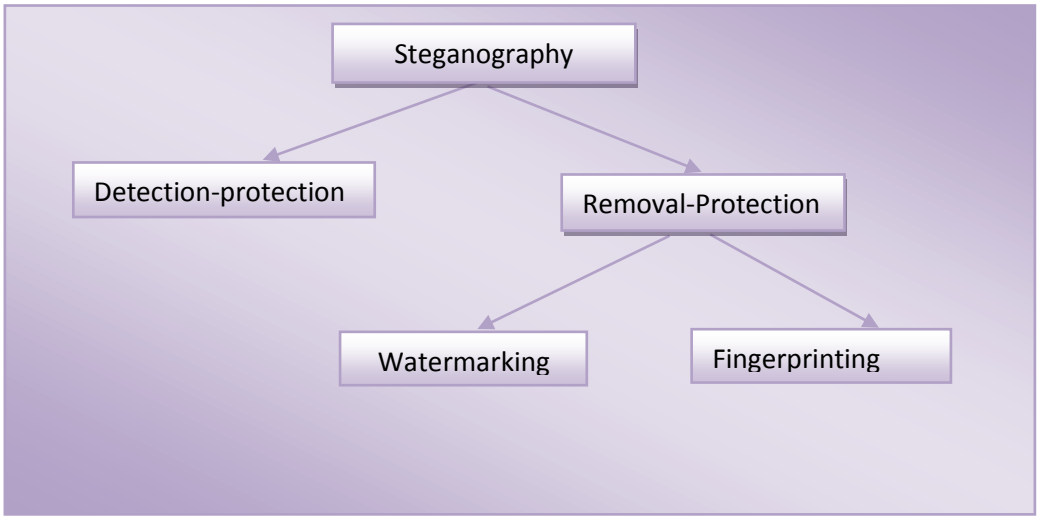

Figure 1: Types of Steganography picked from [30]

2.2. Digital Watermarking

In the age of digital information, internet communication and digital transmission via public media, many cryptography techniques were used to transmit secret information through global communication media.A secret key must be known for a two parties that participate with secret communication, and this causes the low level secrecy that resulted from such insecure old techniques.

The need for an authenticity of digital information has appeared, so to solve the present needs a Watermark technique was invented. The Steganography idiom is not known as by other terms (like watermarking, embedding information or hiding data). The watermarking term rates as the most popular concept [1]. The Watermark is related to Steganography in its concept but differs in its goal [4]. Steganography has a digital signal with no relation to the message embedded in it. Actually, it makes use of the digital signal to cover the important message, unlike the Watermark that embeds a message that has a factual relationship to the digital signal.

Watermark, which is also called watermark embedding is a technique through which secret information is embedded (or in other words, some new data is added to the original [1]) in the carrier signal without perceptual degradation to the host multimedia.

The hidden information or the new data that are intended to be added to the cover data or original information can have a variety of types; it could be a binary-image or a grayscale-image, and it also could have the type of text, signal control, audio, serial numbers or random number PRN, and it could be copyright messages, and so on [1].

2.2.1. Watermark Block System

All of the watermark techniques are sharing two- common-constructing-blocks, a watermark embedding system and a watermark extracting system [5]. In the embedding scheme a secret message is hidden in the original signal; while in the extraction scheme, the watermark should be extracted or detected in order to prove the ownership and copy control.

The inputs in the embedding process are an original signal (I), a watermark (W) (a binary image, PRN number or a text), and optional secret key (K), on the other hand the output of this system is a watermarked signal (I'). Fig (2) represents the Embedding System.

In the Extracting system the inputs are the tested signal (I'), sometimes the watermark (W) and the secret key (K) (optional). The output of this process is the extracted watermark (W'). Fig (3) represents the Extraction System.

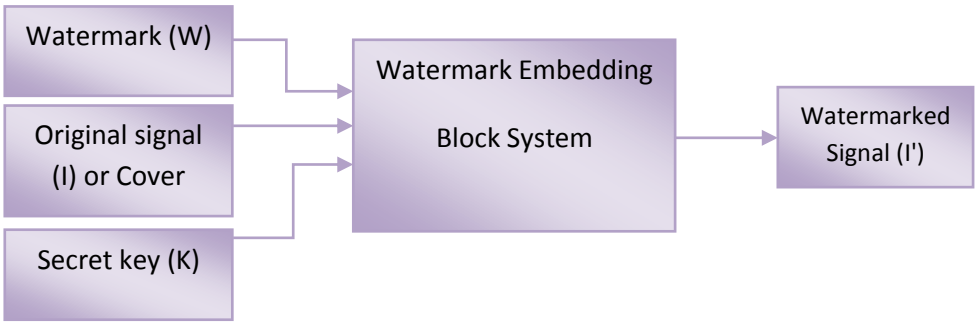

Figure 2: Watermark embedding system

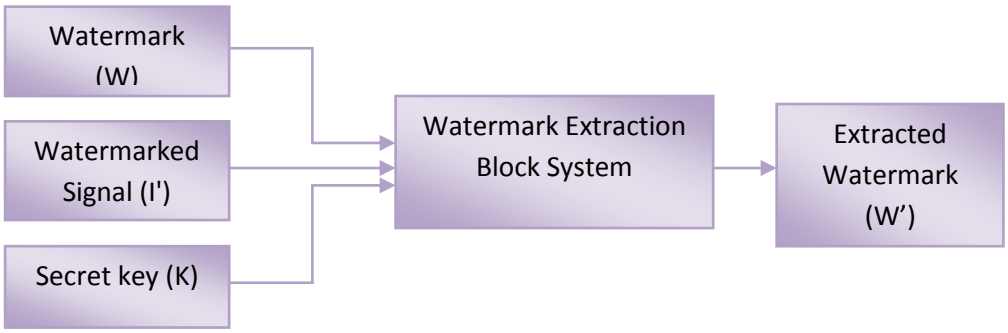

Figure 3: Watermark extracting system

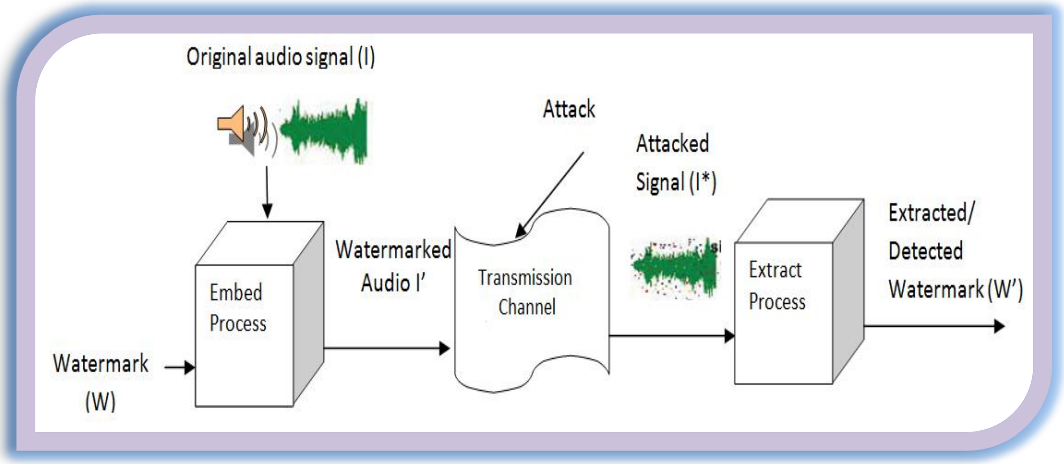

Figure 4: The watermarking system

The Watermark System can be classified according to the combination of inputs and outputs that separate the Embedding and Extracting Systems into three different types, these systems are classified as follows [6]:

1. Blind Watermarking (Private Watermarking):- Needs at the very least the host signal (I) in the detection process for the watermark $(I, I', K, W \rightarrow \{0, 1\})$.

2. Semi_Blind Watermarking (Semi- Private Watermarking):- to detect for the watermark this system doesn't require the original signal (I), but may need any of the following parameters: $(I', K, W \rightarrow \{0, 1\})$.
3. Non_Blind Watermarking (Public Watermarking):- this system is the most difficult and challenging one because it needs neither the original signal (I) nor the watermark signal (W), but may need these (I', K, W').

2.2.2. Requirements of the Efficient Watermark Technique

According to **IFPI** (**International Federation of the Phonographic Industry**) [7], an efficient watermark should satisfy some important requirements. The most historic requirements and the trade_offs among them are as follows [8]:

1. Transparency (imperceptibility, fidelity**)**

Imperceptibility refers to perceptual transparency, a certain feature that an efficient watermark should have, through which the watermarked signal should not be distorted or changed significantly from the original signal. For an audio signal the watermarked audio should be inaudible by the human auditory system (HAS).
According to IFPI, the **PSNR,** which is the (PeakSignal_ to _Noise Ratio) should be preserved at over 20 dB to satisfy the Watermark-Transparency requirements. Whenever the Transparency is increased, the robustness and security against attacks will be decreased:- this is the challenge among the requirements in trade offs.

2. Reliability (Robustness)

Reliability or Robustness is another important property of an effective watermark technique [4]. The work of the watermark may be changed throughout its lifetime, and these changes may result from transmission media or by attacks. The attacks can be summarized by Secure Digital Music Initiative (SDMI) [7], which is an online form that contains the digital copyrights. These attacks are like noise addition, lossy compression, digital-to-analog and analog-to-digital conversions , time scale modification (TSM), band-pass filtering, sample rate conversion and echo addition [9].

This feature means that the watermark should be robust against processing and malicious attacks and the watermark should be detected only by authorized parties to prove their ownership.

3. Capacity (or Payload Size)

This property refers to the amount of information to be embedded in the original signal, and the efficient watermarking technique should be able to hide a large amount of data within its host signal taking into account that distortion (that occurs from adding new information of the watermark to host signal) should be kept as minimum as possible, and it is better to include the watermark all over the host signal to avoid for example the crop attack.

This feature struggles against the Robustness and the Imperceptibility properties:- any increase in the capacity size will reduce the effectiveness of the two mentioned features [10].

4. Security

The Security feature denotes that the embedded watermark should not be removed or modified without causing damage to the host signal; the strength of this property depends on the proposed embedding algorithm [10].

This requirement tries to keep the watermark secure and solid against malicious attacks like the other requirements.

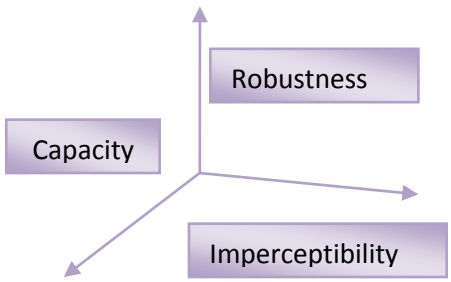

Figure 5: The Robustness, capacity and imperceptibility and trade off among them

2.2.3. Watermark Applications

1. Owner Identification and Ownership Protection

In the owner identification (owner proof) application, a unique- secret watermark is hidden in the host media, and in the case that any unauthorized parties change and claim that this digital media belong to them, then the genuine owner of this multimedia file can show the original watermark to prove ownership [9].

2. Fingerprinting

Control and owner identification applications embed the same watermark in all copies of a multimedia file, but with electronic_content_distribution permits each copy has a unique watermark and customized to each individual related recipient, this operation enables the distribution party to identify illegal usage of this content, this application is called Fingerprinting [11].

3. Broadcast Monitoring

This application is concerned with the broadcasting. A unique watermark is embedded in each audio or video before broadcasting the multimedia file, in order to monitor the broadcast operation. Several organizations like musician, advertiser and film producers are interested in this field [11].

4. Authentication

This application is concerned with the cryptography problem, in which the cryptographic signature is computed related to an image. Any bit, altering this image will affect the signature, and this implies detecting tampering in that image. Also, if the metadata signature is embedded in the header of an image, any copy of this image will cause signature loss, so this problem is avoided by embedding a signature in the desired image through the watermarking process. This will ensure the signature will be firm in the image. JPEG compression is one of the application systems that allows making rough changes to the digital signal.

5. Copy Control

The proof and monitoring ownership watermarks do not restrain illegal copying of a digital signal. When an audio clip is recorded and it is watermarked so this copy will fail, all of the manufactured recorders should contain detection circuitry of the watermark.

6. Monitoring Air Traffic

To avoid a problem occurring from a message sent from a pilot to the ground monitor that this message may be attacked, a flight-number is sent with the message voice as a watermark to the ground monitor and this will offer more security to this operation[9].

7. Medical Applications

This is applied in medical fields by embedding the patients' names in their X-ray file as an example and all patients have their unique watermarks [9].

2.2.4. Watermark Types

According to the type of document or the file that has been chosen to be watermarked, the Watermarking Techniques can be classified into four groups, as follows [10]:

1. **Image Watermarking.**
2. **Audio Watermarking.**
3. **Video Watermarking.**
4. **Text Watermarking.**

Our work will be concerned with Audio Watermarking, so we will discuss only this technique in chapter 3.

2.2.5. Watermark Attacks (Audio Attacks)

There are many attacks that try to remove or alter the watermark that is embedded in the host signal or tries to make the watermark undetected so the ownership of this multimedia will be lost. Some of these attacks are performed on all the types of watermarking and the others specialize in one type of watermarking. Because our work in watermarking is about Audio Watermarking as we shall notice in chapter 3, we will focus on audio attacks.

1. Dynamics

This occurs result from the amplitude modification and feebleness, like compression, expansion and re-quantization [9].

2. Ambiance

That's resulted when some groups try to record an original host of others and they claim that this signal belongs to them [9].

3. Filtering

This type of attack is like a high and low pass filter, in which it tries to amplify or lessen some part of the host signal [9].

4. Noise Additive

Like Additive White Gaussian Noise (AWGN), it is noticeable that through the signal transmission, the signal is affected by noise attacks that result from transmission; therefore to make the watermark robust it should be tested against such attacks [9].

5. Conversion and Lossy Compression

The audio in this structure depends on sampling frequency and bit rate. Some programs are invented to change these features, like MP3 compression and re-sampling, so a good watermarking technique should enable the watermark to sustain this kind of attack with a minimum rate of distortion [9].

6. Time Stretch Modification (TSM) and Pitch Shift

This kind of attacks tries to increase or decrease the length of the host audio without changing its pitch, or to change its pitch without changing its length. These attacks occur through the transmission operation like Jittering, which is an attack of this type [9].

2.2.6. Overview of Audio Watermarking Techniques

"The audio signal in its original format is a mixture of high and low frequencies that form semi-sinusoidal waves with uneven amplitudes and wavelengths" [12], taking into consideration the low amount of information to be embedded in the audio signal that belongs to its sensitivity against noise rather than video or image signals, i.e. the Human Auditory System (HAS) is much more sensitive than the Human Visual System (HVS), the audio watermarking techniques faced many challenges. One of them, as we mention, has a higher sensitivity nature against noise; also the audio clip is shorter than video and the amount of information that the audio clip owns is less than the other multimedia. This will make the audio watermarking technique much more difficult than the others, so to choose the best watermarking algorithm with inaudible distortion in original audio is a hard mission. Generally the watermarking techniques can be classified according to the work domain into two major embedding algorithm groups:- the first group is interested in working with the **Spatial (Time) Domain** and the second groups prefers to work in **Frequency Domain**. These groups challenge the watermark requirements trade off and try to balance among them to have the most imperceptible, robust, secure algorithm.
We will review the work of others in both fields in our literature review as follows:

2.2.6.1. Spatial Domain

The most straightforward watermarking technique through which the watermark is embedded into a host signal directly with no need to transform the domain, this method is so simple to achieve and it requires a low computation time, but this technique is not robust against attacks, especially crop attack and the embedded watermark is removed easily. Therefore, this method is not useful in copyright protection or ownership proof. There are several algorithms in Time Domain, and the LSB is an example of this technique. In our work on audio watermarking, we include the work of others in Time Domain. Through reviewing their work, we discuss the whole work and explain the advantages and disadvantages as follows:

1. Xiaoming, Xiong and Zhaoyang [13]. The researchers of this paper worked on the histogram specification in time domain to improve the watermark-resistance towards common attacks. The team in their algorithm depends on audio data analysis, which consist of analysis of invariant features, analysis of the histogram specification with four consecutive bins and data range analysis and design. In an analysis of invariant features and, the researcher found the values of audio mean and standard deviation before and after LPF (Low Pass Filter) processing and they were less than 2% and 0.5% respectively.

In the analysis of the histogram specification with four consecutive bins, they found the relationship among four consecutive bins using a special formula, then found the histogram depending on this relationship after applying LPF and ± 10 TSM attacks.

They also worked on **the Segmenting hiding idea for audio data** and **Segmenting watermarking algorithm based on histogram specification**. They segmented the audio signal to choose the best region for embedding depending on the histogram. Also, they segmented the watermark into parts, and then each part was embedded in different region.

In the Segmenting embedding algorithm, they use the following formula:

1. $(a + d)/(b + c) \geq T \; if \; W_i = 1$

2. $(b + c)/(a + d) \geq T \; if \; W_i = 0$ Where a, b, c, d are the number of samples in the four watermarked bins.

T threshold to control the embedding distortion and watermark robustness, T should be not less than 1.05 or 1/T should be less than 0.95.

If the bit =1 and it satisfies the equation (1), it means that the sample is suitable for embedding, or the algorithm alters the four consecutive bins until it satisfies formula 1, and by extracting a few samples from bin2 and bin3, they will be added to bin1, bin4 respectively.

If the bit=0 and satisfies eq. (2), then this is the desired region or the values a, b, c, d will be altered, and by extracting a few samples from bin1 and bin4, they will be added to bin2, bin3 respectively.

The extraction method doesn't need the previous work: it will extract bin1, bin2, bin3 and bin4 using a special formula.

The team used an audio of the 20s, mono, 16 bits per sample, and 44.1 kHz. The audio is segmented into 6 equal-sized parts for a watermark of 60 bits.

If we want to discuss the advantages of the proposed algorithm in the robustness field, the researcher makes a noticeable comparison of their work with other papers on LPF attack. As a result the proposed algorithm shows good resistance more than in other papers. They use BER to evaluate their work robustness:- the Cutoff of LPF (10_orderButterworth, kHz) in the Comparison-Table when it was equal to 7, 6, 4, BER will equal (0) for this paper algorithm while the other equal respectively 2/60, 9/60, 16/60. They attribute this to the usage of 4 consecutive bins.

Another advantage reported by the writers was the high capacity of 2 watermarks to be hidden in the desired signal which were Chinese letters, which increases the robustness for watermarked audio.

The segmentation and use of histograms depending on four consecutive bins, make this work good toward LPF and fair against some other common attacks, but that is taken into consideration that there is no evaluation for common attacks. Also in spite of the good resistance using two watermarks, I think that the audio is very sensitive to undertake such a capacity. Finally, working in time domain still has many challenges to prove its good resistance while working in frequency domain gives many guarantees in robustness and in inaudibility (audio). We shall notice in this

research through the next chapter, the watermark improvement through using such an algorithm of frequency domain.

2. Wen-Nung Lie and Li-Chun Chang in [14] worked on the Time Domain in embedding procedure. They perform their experiment on three different kinds of music:- Dulcimer, Symphony and Popular. Each audio signal is about 30-60 seconds and has 44.l kHz, L = 300 samples and initial d' = 0.05 for the psychoacoustic model test are an example parameter were used in embedding stage, they applied the bind extraction algorithm (no need for the original signal).

Watermark Embedding Scheme: In this method the researchers portioned the original audio signal into three consecutive sample sections, each sample of length L, the three samples indicated as sec_1, sec_2 and sec_3 . The proposed algorithm of embedding depend on the energy relation among these three consecutive samples.

These different energies of the three samples were denoted and arranged according to their energy as E_{max} , E_{mid} and E_{min} , then the group working on this paper tried to find the difference between these energies as the following:

A=E_{max}- E_{mid}

B= E_{mid}- E_{min}

Using a watermark image that has been transformed into a binary bit stream, the watermark was embedded into an audio signal in the selected three energy consecutive samples according to two formulas, one of these formulas for a bit (1) and the other for bit (0). The group tried to embed only one bit of information in each of the three consecutive samples.

Watermark Extraction Scheme: Wen-Nung and Li-Chun in the extraction method tried to segment the watermarked audio as in embedding and worked on each three consecutive samples to extract bit (1) and bit (0), dealing with the three energy parameters and the special difference equation among them. Here as we mentioned previously in this paper, the extraction is blind; in other words, there is no need for the host audio.

Continuity of Audio Waveform: To overcome the problems that occurred from discontinuities occurring between the boundaries of adjacent sections that affected the quality of an audio, there was progressive weighting near section boundaries.

Psychoacoustic Model Test: to curb the disturbances tangible to the human ear, the key for this solution is to compel the watermark energy to be under the masking thresholds and this was done by trying to get benefit from the frequency domain. In this paper they used the FFT transform.

Error Correction Coding: Before embedding error correcting codes (ECC) was used to increase the watermark robustness, a seed of the pseudo random number generator is used to permute binary bits in order to increase the watermark security. This paper has advantages in many fields, the researchers worked on three different music's signals, they tried to embed watermarks in consecutive samples, and by doing this they will overcome the synchronization problem, i.e. any attempt to add or to cut from such a sample will be difficult. Also, they try to benefit from the FFT frequency transform to overcome the audibility problem. In addition, the worker-group on this paper perform a variety of attacks like MP3 compression, low-pass filtering, amplitude normalization and digital-to-analog/ analog-to-digital. They used Bit Rate of different watermark size with MP3 compression, and the result they obtain for correlation is about 98%, which is magnificent result. They found that for LPF(Low Pass Filter) Popular music has a higher error rate (1.9-2.8%). They also applied DA/AD attacks, and they found the correlations were 100% for all tests in spite of alignment errors. In spite of all the good results that were obtained from working on such algorithms in the Time Domain and getting benefit from the Frequency Domain, still embedding in the frequency domain has the best result in embedding and in robustness. The frequency domain keeps a balance between the imperceptibility and robustness with less effort performed than Time Domain. This what we shall notice in chapter three of my research.

3. Md. Rifat Shahriar et al. in [15] worked in Time Domain. The global embedding scheme was done in two different marking spaces which were obtained from the original audio. They invested the properties of Polar coordinate system of the host audio. The group, which worked on this paper used two different watermark messages to be embedded in the two different marking spaces.

Proposed Audio Watermarking Scheme:

The two different marking spaces were produced from examining the effect of MPEG 1 Layer 3 compression of the host signal. The two different watermarks were embedded in this two marking space after they decomposed into two more marking

spaces. After embedding is done, the two spaces is combined with original audio using a special method.

Embedding Watermark Scheme: The audio signal $C_{o(i)}$ consists of M samples:

Step1. Applying MPEG 1 Layer 3 compression and decompression of the original audio $C_{o(i)}$ to produce $C'_{o(i)}$.

Step2. Find out the difference between $C_{o(i)}$ and $C'_{o(i)}$ to obtain points in mark spaces $V_{o(j)} = C_{o(i)}$ that will satisfy some desired conditions.

Step3. The mark space obtained from Step2. $V_{o(j)}$ is decomposed into $V_{ox(j)}$ and $V_{oy(j)}$ by applying the two equations that depend on the relationship between Polar and Cartesian system, the $\emptyset$ value of 45^0.

Strep4. $M1(s)$ And $M2(S)$ two different messages embedded in $V_{ox(j)}$ and $V_{oy(j)}$; a=embedding constant is set to the value of 0.13.

Extraction algorithm: The proposed algorithm for watermark extraction was performed with informed detection.

Step1. $V_{o(j)}$ And $V_{wn(i)}$ are obtained from $C_{o(i)}$ and $C_{wn(i)}$ respectively using the watermark key1.

Step2. $V_{ox(j)}$, $V_{oy(j)}$ and $V_{wnx(i)}$, $V_{wny(i)}$ respectively constructed from $V_{o(j)}$, $V_{wn(i)}$ using key 2.

Step3. By applying some special equation on the constructed mark space from the former step, watermarks are detected.

To discuss the present paper, the researchers examined ten different audios of 4 Sec. sampled at 44.1 kHz, 16 bit quantization and mono channel, the two watermarks of 1024 lengths.

The performance and the imperceptibility of the proposed algorithm were measured by using the SNR (signal_ to_ noise ratio) method. We can notice that almost all values of SNR for the ten clips produce a good result. The values ranging from Trumpet=25.66, which is the lowest value compared with Classic1=31.68, the highest value. For the robustness, the researchers experienced the proposed algorithm against twelve attacks such as Amplitude Compression, BPF, Crop, Echo, FFT- Real Reverse, LPF and, Noise Addition. Most of the recovery rate of robustness for these attacks have good robustness, but FFT Real Reverse has bad results, and Echo has some fair values and some bad values.

If we want to evaluate this work that was done by Md. Rifat Shahriar, Sangjin Cha and Vi-pil Chong, it was good work in Time Domain and that was clear from the SNR values and robustness against attacks, but the embedding of two watermarks in the same audio will harm this audio signal and they could discover a better way to embed optimum watermark in the host audio. Also the Time Domain is not desired to work, unlike the Frequency Domain.

4. Martinez-Noriega Raul, Mariko Nakano and Kazuhiko Yamaguchi in [16] were interested in Time Domain audio watermarking based on a self- synchronous decoding algorithm which uses low-density parity- check (LDPC) codes. The team of this paper depended on their work on another paper of W. N. Lie and L. C. Chang "Robust and high-quality time-domain audio watermarking based on low-frequency amplitude modification", by making improvements in the mentioned paper. The Lie et al. paper was used by this paper's researchers and in brief words the audio segmented into GOS group of samples which defines audio segments, of consecutive L samples. Each of GOS^iconsist of three sections S_1^i, S_2^i, S_3^i , of length L^i=L_1^i+L_2^i+L_3^i respectively, each bit will be embedded in different GOS and this can done by changing the average of absolute amplitudes (AOAA) of S_1^i, S_2^i, S_3^i, sections, the (AOAA) compound of a E_1^i ,E^i_2, E_3^i, that could be founded by using special formula for each parameter. These parameters are sorted to E_{max} ,E_{mid} , E_{min}, and by applying some other formulas depending on the resulted ones, bit $'1'$ is embedded according to the suggested approach and bit $'0'$ to different one to the host audio. The watermark is protected with a half-rate convolution-code, the Viterbi algorithm is used to recover the watermark, while the proposed algorithm used (LDPC) code in encoding the watermark.

Lie's algorithm used synchronization code by concatenating it at the beginning of each bit stream unlike the proposed algorithm which used a self- synchronous algorithm.

The proposed algorithm uses the watermark detection. The researchers of this paper did not know where the watermark is embedded; therefore, they need the synchronization code.

Now if we want to review this paper, the proposed algorithm picks out 30 random segments from three different audio files: "Egmont Op. 84" (7 min.), "Billie Jean" by

M. Jackson (4 min.) and “A Change of Seasons” by Dream Theater (21 min.). The watermark will be embedded in these 30 selected segments. These audios are mono-Wave-format sampled at 44.1 kHz and quantized at 16 bits, and the LDPC codes of length 96.The present work of this paper made improvements in the algorithm suggested by Lie et al. by using a self- synchronous decoding algorithm which uses LDPC codes, by achieving those improvements, a higher payload will be obtained depending on avoiding synchronization codes, and the watermark will be more robust if it will be compared with the Lie paper. The proposed algorithm shows BER values in the comparison table that contains attacks of the LPF and MP3 compression, and the table proves that the proposed algorithm of the present work was more robust in these attacks. In our opinion, to evaluate this work, it was good work to improve former papers on time domain by avoiding some problems mentioned in our review, but they used small LDPC codes that decrease the robustness of the watermark. We advise increasing the LDPC codes without increase the embedding complexity and audibility. Besides that it was better to improve the work in the frequency domain, a more robust field than the time domain.

5. Bassia Paraskevi, Ioannis Pitas, and Nikos Nikolaidis in [17] worked on audio watermarking in Time Domain processing by modifying the amplitude of each audio sample. To determine the characteristics of the modification they depend on the host audio and the copyright key. The watermark method of the proposed algorithm does not need the host audio signal in its detection of the watermark. To generate the watermark, the researchers used a special key which is a single number known to the copyright owner only. The watermark of this paper shows robustness against common attacks like MPEG audio coding, time shifting and, re-quantization, filtering, cropping, time shifting, re-sampling.

Five durations of two classical, pop and ethnic music themes were used in a subjective quality evaluation table to certify the inaudibility between the original and the watermarked signal. All listener evaluations had the score of 5 for all subjects, which is equivalent to a mean opinion score (MOS) of 5. Making a comparison with other papers from this table, the proposed work has a very good result. One of the disadvantages of this work is that the group working on this paper could not detect a watermark in an audio, attributed to a change in the time scale.

6. In the paper of [18] Tsai Hung-Hsu and Ji-Shiung Cheng offered an algorithm based on the characteristics of HAS (human auditory system) and neural networks. They propose (ASDAW) an Adaptive Signal-Dependent Audio Watermarking method. This method depends on both temporal (either pre-masking or post-masking) and frequency domain, which are the characteristics of HAS to generate the watermarks. The watermark of (ASDAW) was embedded in the time domain of the original audio.

Tsai Hung-Hsu et al. used another technique which is called (ANN), an artificial neural network ANN is trained in the ASDAW method to have the TANN technique to memorize a relationship between the watermarked and the original audio.

Based on the TANN and ASDAW techniques, the signal-dependent watermarks were extracted with no need for the original audio signal. The proposed work overcomes the deadlock problems affected by the works of other papers (this means the others try to detect the watermark rather than extract it). They extracted the watermark with no need for the presence of the original signal. The team of this paper used the temporal masking to eliminate the pre-echoes while generating a watermark. They get benefit from temporal and frequency domains to enhance the inaudibility. Another advantage of this work the fabulous feature of the ASDAW technique, is that to enable each audio to have its own identifiable signal-dependent watermark, i.e. a unique fingerprint for every audio.

The performance of watermarking techniques was evaluated using MAE (Mean Absolute Error). Also the present work evaluated the robustness towards signal attacks like MP3 compression/decompression (ISO/MPEG-I Audio Layer III), filtering, multiple manipulations and temporal re-sampling. This work has some advantages of using new techniques to enhance the inaudibility and to find a single watermark depending on the HAS system, but on the other hand the proposed work shows lower robustness against former mentioned attacks which belong to the work in the time domain.

2.2.6.2. Frequency Domain

The second group interested in working in the Frequency Domain, the watermark, is embedded in the original signal after the frequency transformation is applied to the

original signal, and it is better to hide the watermark in the low frequency [10], since the compress scaling affects the high frequency coefficients, i.e. it is better to make use of the most important information of a signal in the embedding process. That is because any attack cannot remove the watermark without causing significant damage to the watermarked signal.

Generally working in the Frequency Domain offers the robustness and imperceptibility watermarking requires unlike Spatial Domain, but the Frequency Domain has a high computation cost while Spatial Domain has less computation cost. The Frequency Domain Watermarking Algorithm is preferred to be used in copyright protection.

The Frequency domain can be classified into three algorithms, DWT (Discrete Wavelet Transform), DCT (Discrete Cosine Transform) and FFT (Fast Fourier Transform), with each having its special characteristics.

DCT is a mathematical function which is based on JEPG compression that is used to transform data into a summation of cosine waves of various frequencies, and it is usually used in audio or image compression.

FFT is also used to transform a signal into a frequency domain and it's similar to DCT, but it differs in using both the sine and cosine function, also it uses complex numbers while the DCT uses the real numbers.

DCT is very useful in audio lossy compression, but it is not good enough when compared with DWT. Which is very popular in frequency transformation because of its good features in allowing good localization. Furthermore , it easily divides the input signal into blocks according to their frequency, unlike DCT which that needs to divide the input signal into 8x8 block size, in order to do the transformation on the resulting blocks.

DWT is a better identification of relevant-human-perception-data, and also DWT transforms the whole signal that introduces inherent scaling. All of these advantages of DWT enable the watermarking algorithm related with DWT to be more robust against attacks than the other frequency algorithms, and as a result it is the more robust algorithm in the time and frequency domain.

Depending on the good abilities of DWT our work will be in this field to improve the watermarking algorithm against attacks and to provide watermarking requirements.

In our review for watermarking algorithms in the Frequency Domain, we discuss the work of all of the three frequency algorithms and also one of the discussed

algorithms combines between two algorithms of the DCT and DWT method, we discuss the advantages and disadvantages of each work as follows:

2.2.6.2.1. DWT (Discrete Wavelet Transform)

1. Al-Haj, Ali, Christina Twal, and A. Mohammad [19], have worked on the branch of DWT Frequency Domain and also applied the SVD (singular value decomposition) in their attempt to embed a watermark into an audio. In the first step they use a binary image as a watermark; next they samples the audio at a sample rate of 44100 samples per second, applying DWT discrete wavelet transform and SVD to the host audio. They have achieved a 4-level DWT transformation, re_sampling the watermark into a one-dimensional vector W (watermark), then embedding a watermark to the converted spectrum of the original audio according to a specific formula. Finally, after embedding, they reconstructed the final one, then extracted the watermark in reverse approach to the embedding procedure. They (the research team) discuss the algorithm that they have applied to the results of that algorithm. They found that for **"Fidelity"** they obtained an SNR value (signal to noise ratio) equal to 28.55, which is a good result. For **"Imperceptibility (Inaudibility)"** they performed a method which is called **"Perceptual Evaluation of Audio Quality",** which is an evaluation of the real world listening. In this method they classify five grades. They gave a value of 1 to the annoying audio and a value of 5 to the imperceptible audio. They obtained the approximation grade for their work equal to 5.0. (That means it was a very imperceptible audio). Finally, for the **"Robustness"** they have applied some **Attacks** (Add/Remove Attacks, Filter Attacks, Modification Attacks, MP3 Compression, ADOBE® Attacks) on the watermarked audio, and after extracting the watermarks from the watermarked audio after applying those attacks. They discovered that most of them were distinguished to some extent. If we want to evaluate this work we could notice that the team used two powerful transforms (DWT transform and SVD) and benefit from combining these two transforms over the time domain, which is a very crisp algorithm. Also the results that they gained from their work were good, but still they were able to apply this work to more than one audio to have more reliable results [2].

2. The writers of this paper [20], Tianchi, Liu, Yang Guangming, and Wang Qi, used a new algorithm depending on DWT (Discrete Wavelet Transform) and LSB (The Least Significant Bit). Also, we shall notice that the worker-group on this paper interested in using a multiple digital watermark with different characteristics to be embedded in the host digital audio said that the use of a multiple watermark will increase the security of the watermark.

They used two watermarks in their algorithm the chromatic image Lena (a) of $256bit \times 192bit \times 8bit$ to produce a robust watermark and grayscale image (**Copyright (b)** $128bit \times 128bit$) to get the fragile watermark.

The embedding algorithm was grouped into two major procedures; each one has many steps and each procedure was applied upon the two watermarks separately.

The sequential steps of the embedding algorithm of robust watermark (Lena) are as follows: a) Scrambling the data, they applied the Fibonacci function to scramble the whole image. b) Reducing dimensions $M\ x\ N$ bit of the chromatic image Lena converted to $M\ x\ N\ x\ 3$ bit. c) Discompose the audio frame. d) DWT transformation: they applied DWT transform on the audio frames, then choosing the most maximum value from the low frequency coefficients to embed the watermark in. e) IDWT transformation.

On the other hand the embedding algorithm of the fragile gray scale image (Copyright) was achieved after embedding the chromatic image (Lena) and the embedding of the fragile image was at LSB (The Least Significant Bit) as follows: a) Scrambling the data of chromatic image Lena. b) Reducing the dimensions for the watermark. c) Hex conversion.

d) Discompose the digital audio. e) Select the least significant bit to embed in.

After the embedding process, the extraction algorithm was applied to the watermarked audio in reverse steps, the team-work extracted the fragile image before the robust one, and the algorithm briefly is:

a. Extract the least significant bit:
   a.1. Translate the bit to the pixel value.
   a.2. Increase Dimension and Inverse scramble.
b. The extraction of robustness watermark:
   b.1. Perform DWT transform on watermarked audio.
   b.2. Extract the watermark.

b.3. Increase Dimension and Inverse scramble.

We try to compare the results of this algorithm with the watermarking requirements. First of all, we have the Imperceptibility, the team-worker enforced the SR (which is a familiar phrase in watermarking techniques used for finding the similarity between the host multimedia with the distorted one). It shows a convincing result for one audio, if this result compared with some other algorithms done in their paper.

Secondly, the Vulnerability (sensitiveness):- the writers found that any tampering for the watermarked audio will surely affect the watermark.

Finally, the Robustness (resistance):- the watermarked audio passed a number of attacks successfully, such as linear filtering and lossy compression.

Our point of view of the complete work that it has advantages and disadvantages, it is a good idea to use different watermarks with different features, but it will experienced challenges, for instance the two watermarks will have much information to hide in the audio that will may be affect the imperceptibility especially if we will take another audio, also the robustness undergo only fewer attacks, we expect that they try some other attacks [20].

3. Al-Yaman et al. in this proposed paper [21], which has the subtitle of ″Audio-Watermarking Based Ownership Verification System Using Enhanced DWT-SVD Technique″, from the subtitle we derive that the writers depend on the DWT and SVD in their algorithm.

The embedding stage begins with using an image as a watermark. They sampled the host audio, framing each one of these samples, then applied DWT transform of 4-Level and using the SVD for decomposing, SVD $(n\,x\,n)$ matrix A is equal to $U \times S \times VT$.

The digital watermark is encrypted by performing the SHA-1 hash algorithm, after this operation the encryption of the watermark is integrated to the DWT-SVD decompose by using the following formula: $S_{1,1w=}\, S_{1,1}\, [1 + \alpha \times W_n]$, where the W (n) is the hash bit of watermarked image containing 2 values either 0 or 1, $S_{1,1w}$ is the top-left of the S matrix.

The extraction stage has similarity with the embedding stage; they are framing the watermarked image, performing DWT frequency transform, matrix formation, then the SVD decomposition, and finally the extracted bits are obtained.

Ownership Verification Process is done by comparing the extracted bits from previous approaches to those gained from applying hash-1 on the watermark.

The enhancements that have been suggested by the writers of this paper can be done by the following steps:

a) Framing of Audio Signal: the team-worker avoided the watermark to be added to all frames that were obtained from the original audio. They try to embed the digital watermark to randomly- selected frames in order to decrease the noise.

b) Matrix formation: a five-multiple DWT sub-band applied to each frame to obtain the matrix formation in the following format:

$D_1 = [A_4, D_4, D_3, D_2]$

Where the D_1, D_4, D_3, D_2 are the fifth DWT decomposition of the high frequency and A_4 is the fifth level of low frequency DWT decomposition

$Matrix\ Size = 2(rows) \times \frac{1}{2} L(columns)$, where the L is the frame-length.

According to the five level DWT sub-band to the D in the foregoing speech, the team avoided the repeated D as in other works, so they will not stick in the SVD reverse problems.

c. Embedding Process: try to use different values of α which is the watermark intensity (if it is high value it will be impervious to attacks and in the minimum value it will be more Imperceptible but easy to damage), in order to find the most suitable α value for embedding and to obtain the perfect, robustness, Imperceptible watermark.

A good amount of attacks has been achieved in this paper by the team and they try to evaluate their work by finding the **BER** value which is the meaning of (Bit Error Rate), in other words the ratio of extracted errors to the total bits.

In the final review of this paper we found it a good algorithm in DWT frequency transform because of the Framing of Audio Signal, Matrix formation and Embedding Process, in each one they explained how they enhanced the algorithm; in summary the Framing of Audio Signal was attempted to be embedded in random frames which increase the Imperceptibility, reduce the cover size and obtain the inaudibility. In the Matrix formation the team avoided problems in SVD reverse, and finally in Embedding Process they tried to find the optimum value to embed watermark with

less BER, and they applied a reasonable number of attacks and they had obtained results in most of them [21].

4. A.R.Elshazly, M.M.Fouad and M.E.Nasr, the researcher group of this paper [22], were interested in the work with the frequency module DWT transform like the former researcher in [19], [20] and [21].

The group worked on the binary image as a watermark to be embedded by the team in the original audio. After they segmented the host audio into frames, they applied three levels of DWT decomposition on each frame. They embedded the watermark after they were encrypted. It depended on **Logistic maps** to the low frequency coefficients of DWT decomposition, and a watermark with chaos encryption and less complexity was added to the high frequency of 3 levels of DWT decomposition.

The encryption approach of watermark used chaotic iteration to gain the secret key, after that XOR achieved with the plain text in order to alter the image pixels values, the **Logistic maps obtained** using the formula of: $Xn + 1 = aXn\,(1\_Xn)$.

Embedding algorithm:

a. Using two dimensional binary image as a watermark of size=MxN.

b. Encrypt binary digital image using chaos relay on a secret key.

c. Reshape the watermark from 2 dimensions to one dimension to be suitable for embedding in the audio signal.

d. Performing 3 levels of DWT decomposition on the host audio to obtain low frequency and high frequency coefficients on each frame of the audio signal.

e. Choosing low frequency coefficients to embed watermark in, and they are ordered in matrix of FxL, F is the size of each frame, and then apply mean quantization.

f. The result of former steps was reconstructed in the following order: Á= $\{CA_3, CD_3, CD_2, CD_1\}$.

Extraction algorithm:

The extraction algorithm of this paper was performed without the needs to the host audio, which means a blind detection for a watermark.

The extraction method is similar to the embedding process with differences in decrypts of the watermark and the mean value.

a. The 3 levels of DWT decomposition achieved on the watermarked audio to get the low frequency($C'A_3$), ordered in matrix of size FxL.
b. The encrypted recovered watermark was obtained using a special formula.
c. Reshape the encrypted watermark was obtained from previous work from 1-dimension to 2-dimension.
d. Using a secret key to get the watermark from encrypted watermark.

A.R.Elshazly et al. try to evaluate their work through discussing some points. They use as they mentioned an encrypted binary digital watermark and hide it in the original audio without perceptual difference between the host and watermarked audio. For subjective substantiation they applied the BER (Bit Error Rate), SR (Signal-to-Noise Ratio), PSNR Peak Signal_ to_ Noise Ratio, and NCC Normalized Cross-Correlation on the watermarked audio. All of these functions are used to find if there are any perceptual errors between the original audio and the watermarked one.

If it is our place to evaluate the team work, we see that it is an effective idea to use the DWT transform as well as the usage of encrypted watermark and to hide it in low frequency coefficients because it is the region with the most energy, but using the high frequency sub-band will affect the inaudibility because it is a very sensitive area to work with, especially in the audio. May be the high frequency in an image or video is reasonable, but in the audio we don't think so.

It is a good work to achieve their experiment on a fine number of clips to evaluate their work on different audios with diverse features to support their algorithm. Also, they were getting benefit from performing a reasonable number of functions to discover the amount of errors between the original and watermarked audio after attacks. They also used a fair number of attacks to prove the power of their algorithm.

5. Shaoquan Wu et al. in the present paper ″Efficiently Self-Synchronized Audio Watermarking for Assured Audio Data Transmission″ [23] propose a self_synchronization method for watermarking audio signal, actually they were interested in DWT Frequency Domain Transform and try to embed the synchronization codes as well as the formative data (watermark) in low frequency sub-band of the host signal, the watermark is denoted as a sequence of ECC (Error Correct Code).

Shaoquan Wu et al. benefit from the good localization of DWT in eliminating the load to find the synchronization codes. Thus the purpose of the algorithm makes some balancing between low complexity and robustness, also they used SNR (signal to noise ratio) and BER (bit error rate) to certify the performance of their work.

They tested their algorithm on two Wave audio formats of length 15 Sec. that has different characteristics, quantized at 16 bits, sampled at 44.1 kHz and they used a watermark of 256 bit sequence.

The propose of the algorithm shows good robustness towards common attacks like re-sampling, MP3 compression, Gaussian noise corruption, re-quantization, cropping.

6. Fallahpour Mehdi and David Megías in [24] proposed a watermarking technique that is based on frequency domain of the wavelet of the DWT.

The original audio signal is segmented into frames and the mean value of each frame is used as a key in the watermarking procedure.

The researchers of this paper decomposed the frame into second-level decomposition and take the high frequency sub-band of the second decomposition to embed the watermark in.

The proposed algorithm was evaluated using the following functions: $SNR = 30$, which mean good result for imperceptibility, ODG (Objective Difference Grade) in the range $[-1, 0]$, which means inaudibility.

The experimental results have an excellent capacity that equals about 11 kbps. This work shows good resistance to attacks like MPEG compression (MP3), additive noise and echo.

One thing to mention in this work is that the high filter may erase the watermark.

Pitch Shift and Time Stretch attacks affected this work, but they also damaged the audio signal.

7. Peng Hong et al. in [25] proposed a novel watermarking algorithm based on the Kernel fuzzy c_means (KFCM) that control two quantization steps, the mean and energy quantization. The original audio is segmented into frames, and each frame is partitioned into two sub_frames. By using the mean quantization, a synchronization code is embedded in the first frame sub-part and the second sub_part of the same frame was embedded by using the DWT with the desired watermark.

The proposed algorithm in watermark detection tries to extract the synchronization code from the first part of each frame as the first step. By finding the position of this code, the watermark is extracted from the low frequency of DWT of the next sub-part of the frame.

The work was tested on four different audio signals. It shows robustness against ten attacks, 5 de_synchronization attacks (pitch shifting, amplitude variation, random cropping, jittering and time-scale modification) and 5 common signal processing attacks (re_sampling, re_quantizing, low_pass filtering, additive noise and MP3 lossy compression).

KFCM the machine learning technique that is used by this paper, is good to control the quantization steps and also to control the strength of the watermark because of finding the optimum parameters and best robust location to embed the watermark according to the features extracted for each frame of the original signal ; this will provide low computational complexity, unlike other papers that use other learning machine methods like ANN (artificial neural network) and SVM (support vector machine) that need long training time and a complicated training algorithm.

This experiment is also tested using PSNR and BER functions to show the improvement of the present work. Most of the values of these functions were better than the other proposed work in the machine learning method.

Besides all the advantages mentioned in the former paragraphs, this work has robustness for two attack-groups (common signal processing and de_synchronization attacks), which is very difficult to achieve both of them, this work keeps balance between robustness and imperceptibility which is very difficult to achieve, in brief words it is very fabulous work.

8. Chen S-T., G-D. Wu and H-N. Huang in [26], optimization_based group_amplitude method with DWT were performed for this paper.

In order to increase and to enhance the robustness against crop and shifting, a synchronization code with the watermark were hidden in the lowest frequency of the 7_level decomposition of DWT.

The parameter of the group-amplitude quantization that was used in this paper is equal to 13 500.

The extraction method tries to find the synchronization code in order to find the watermark according to that synchronization code.

In this work, optimization based on the quantization method was used to keep the balance between the SNR and BER which were used to measure the robustness and imperceptibility of the watermarked audio.

This method was tested on two audio clips, the SNR for embedding quality was performed on these two clips and they were compared with the results of paper [27] and [28].

The proposed work tested against the five attacks: 1) MP3 compression 2) Low Pass Filtering 4) Re-sampling 5) Amplitude Scaling.

This work shows good resistance to MP3 compression for two clips when compared with [27] and [28] papers, also it shows good resistance against the re_sampling attack but for the others it was not as robust against the remaining attacks.

This work was very good in providing new ideas to balance between BER and SNR, and embedding in DWT is very effective work, but it still has some weak points. This proposed algorithm should be developed to be more robust against different attacks.

2.2.6.2.2. DCT & DWT

1. Xiang-Yang Wang and Hong Zhao in their proposed algorithm [29], produce an algorithm that combines between two Frequency Domain algorithms (DWT and DCT).

They get benefit from using the two frequency algorithms; from DWT they make use of the multi-resolution characteristics and the energy-compression characteristics of DCT in order to amend the transparency of the watermark and the main aim is to find a violent resistance towards common attacks, especially synchronization attacks like cropping. The16-bit Barker code as synchronization mark 1111100110101110 and

64x64 binary image was hidden as a watermark in the low frequency amplitude depending on human auditory masking. Also in extraction they used a blind algorithm with no need for the original audio.

The proposed algorithm shows robustness against attacks such as: MPEG_1 Layer III (MP3) compression, noise adding, re_quantization, re_sampling and random cropping.

It was a good idea to use two algorithms to offer more robustness; one of the disadvantages of this work is the weak resistance against pitch invariant time scale modification.

2. Dai Hua-liang and Di He in [30] proposed a zero-watermarking method that depends on the steady sign of certain DWT-DCT coefficients with maximal absolute value. The major features were picked out from the original audio. The XOR operation was achieved between the extracted features and the host watermark to have a key that is used in the detection for the watermark.

Actually the host audio is segmented into a number of segments and the higher energy segments were selected. After that the DWT was applied to the selected segments to have the coarse signal which was partitioned into frames. The DCT was performed on these frames to get the DWT-DCT coefficient that earned the maximal absolute value.

The proposed work increases the robustness of the watermark against attacks by using the DWT and DCT to get high energy segments. Another advantage of this work was the key which is used to detect the watermark that will avoid mistakes of false detection.

This algorithm tested using NC, BER, SR functions and also was compared with [29] and [31]; in most of its results it shows higher robustness and imperceptibility that was achieved naturally by the proposed work unlike in [29] and [31].

Nonetheless, the preset algorithm has the high computational complexity that resulted from using two frequency algorithms DWT and DCT.

3. Ren Keqiang et al. in [32] was interested in the work that was based on two important frequency algorithms: the DCT and DWT.

The Arnold transforms and DCT (Discrete Cosine Transform) was used by the proposed algorithm to perform scrambling encryption and compression on a

watermark, then the 3 levels DWT applied to the segmented audio, the watermark was embedded in the low frequency domain of the original audio.

This algorithm was also based on the large capacity of the watermark, which is a 24 bit true color image of 40×40 pixels.

The scrambling used in this work tries to reduce the correlation among pixels in the watermark, which has an important role in increasing the security and the invisibility of the watermark, also the Arnold algorithm performed the recovery of scrambling images. This algorithm was applied on the watermark 10 times to get the scrambled watermark with an Arnold cycle of 30, in order to extract the watermark Arnold transform must be done 20 times. As we see this will increase the security of the watermarking.

Because of the high capacity of the color image the DCT was used to compress the watermark that has the good characteristics of energy concentration and de-correlation.

The experiment was tested using the NC and PSNR functions; it shows robustness towards the following attacks depending on the PSNR and NC values:

Up-sampling, down-sampling, low-pass filtering, median filtering, white noise, colored noise, de-noising and compressing.

This work is good in imperceptibility, robustness of the mentioned attacks and in the enhancement of the large capacity of watermark, but it will not be robust enough against the de_synchronization attacks like a crop or shifting, and also there is the high computational cost.

2.2.6.2.3. FFT (Fast Fourier Transform)

1. Dhar Pranab Kumar and Isao Echizen in their proposed work [33], perform the copyright protection by applying the FFT (Fast Fourier Transform). They applied the FFT on the original audio in the first step; then they partitioned into segments. The higher energy segments were selected, after that the watermark was embedded in these picked out segments. The watermark detection was done in reverse steps of the embedding procedure.

This work for imperceptibility exceeds Cox's method, and if it is compared with Cox's method for robustness, the SNR function was performed and the range was

between 20-31 dB for this paper's algorithm, unlike the Cox's method that SNR have the range of 11-23 dB, which means that the proposed algorithm shows more robustness for most of the tested attacks which include the following: Noise addition, Re_sampling, Cropping, Re_quantization and MP3 compression.

It is a good work if we look at SNR values. To improve this work, the team worker should increase the amount of attacks to overcome the ownership problem.

2. Kang Xiangui, Rui Yang and Jiwu Huang in [34], a multi-bit spread-spectrum audio signal watermarking was proposed by this paper depending on the geometric invariant log coordinate mapping (LCM) feature.

The watermark was embedded in LCM features; in fact the watermark was embedded in DFT frequency domain with no need to the interpolation, so this will reduce the degradation that generated from non-uniform interpolation mapping.

Effectively the synchronization of watermark is done using one FFT and IFFT, also the mixed correlation between a key-generated PN and LCM features was used to stratify the log-coordinate mapping.

This work offers subjective and objective high auditory quality; the objective quality was evaluated using the SNR which is greater than 33dB, which is a very good result, according to IFPI and ODG (objective difference grade) was achieved and it was equal to 0.1.3 alpha which is greater than -1. That is means that the watermarked audio is similar to the original signal. On the other hand the subjective field was performed by asking 11 persons to distinguish between 2 audio signals. One of them with watermark and the other without. The discrimination rate was equal to 52%. That means that both audio could not be discriminated.

If we compare the present work with others, we would obviously notice that other papers used the ILMP interpolation on the watermark in the embedding scheme that could add a distortion to the watermark through the mapping procedure, unlike this paperwork that used the LCM with no interpolation in embedding. In fact the embedding will be in FFT transform domain that will not cause the watermark degradation.

This paper shows good resistance against common signal processing operations like Low Pass Filtering, echo addition, MP3 decompression, normalization and volume

change. Also, it has good resistance towards the Stirmark benchmark and DA/AD conversion.

This work also shows effective robustness to the geometric distortion attacks like TSM (time-scale modification) of $\pm 20\%$, pitch shift of $\pm 20\%$, random cropping of 95% and re_sample with scaling factor $75 - 140\%$.

In the extraction process, the BER was performed on the extracted watermark for the suggested attacks, and it was equal to 0-1.5, which was very efficient in robustness against attacks. It is really good work for all the advantage points that were mentioned in the former paragraphs.

2.2.6.2.4. DCT (Discrete Cosine Transform)

1. Liu Ji-Xin, Zhe-Ming Lu, and Jeng-Shyang Pan in the proposed algorithm of [35], protect the watermarking with an algorithm based on the DCT transform and Vector Quantization (VQ) method which is an efficient lossy compression technique for multimedia that uses a high compression ratio.

In this suggested work, the audio is segmented into frames, and the DCT algorithm is applied on these frames. Then the middle_frequency coefficients are selected that formation of a vector in order to be modified. This middle_frequency is used to generate a codeword_labeled VQ codebook depending on LBG codebook design method.

The labeled_codeword is used to quantize the obtained vector from the middle-frequency depended on the Vector Quantization (VQ), this is done according to the watermark bit, then the reconstruction method of the inverse of DCT is applied to the selected coefficients with the unwanted ones to obtain the watermarked audio.

The extraction method is applied without the need of the host audio i.e. blind extraction method; the extraction method here depends only on the Vector Quantization (VQ) method.

One of the important things is that the VQ method was used for a variety of papers that applied it to the video and image watermarking, but no one used this kind of method in audio watermarking which is used in this paper.

The SNR (signal_ to_ noise ratio) and NC (Normalized Correlation) methods are used to evaluate the proposed work against the suggested attacks of Stirmark Benchmark.

Through the values of SNR and NC, this work is robust against AddBrumm_100, AddBrumm_1100, AddBrumm_9100, AddSinus, Compressor, DynNoise, FFT_RealReverse, FlippSample, LSBZero, Normalize, ExtraStereo_30, ExtraStereo_70, RC_HighPass, RC_LowPass, Smooth and ZeroCross, and it shows less robustness against ZeroLength, Echo, Invert, FFT_Stat1, FFT_Test, Exchange, FFT_HLPass, FFT_Invert, CopySample, CutSample, Stat1, Stat2, ZeroRemove and Amplify.

This method is efficient against some attacks, but it is not for many others and the team applied this method to just one audio. If they try to apply it to more than one, the working in DCT is more complex than the work in DWT that can offer more reliable results in attacks robustness.

2. Yongqi Wang and Yang Yang proposed an algorithm in [36] that is based on the chaotic encryption that depends on DCT transform algorithm.

Synchronic signal is used to find the watermark and then both of them (synchronic signal and watermark) are embedded in the original audio of the low frequency by using the quantization method.

A watermark serial number which is the synchronization code (Bark code) and the watermark is embedded in audio to enhance the watermark robustness against attacks, especially the synchronization attacks.

The watermark extraction does not need the original audio signal; the complete process is done by extracting synchronic code, extracting the serial number and the watermark.

The proposed algorithm improves the transparency between the original and watermarked audio.

The present work tested against a number of attacks like adding Gaussian White Noise, re_sample, low pass filter, re_quantization, random cutting, the NC and BER evaluate the robustness of a watermark is used in this work and it shows robustness for all suggested attacks except re_sampling.

By having a look at all the work, we have got some of the disadvantages of this algorithm in that it is tested on only one audio, besides the huge quantity computation that is done by using the DCT frequency method.

3. Guo Qijun et al. in [37] submit an algorithm that is based on the second level decomposition of DCT and select the low frequency sub-band of the second level to embed a watermark in the host audio.

They perform the BCH coding on the watermark and the resulted encoded information of the watermark with synchronization signal is embedded in the audio carrier with large capacity of watermarking information.

The proposed work is designed to be robust against attacks, especially D/A and A/D transform which were affected during the cable channel transmission with the cable channel noise. This work tries to solve the former difficult mentioned attack by choosing to work on the frequency domain which is more robust than the time domain.

The proposed experiment shows robustness against the common signal processing attack like, Re-quantization, Normalize and Low Pass Filter. We can notice that the usage of BCH error correction coding will reduce the BER (bit rate error), which is a benefit for the suggested work, this work is also tested against re_sample attack and it shows robustness for approximately 30% of re_sample attack by using BCH code on the watermark.

For recorded audio, the StirMark Benchmark attacks for Audio were tested and it shows good results, also this algorithm is tested by using a different sound card to determine how this change in sound card will influence the watermarked audio, and it proves that there is no noticeable difference among them.

By calculating the BER of the influence of a different player, we can see that there is no big difference between the BER result of different players.

The strength of the embedded synchronous signal could control the synchronous code algorithm for a variety of broadcast delay, which makes the watermark not easy to be discovered by others, so it will make the watermark more robust.

This algorithm holds BER below 6.2%, which effectively can be performed on the field of cable channel transmission.

Finally the suggested work is good for many advantages that is mentioned in the former paragraph, but to improve this work it may be tested with the works of others

to see how much the proposed work is better than other works in the same aspect, and also the work on DCT will make the work more difficult in time computational complexity.

4. Xiong-Hua Huang, Jiang Wei-Zhen and Jing Xing-Xing in the present paper of [38] propose an algorithm that based on the non-uniform DCT, the watermark that is used to be embedded in the host audio, is obtained by the Henon chaos system with the key.

The generated watermark is embedded into the quantified statistical coefficients mean of the NDCT non-uniform DCT.

Chaos generates the NDCT frequency sampling positions that increase the watermark robustness.

The proposed algorithm uses 5 keys in the operation of watermark embedding, and the length of the key is limited depending on the effect of the computer word length, so this will increase the security of this work- algorithm.

The SNR in the IFPI (International Federation of the Phonographic Industry) should be greater than 20 dB and in this work the SNR values for the three tested audio signal were 36.07dB, 36.75dB, 29.34dB which is good results.

The proposed algorithm is tested against MP3 compression, filtering, Gaussian-noise, re-quantization and re-sampling attacks using BER function and in all of them in this work shows perfect result.

The proposed work is good for the advantages that are mentioned in the previous paragraphs, but to increase the power of this algorithm, it should put the synchronization attacks in considerations, which are the main attacks against audio signals and also to try to work in other frequency algorithm which is much easier in work like DWT.

CHAPTER 3

PROPOSED METHOD

3.1. Data Source

We apply our watermarking algorithm on two different features audio clips in wav format, Pop and Classic. The first one has a high tempo while the second one has low tempo. Each of the two audio files was sampled at 44.1 and 48 kHz for Pop and Classic audio clips respectively, and quantized to 16 bits per sample. We will use a blind technique in our watermarking method which means that we need for the original signal in our algorithm. We also used a binary image watermark of 265×256 pixels. The original audio frames, the watermarked frames for the both audio clips and the original binary image will be illustrated in Fig (6), Fig (7) and Fig (8) as follows:

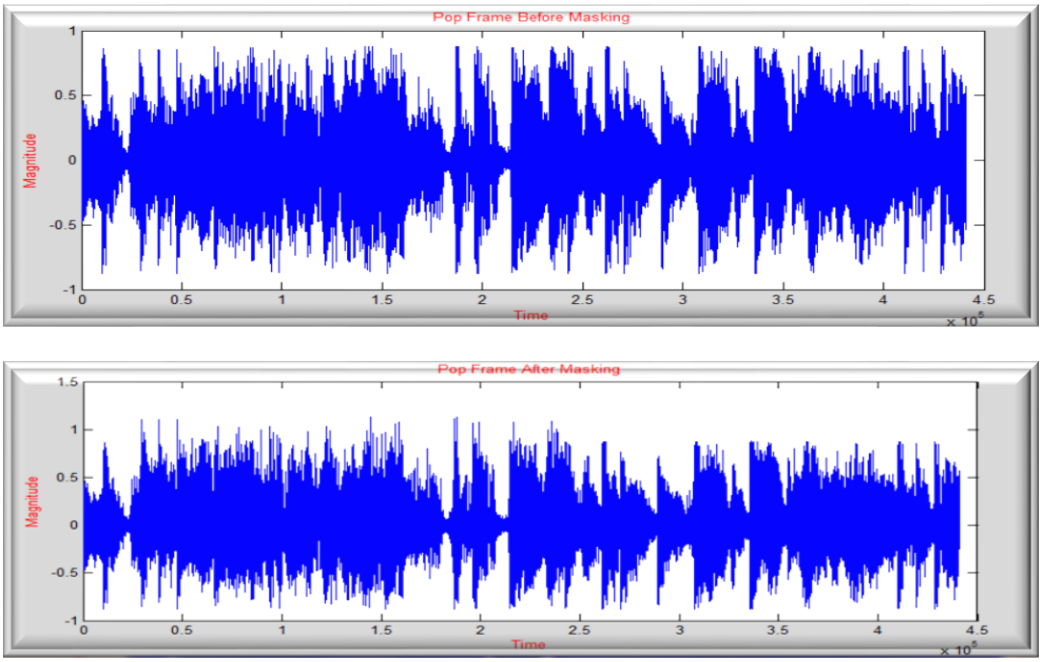

Figure 6: Pop original frame and Pop watermarked frame

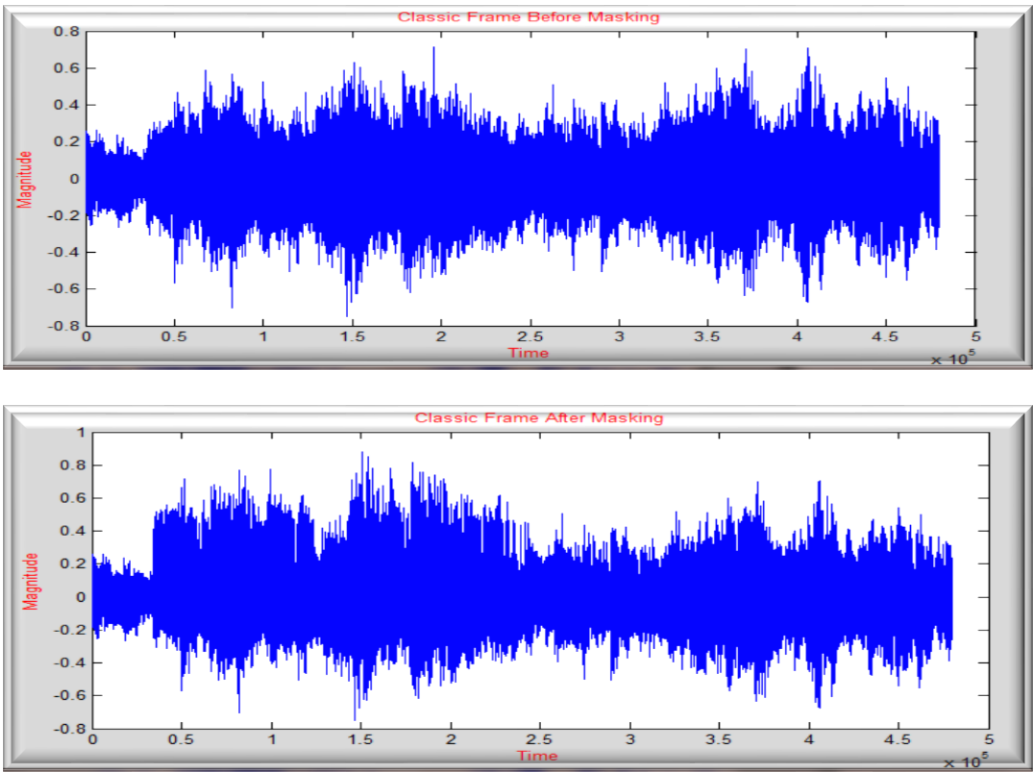

Figure 7: Classic original frame and Classic watermarked frame

Figure 8: Binary image watermark $\mathbf{256 \times 256}$

3.2. The Proposed Frequency Audio Watermarking Technique (DWT)

Audio watermark is embedded either in the time or frequency domain; both techniques have different characteristics, as we mentioned in chapter 2. Frequency domain watermarking algorithm try to offer a good versatility to control and balancing between the requirements of audio watermarking (robustness and inaudibility) [39].

The frequency domain audio watermarking techniques take the advantages of Human Auditory System (HAS) to preserve the inaudibility, especially when working with the wavelet transform DWT.

The groups that perform watermarking and select the DWT technique, tend to exploit the great performance of DWT that offers a multi - resolution, simultaneous spatial localization and also spread spectrum of an original signal.

Our work focuses on DWT algorithm, so our work belongs to the groups that interested in DWT as we mentioned in 2.2.6.2.1. DWT (Discrete Wavelet Transform) in chapter 2.

The DWT divides the original signal (as an image signal) in the first level decomposition into four sub-bands these sub-bands are the Low Frequency sub-band (LL), the Mid Frequency sub-bands (HL, LH) and High Frequency sub-band, the larger the magnitude DWT coefficients can be provided by the LL, the following figure will illustrate the four sub-bands of DWT, Fig (9).

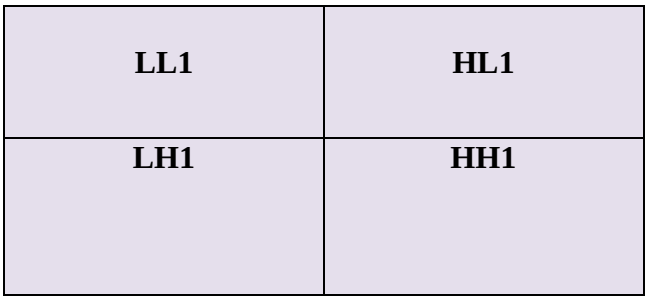

Figure 9: DWT-first level decomposition of an image

Mathematically One-Dimensional DWT is illustrated as follows:

$$H(w) = \sum_{k} h_k \times e^{-jkw} \quad (3.1)$$

$$G(w) = \sum_{k} g_k \times e^{-jkw} \quad (3.2)$$

The $H(w)$ and $G(w)$ represent the low and high pass filters and they should satisfy the following equation:

$$H(w) + G(w) = 1 \tag{3.3}$$

We shall use the DWT of Haar Filter; The Haar Filter in DWT can be represented using the following formula:

$$H(w) = \frac{1}{2} + \frac{e^{-jw}}{2} \qquad , G(w) = \frac{1}{2} + e^{-jw} \tag{3.4}$$

By using the following two formulas, $F(n)$ of one dimensional signal is represented:

$$f_{j-1}^{low}(k) = \sum_{n} h_{n-2k}\, f_j(n) \quad , \quad f_{j-1}^{high}(k) = \sum_{g} h_{n-2k}\, f_j(n) \tag{3.5}$$

The construction formula of $F(n)$ can be represented as follows:

$$f_j^{low}(n) = \sum_{k} h_{n-2k}\, f_{j-1}^{low}(k) + \sum_{k} g_{n-2k}\, f_{j-1}^{high}(k) \tag{3.6}$$

For an audio, the first level decomposition of DWT is performed in order to obtain the approximation coefficients A1 (low frequency coefficients) and the detail coefficients D1 (high frequency coefficients), and for our work the low frequency coefficients A1 is decomposed again in second level decomposition to have as a result A2 and D2. Fig. (10) illustrates the second level decomposition for an audio signal.

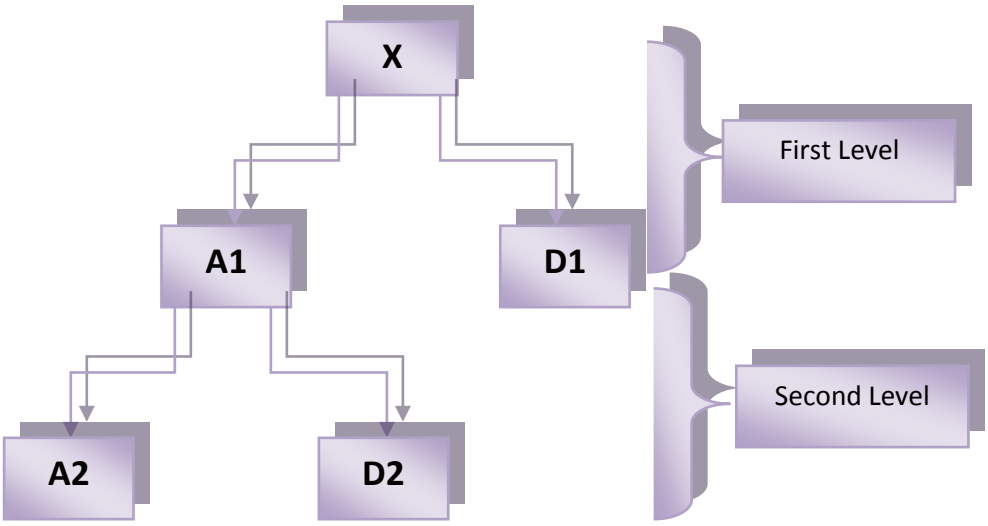

Figure 10: DWT second level decomposition of signal X

3.3. Imperceptibility Measurement Terminology (Used in Embedding System)

We have used in our thesis for inaudibility (imperceptibility) evaluation, two different methods, the first one which is denoted by an objective evaluation, the PSNR function, while the second one, a subjective evaluation that related to Human Auditory System (HAS).

3.3.1. PSNR (Peak Signal -To- Noise Ratio)

An engineering idiom that attempts to measure the ratio between the maximum-possible-signal-power and the magnitude of distorted signal, it is used widely in quality measurement of lossy compression, PSNR can be defined in an easy manner via MSE (Mean Square Error) which can define through the following formula:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2 \qquad (3.7)$$

Where (I) is the original signal and (K) is the corrupted signal.
PSNR formula can be defined as follows:

$$PSNR = 20.\log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right) \qquad (3.8)$$

It is difficult to approximate human evaluation, the Objective evaluation model, including PSNR method trying to achieve this evaluation.

According to **IFPI** (**International Federation of the Phonographic Industry**), PSNR evaluation must preserve over than 20 dB to provide a great inaudibility between the original and distorted signal, we shall use the PSNR method in an embedding system to evaluate the perceptibility, and the results of this method will be illustrated in Table (2) in chapter (4).

3.3.2. Subjective Quality Evaluation

The Subjective Quality Evaluation Model is performed based on ten people's observations, i.e. this model is based on Human Auditory System (HAS) to detect if there is an audible noise between the host and the watermarked audio clips. We will include a table in chapter (4) that represent the proposed evaluation. In this table, we have a mean opinion score (MOS) which gives a reasonable evaluation for a perceptual distortion in watermarked signal; we gave the 5 points score for inaudibility, the 4 points for perceptible but not annoying, the 3 points score is given to partially annoying, the 2 points score for annoying and finally the 1 point score is given to a corrupted signal. We illustrate the result of this model in Table (3) in chapter (4). This method will be used in an embedding system technique of our proposed algorithm.

3.4. Robustness Measurement Terminology (Used in Extraction System)

In our proposed algorithm of watermarking extraction, we shall use two methods to evaluate the watermark robustness against malicious attacks or signal modifications; we illustrate these two methods in the following paragraphs:

3.4.1. SR (Similarity Ratio)

This term denotes for the ratio of similarity between the original watermark and the extracted ones, we can define the SR formula as follows:

$$S = S + D \; / \; S \tag{3.9}$$

Where S is the matching pixels and D is the different pixels.

We shall use the SR model in our objective evaluation of the extraction system for an extracted watermark: this model will be explained in Table (1) in chapter (4).

3.4.2. Subjective Evaluation for Extracted Watermark

We shall use subjective difference grade (SDG) that gives an evaluation of watermark robustness against attacks. This will give an evidence for similarity between the host and extracted watermark depending on Human Visual System (HVS), for this we ask two people about their opinion in evaluating the robustness of an extracted binary images. We will give a point score of 5 to the imperceptible watermark, 4 points score is given to slightly distorted watermark, 3 points is given for partially distorted, 2 points to a watermark can hardly be seen, finally 1 point score is given to totally distorted ones, this evaluation quality will be explained in Table (4) of chapter (4).

3.5. Embedding Procedure

We have $A\{a(i), 0 \leq i < Length\}$ that represent the original audio signal to be watermarked, and let $W = \{w(i,j), 0 \leq i < M,\ 0 \leq j < N\}$ be the watermark binary image of 256×256 pixels, that is to be embedded in the host digital signal.
The embedding system is illustrated as follows:

Step1. The audio signal A at first is divided into frames, each frame has a time of 10 seconds, the frame length is more suitable to include the whole binary watermark in one frame and also it is ideal if it is compared with the audio clip length.
Frame-No. can be obtained by:

$$Frame\ Length = Frame\ Duration\ \times\ Fs \qquad (3.10)$$

Where Frame Duration=10 and Fs=44100 for Pop and Fs=48000 for Classic.

$$Frame\ No. =\ Audio\ Length/\ Frame\ Length. \qquad (3.11)$$

Step2. Since the audio signal is one dimension, we reshape a watermark W of $N \times M$ pixels into one dimensional signal to be educated to be embedded in the host audio.

Step3. We apply two level decomposition of DWT of Haar Filter on each frame of the signal, we will have $cA(i)$ which denotes the approximation coefficients of the first level of DWT, while $cD(i)$ represents the detail coefficients of the first level of DWT.

The second level decomposition is applied on $cA(i)$ to obtain the $cA2(i)$ the approximate coefficients of second level decomposition of DWT and the detail coefficients $cD2(i)$.

Step4. RMS (Root Mean Square) model is applied on the $cA2(i)$ for all the frames, the RMS formula is as follows:

$$RMS = \left(\frac{\sum_0^i cA2(i)^2}{LcA2}\right)^{1/2} \tag{3.12}$$

Where $LcA2$ is the length of $cA2$.

Step5. After we apply the RMS on all of the low level coefficients of second DWT decomposition, the RMS for all of $cA2(i)s$ will be arranged in descending order, then the number of frames to embed watermark in, can be calculated as follows:

$$SFN = TotalFrameNo./4 \tag{3.13}$$

Where SFN represents the selected number of frames.

We find SFN in order to obtain the final quarter number of frames from the whole audio frames to embed the watermark in. That is because we do not propose to embed a watermark in all of the resulted frames from the original audio segmentation.

By ordering the frames in descending order depending on RMS of the low frequency $A2(i)$, and taking into consideration SFN the Selected-Frame Numbers, the embedding system will select the lowest level magnitudes from these ordered $cA2(i)$ to embed watermark in.

Step6. The procedure after selecting the desired frames with a larger magnitude of $cA2(i)$ is to embed watermark in these $cA2(i)s$ according to the following formula, that excludes the magnitudes less than 0.3:

$$If\ \ cA2(i) > 0.3$$

$$cA2(i) = cA2(i) + \alpha * W(i) \tag{3.14}$$

Where α equal 0.5 and α controls the strength of the watermark against attacks.

Step7. In order to have the watermarked signal, we are reconstructing the resulted signal after embedding using IDWT for all the sub-band of the watermarked signal of the second level of DWT.

The embedding system of our proposed embedding procedure is illustrated in the following diagram of Fig.(11):

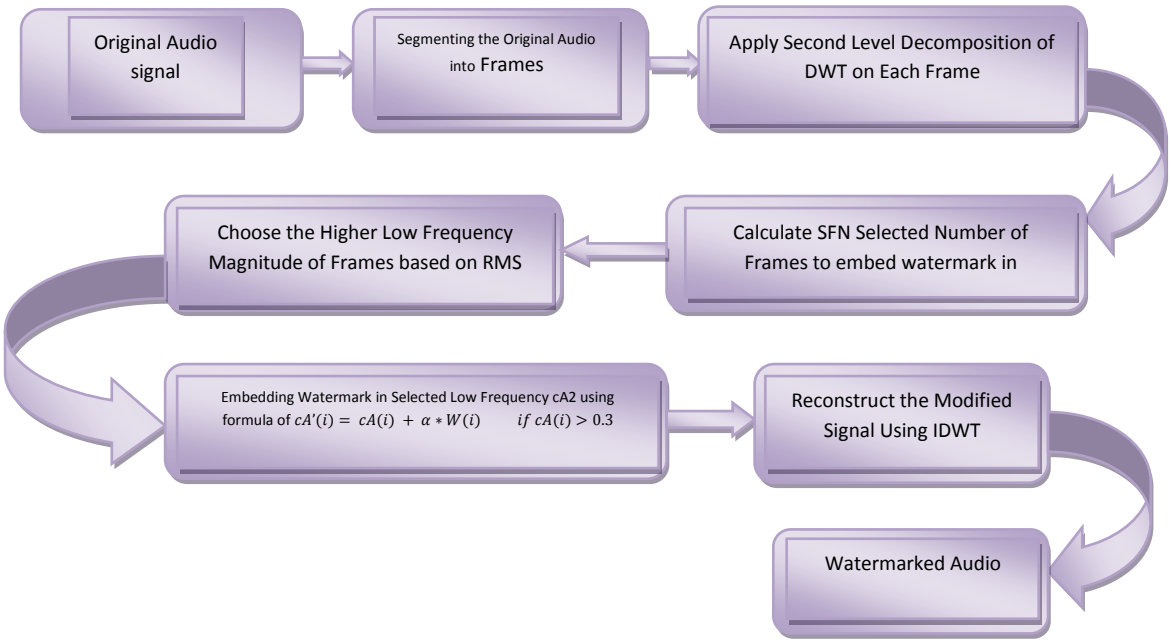

Figure 11: The proposed embedding system

3.6. Extraction Procedure

Let $A\{a(i), 0 \leq i < Length\}$ represents the original audio signal, and $A'\{a'(j), 0 \leq j < Length\}$ that represents the attacked and watermarked audio, the extraction procedure can be easily done according to the following steps.

Step1. Segmenting both of the original and watermarked audios into frames, each audio-frame-duration is 10 seconds and the frequency Fs=44100 Hertz for Pop and Fs=48000 for Classic, the segmentation formula is as follows:

$$Frame\ Length = Frame\ Duration\ \times\ Fs \tag{3.15}$$

$$Frame\ No. =\ Audio\ Length\ /\ Frame\ Length. \tag{3.16}$$

Step2. Apply second level decomposition DWT on all of the resulted frames for both audio clips, in order to obtain low frequency coefficients of $cA2(i)$s and $cA2'(i)$ of original and watermarked audio respectively.

Step3. Each of the resulted frames from segmentation for both audio clips is compared with each other to obtain the threshold parameter according to the following:

If $(cA2(i) >\ 0.5)$

$$ThresholdX = cA2(i) - \frac{cA2'(i)}{\alpha} \tag{3.17}$$

$\alpha = 0.7$

Step4. The watermark is resulted from the following:

If $(ThresholdX\ >\ 0)$

$$W\ (i) =\ ThresholdX \tag{3.18}$$

Reshape the resulted one-dimensional watermark into two dimensional watermark $W(i, j)$.

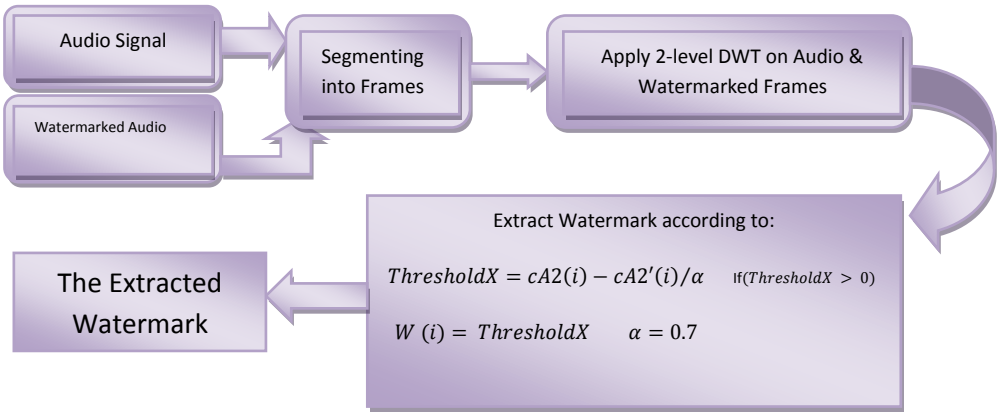

Figure 12: The proposed extraction system

CHAPTER 4

EXPERIMENTAL RESULTS

4.1. Data Source and Environment Techniques:

For data we have used two audio clips of WAV format and Mono channel. The first one is named Pop and the second one is named Classic,. These two audio clips that were used as tested data have different characteristics: Pop audio is a high tempo audio while the Classis is a low tempo audio, Pop and Classic are sampled at 44.1 kHz and 48 kHz respectively, both of the tested audios were quantized at 16 bits, Pop and Classic are about 3.36 and 2.86 minutes respectively, Pop is about 16.9 MB and Classic is about 15.5 MB. We segmented the two audio clips into frames. Each frame is 10 Sec., so Pop is segmented into 20 frames, and Classic is segmented into 16 frames. The number of the selected frames to embed watermark for Pop and Classic are 5 and 4 frames, respectively.

The selected watermark as illustrated in Fig. (8) in chapter (3) is about 256×256 pixels, also to achieve such a work we used the 2009 version of Matlab, the selected watermarking technique for embedding watermark is DWT (Discrete Wavelet Transform) of second level decomposition of low frequency, the watermark is embedded in the second_level_low_frequency sub_band. We apply 17 attacks on both watermarked audio clips (Pop and Classic), and finally propose a blind digital audio watermarking scheme to resist the 17 attacks to extract the watermark.

4.2. Extracted Watermark for Each Attack

After applying 17 attacks on two selected watermarked audio clips (Pop and Classic) , the extracted watermarks for each attack from both audios are illustrated as follows:

4.2.1. Standard Deviation Attack

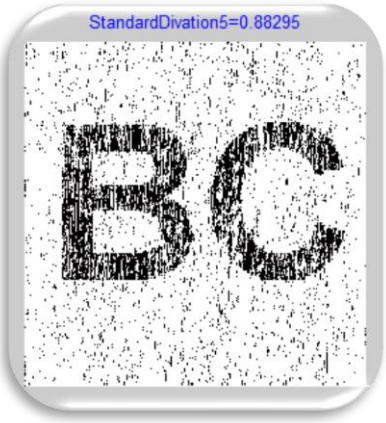

Figure 13: Pop standard deviation

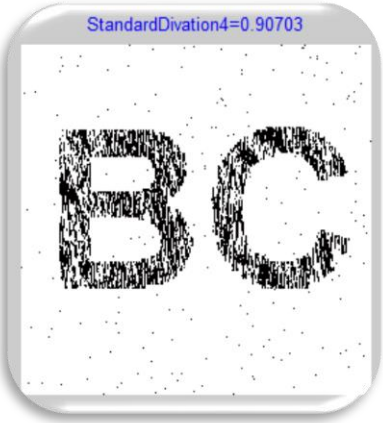

Figure 14: Classic standard deviation

4.2.2. Quantization Attack 8 bit

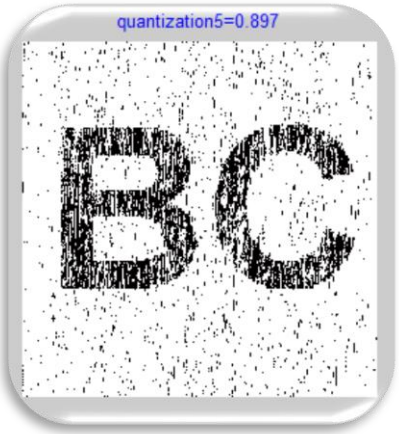

Figure 15: Pop quantization

Figure 16: Classic quantization

4.2.3. Pitch Shift Attack

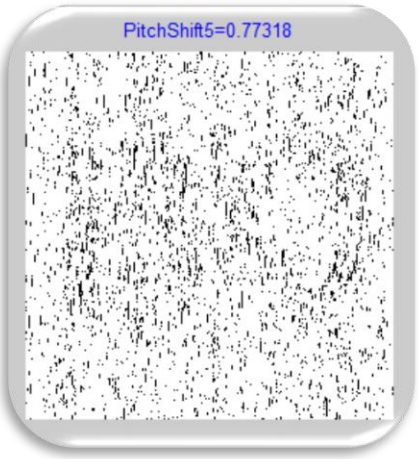

Figure 17: Pop pitch shift

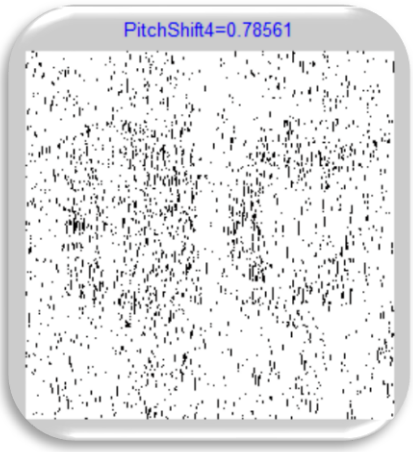

Figure 18: Classic pitch shift

4.2.4. NSR_15 Attack

Figure 19: Pop NSR_15

Figure 20: **Classic NSR_15**

4.2.5. NSR_11 Attack

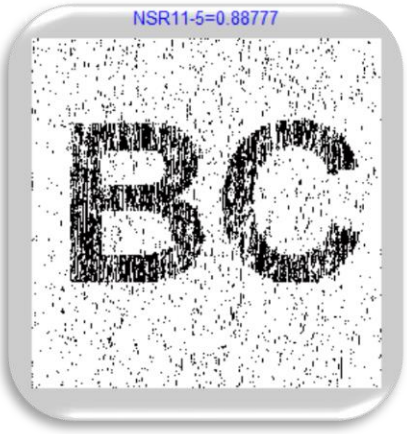

Figure 21: Pop NSR_11

Figure 22: Classic NSR_11

4.2.6. Low Pass Filter Attack

Figure 23: Pop low pass filter

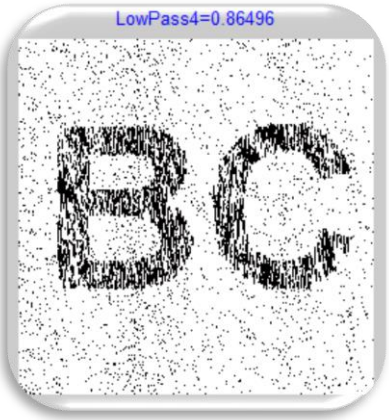

Figure 24: Classic low pass filter

4.2.7. Low Lossy Compression Attack

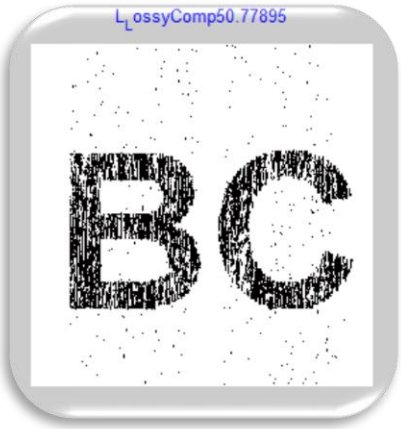

Figure 25: Pop low lossy compression

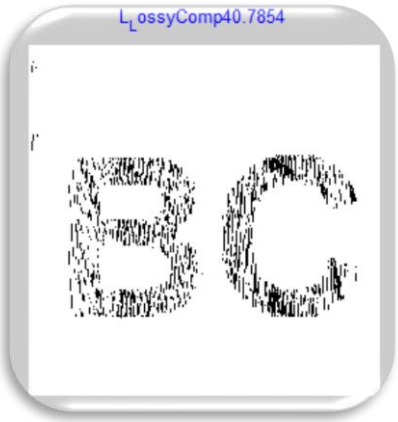

Figure 26: Classic low lossy compression

4.2.8. Medium Lossy Compression Attack

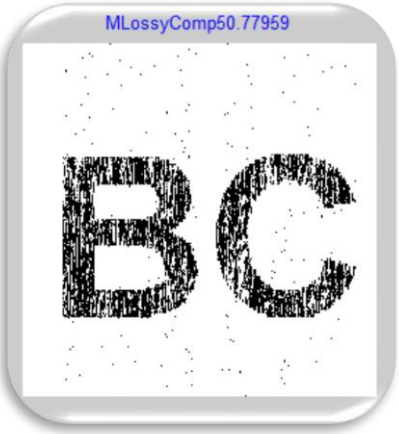

Figure 27: Pop medium lossy compression

Figure 28: Classic medium lossy compression

4.2.9. High Lossy Compression Attack

Figure 29: Pop high lossy compression

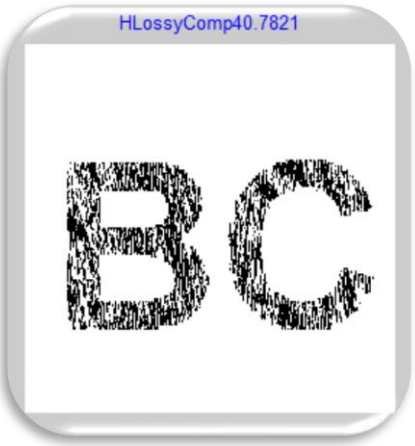

Figure 30: Classic high lossy compression

4.2.10. Gaussian Noise Attack

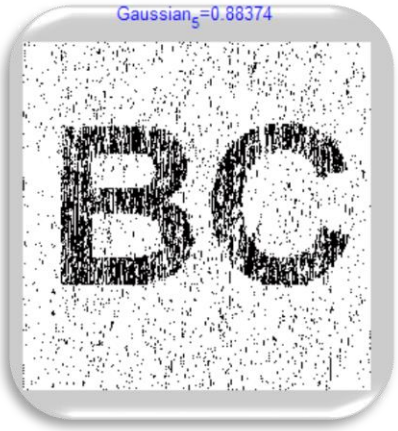

Figure 31: Pop Gaussian noise

Figure 32: Classic Gaussian noise

4.2.11. Amplitude Modification FC =5 Attack

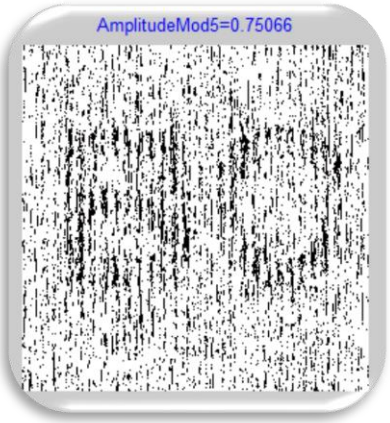

Figure 33: Pop amplitude modification FC =5

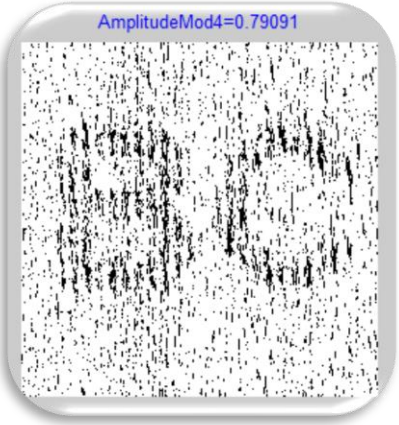

Figure 34: Classic amplitude modification FC =5

4.2.12. Amplitude Modification FC =4 Attack

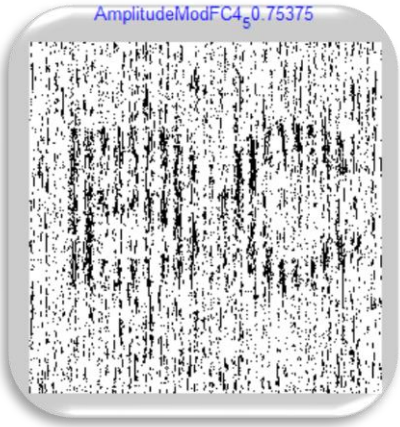

Figure 35: Pop amplitude modification FC =4

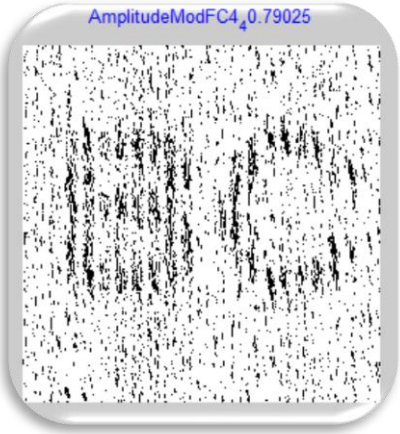

Figure 36: Classic amplitude modification FC =4

4.2.13. Amplitude Modification FC =2 Attack

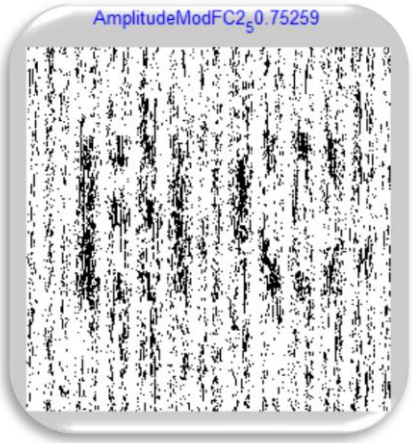

Figure 37: Pop amplitude modification FC =2

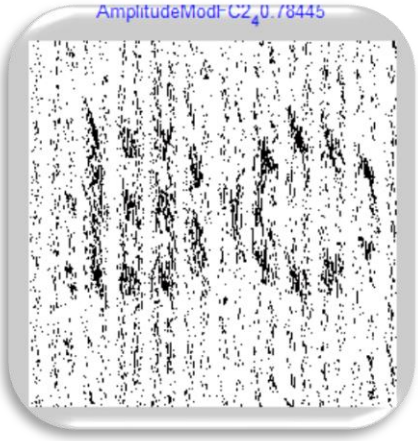

Figure 38: Classic amplitude modification FC =2

4.2.14. Resample Attack 22050

Figure 39: Pop resample 22050

Figure 40: Classic resample 22050

4.2.15. Re-sampling 88200

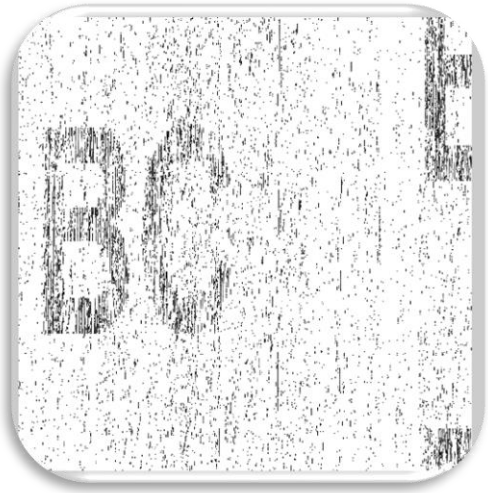

Figure 41: Pop resample 88200

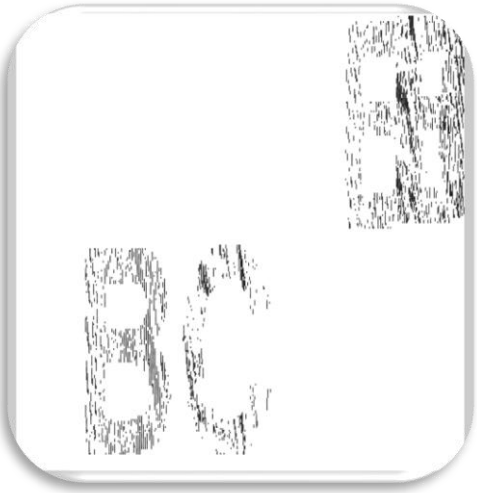

Figure 42: Classic resample 88200

4.2.16. Re-sampling 44000

Figure 43: Pop resample 44000

Figure 44: Classic resample 44000

4.2.17. Time Stretch

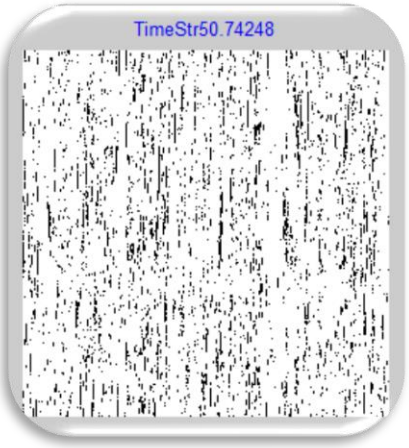

Figure 45: Pop time stretch

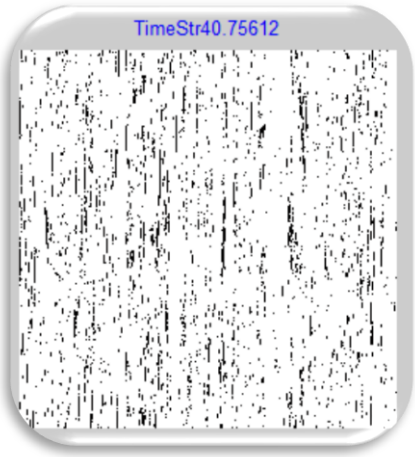

Figure 46: Classic time stretch

4.3. SR (Similarity Ratio) Results

Attack	Pop SNR	Classic SNR
Standard Deviation	0.88295	0.90703
Quantization 8 bits	0.897	0.80203
Pitch Shift	0.77318	0.78561
NSR_15	0.8931	0.91017
NSR_11	0.88777	0.91011
Low Pass Filter	0.77895	0.86496
Low Lossy Compression	0.77895	0.7854
Medium Lossy Compression	0.77959	0.7821
High Lossy Compression	0.8134	0.7821
Gaussian Noise	0.88374	0.80876
Amplitude Modification FC=5	0.75066	0.79091
Amplitude Modification FC=4	0.75375	0.79025
Amplitude Modification FC=2	0.75259	0.78445
Resample 22050	0.93901	0.78018
Time Stretch	0.74248	0.75612

Table 1: SNR of the Extracted Watermarks of Both Audio Clips (Pop, Classic) for All Applied Attacks

In the previous table, we calculate the SR module for all the applied attacks for both watermarked audio clips (Pop and Classic audio). The SR, which is mentioned in chapter (3) section (3.4.1.), refers to how much the similarity is between the extracted watermark image after attack from tested audio clips (Pop and Classic) and the original watermark image, as we notice in the previous table. We have applied 15 attacks, with some of them repeated with different parameters, and the SR parameter is considered to be an effective result as much as it is approximated to one value, i.e. the watermark is to be considered as imperceptible with the original watermark as it becomes closer to one value.

4.4. PSNR (Peak Signal -To- Noise Ratio) Results

Attack	Pop PSNR	Classic PSNR
Standard Deviation	23.3250	31.9631
Quantization 8 bits	53.0223	53.1911
Pitch Shift	12.4833	16.3310
NSR_15	28.5148	32.6082
NSR_11	24.5123	28.6080
Low Pass Filter	21.9101	27.4872
Low Lossy Compression	32.8445	30.8418
Medium Lossy Compression	33.0617	30.8564
High Lossy Compression	33.6958	30.8759
Gaussian Noise	15.3888	18.5462
Amplitude Modification FC=5	10.5616	14.6491
Amplitude Modification FC=4	10.5156	14.6132
Amplitude Modification FC=2	10.5067	14.6351
Resample 22050	33.4829	28.9964
Time Stretch	16.1962	17.4544

Table 2: PSNR for Both Audio Clips (Pop, Classic)

The PSNR, which is mentioned in chapter (3) section (3.3.1.), refers to the imperceptibility between the watermarked audio after each attack and the original audio, to evaluate the amount of distortion that affected the audio clips after each attack. For each attack we found the PSNR value for both watermarked audio clips (Pop and Classic). According to **IFPI** (**International Federation of the Phonographic Industry**), the PSNR should be over 20 dB to be considered as an efficient result, and the PSNR parameters in Table (2) show good results for most of them for both audio.

4.5. MOS Mean Opinion Score Results

Persons	Pop Score From 5	Classic Score From 5
Person1	4	5
Person2	5	4
Person3	5	5
Person4	3	4
Person5	4	3
Person6	5	5
Person7	5	5
Person8	4	5
Person9	5	5
Person10	5	4
MOS Grade From 5	4.5	4.5

Table 3: MOS Results

The previous table shows the values of MOS (Mean Opinion Score), mentioned in chapter (3) section (3.3.2.), that is considered as a subjective evaluation based on the HAS (Human Auditory System). This procedure is done by asking ten people about their opinion, which is to give a score between 5 and 1 about whether the watermarked audio is inaudible or not if it is compared with the original one. This is applied for Pop and Classic audio, the two watermarked tested audio clips, as we notice from the previous Table (3). As we mention in the definition of MOS in chapter (3), the score becomes closer to 5 for each person the MOS shows effective results, that is the mean of the difference between the watermarked audio and the original one for each audio clip is not discovered or noticed from the persons. As we notice at the end of the Table (3) the MOS score for Pop is 4.5 and for Classic is 4.5. Which are very efficient results, which meaning that both the tested audios are inaudible. That achieve very important requirements for watermarking.

4.6. SDG Subjective Difference Grade Results

Attack	Person1		Person2		SDG	
	Pop	Classic	Pop	Classic	Pop	Classic
Standard Deviation	4	5	4	5	4.5	4.5
Quantization 8 bits	4	5	4	5	4.5	4.5
Pitch Shift	1	2	2	1	1.5	1.5
NSR_15	4	5	4	5	4.5	4.5
NSR_11	4	5	4	5	4.5	4.5
Low Pass Filter	5	4	5	4	4.5	4.5
Low Lossy Compression	5	5	5	5	5	5
Medium Lossy Compression	5	5	5	5	5	5
High Lossy Compression	5	5	5	5	5	5
Gaussian Noise	4	3	4	3	3.5	3.5
Amplitude Modification FC=5	4	3	3	4	3.5	3.5
Amplitude Modification FC=4	3	4	4	3	3.5	3.5
Amplitude Modification FC=2	4	4	4	4	4	4
Resample 22050	4	5	4	5	4.5	4.5
Resample 88200	4	5	4	5	4.5	4.5
Resample 44000	5	5	5	5	5	5
Time Stretch	1	1	1	1	1	1

Table 4: SDG Results

The previous table shows the opinion of two persons for both audio clips (Pop and Classic). Each person gave a score as mentioned in chapter (3) section (3.4.2).The Subjective Difference Grade (SDG), which is a subjective technique will be an effective result if it comes closer to 5. This module depends on the HVS system of two persons. The first person gives his score from 5 to 1 to the Pop audio, and this score is given as a comparison between the extracted watermark after the attack and the original watermark, i.e. how much the extracted watermark looks like the original one. The same person gives his score evaluation to Classic audio also. The same evaluation is repeated for person, no.2 for both audio extracted watermarks after each attack. The SDG is given at the right of Table (4) for each attack for both audios, most of the SDG of Table (4) shows the mean equal or more than 4.5, which means

imperceptibility between the extracted and original watermark for both audios. This supports one of the watermarking requirements which is robustness.

4.7. The Effectiveness of the Proposed Watermarking Technique:

In our method we have evaluated our proposed watermarking technique by using subjective and objective techniques. The subjective field is done through using the Mean Opinion Score (MOS) and Subjective Difference Grade (SDG), while the objective field is achieved by using the PSNR and SR modules. MOS is used to measure inaudibility between the watermarked audio clip and original digital audio clip; also the PSNR module is used to evaluate the amount of distortion between the original and watermarked audio. On the other hand the SDG and SR module is used to evaluate the robustness of an extracted watermark.

The efficiency of the watermarking algorithm can be evaluated with respect to the watermark requirements which are Capacity, Inaudibility and Robustness. We give a brief description for the results of our proposed watermarking technique in the following discussion:

1. Capacity:

In our watermarking technique we have used a watermark of 256×256 pixels Fig.(8), and this watermark contains a huge amount of bits to be embedded in an audio clip without affecting the other requirements of the watermarking technique. The proposed watermark consists of two letters with very thick edges to perform very effective robustness against hostile attacks, the proposed watermark with this big amount of bits increases the capacity of our watermarked audio without influencing the imperceptibility as we will notice in the following explanation.

2. Inaudibility (Imperceptibility):

2.1. Inaudibility

The two watermarked audio clips (Pop, Classic) are compared with their original audio clips to evaluate the audibility. For this reason ten persons were asked to give their evaluation (that is related with HAS) for each audio as a score from 1 to 5 depending on the imperceptibility between the watermarked and original audio, and the evaluation results for ten persons are illustrated in section 4.5. Table (3), mean opinion score (MOS) is calculated as a mean for the ten persons scores for each audio clip and the result also is as a mean between 1 and 5. As we illustrated in section (3.3.2.), a score of 5 is given to inaudibility and the score decreases depending on the amount of audibility until reaching a score 1. If we concentrate on the Table (3) of MOS for each audio (Pop and Classic), we could notice that the MOS for Pop =4.5 and Classic =4.5. These results are very effective because these values lie between 4 and 5 and since they are bigger than 4. Therefore, they are classified as equal to 5,and this means that both the watermarked audio clips are inaudible because MOS 5 is given to inaudibility. From this we could conclude that our proposed algorithm is very efficient in supporting one of the important watermarking technique requirements, and the high capacity of our watermark does not affect the inaudibility for both audios.

2.2. Imperceptibility

We have used objective evaluation in the PSNR module to measure the amount of distortion between the original and watermarked signal after the applied attacks. This will support the estimate of the amount of distortion on the watermark in the embedding system so as to evaluate the watermark robustness with SNR results against hostile attacks. PSNR is mentioned in section (3.3.1.), and according to the **International Federation of the Phonographic Industry**, the PSNR value should be more than 20 dB so as to be classified as imperceptible; the PSNR results of the 15 attacks are illustrated in Table (2), and we could conclude that for Standard Deviation, Quantization 8 bits, NSR_15, NSR_11, Low Pass Filter, Low Lossy Compression, Medium Lossy Compression, High Lossy Compression, and Resample 22050, the imperceptibility is very efficient for these mentioned attacks since their PSNR value is greater than 20 dB.

3. Robustness

To estimate the robustness of the proposed watermarking algorithm, we depend on subjective and objective evaluation for the extracted watermark. For objective evaluation the SR module is calculated and the results for this module are illustrated in Table (1), and for subjective evaluation the SDG is performed and the results are illustrated in Table (4). By studying both tables and also studying PSNR table, we could deduce that our watermarking algorithm has conquered the challenging of the following attacks in an effective manner: Standard Deviation, Quantization 8 bits, NSR_15, NSR_11, Low Pass Filter, Low Lossy Compression, Medium Lossy Compression, High Lossy Compression, Resample 22050, Resample 88200, Resample 44000 and Gaussian Noise, and remains robust against Amplitude Modification FC=5, Amplitude Modification FC=4 and Amplitude Modification FC=2 and finally it is not robust against Time Stretch and Pitch Shift. If we take a look at Table (1) of SR results, for some values are very efficient but the others are not, and this is due to the nature of some attacks that affect the position of the watermark inside the watermarked signal without distorting it that leads decreasing the SR from 1. By taking the benefits of SDG in Table (4) that depend on the HVS of two persons for both audio clips (Pop, Classic), the watermarks are detected perfectly. As we notice in Table (4) for most of these attacks, the SDG parameter is equal to 4 and 5, which means that the extracted watermarks are imperceptible since it is greater than (4).

Finally, if we compared our proposed watermarking algorithm with the others' work, the proposed algorithm is effective against Lossy compression for different parameters, while the algorithm of Time Domain in Bassia Paraskevi et al. is not robust for such attack [17]. And for the Resample of different parameters our algorithm has an efficient results, unlike in Tsai [18].

For DWT watermarking algorithm in [19] Al-Haj Ali et al., [23] Wu Shaoquan et al. and [24] Fallahpour Mehdi et al. the MP3 compression is very weak while our algorithm is very robust against Lossy compression with different compression parameters (High, Medium and Low).

In the DCT watermarking algorithm, the algorithm is not robust against Resample in [36] Yongqi Wang et al. unlike our algorithm that is very robust to Resample with a variety of parameters (22050, 88200,44000).

In the FFT algorithm of [33] Dhar Pranab Kumar and Isao Echizen, the robustness to Resample, Re-quantization and MP3 compression is not as effective as our results for the same attacks, and in DCT and DWT algorithms of [30] Dai Hua-liang and Di He, the robustness against Additive Noise is very low, unlike our algorithms that shows very effective robustness against Additive Noise of different types (Standard Deviation, NSR-11, NSR-15, Gaussian Noise).

From the conclusion of each of the foregoing, our proposed watermarking techniques shows an efficient robustness against many of processing attacks, especially for Lossy Compression, Additive Noise, Quantization, Low Pass Filter and Resample and is kept robust against Amplitude Modification but it is not robust against Pitch Shift and Time Stretch.

CHAPTER 5

DISCUSSION AND FUTURE WORK

The audio watermarking technique is a sensitive research area, it has emerged by the propagation of multimedia, especially the wide spread of the digital audio on the internet, and the increased need to manipulate the ownership protection problems of an audio clip, leads to solving the former problems by proposing a variety of audio watermarking algorithm techniques.

In this paper, we proposed audio watermarking techniques that utilized the advantages of the DWT (second level decomposition of low frequency) powerful transforms and RMS (Root Mean Square) model to fulfill the inaudible and robust audio watermarking scheme.

The proposed watermarking algorithm is demonstrated by watermarking two different characteristics audio clips Pop and Classic, it is an effective watermarking technique that has the highest payload of watermark 256×256 bits and robustness against many malicious audio outtakes.

Ongoing research is focused on the resistance to the synchronization attack like Time Stretch and Pitch shift by adding a synchronization code to the original signal.

REFERENCES

1. http://www.google.com.tr/books?hl=en&lr=&id=tatqoONPzHUC&oi=fnd&pg=PA1&dq=1.%09Pan+J.,+Huang+H.+C.,+Jain+L.+C.,+Eds,+(2004),+%E2%80%9CIntelligent+Watermarking+Techniques%E2%80%9D,+vol.+7,+World+Scientific.&ots=Rj_KrcWxMS&sig=cI9St-AN3ZoJzL03m3GeXEm9nZM&redir_esc=y#v=onepage&q&f=false
(Data Download Date:09.09.2014)

2. http://123seminarsonly.com/Seminar-Reports/011/56854906-Steganography.pdf
(Data Download Date:09.09.2014)

3. http://www.google.com.tr/books?hl=en&lr=&id=OWhPeGK93VcC&oi=fnd&pg=PR17&dq=Information+Hiding:+Steganography+and+Watermarking+Attacks+and+Countermeasures+3.%09Johnson&ots=0hGg1brOH7&sig=zTOBCbSZqyyJXGNWlxn0aeERI2E&redir_esc=y#v=onepage&q=Information%20Hiding%3A%20Steganography%20and%20Watermarking%20Attacks%20and%20Countermeasures%203.%09Johnson&f=false (Data Download Date:09.09.2014)

4. http://128.232.0.20/teaching/0910/R08/work/essay-ma485-watermarking.pdf
(Data Download Date:09.09.2014)

5. **Chen D. Y., Ouhyoung M., Wu J. L., (2000),** "*A Shift-Resisting Public Watermark System for Protecting Image Processing Software*", Consumer Electronics, IEEE Transactions, vol. 46(3), pp. 404-414.

6. **Kutter M., Petitcolas F., (1999),** "*Fair Benchmarking for Image Watermarking System*", Proc. of SPIE 3657, Security and Watermarking of Multimedia Contents, pp.226-239.

7. http://www.armageddononline.org/PDF/Smuggling%20&%20Caching/Artech%20House%20Information%20Hiding%20Techniques%20for%20Steganography%20and%20Digital%20Watermarking.pdf
(Data Download Date:09.09.2014)

8. **Petitcolas F., (2000),** "*Watermarking Schemes Evaluation*", IEEE Signal Processing Magazine [Online], vol. 17(5), pp. 58-64.

9. http://etd.lsu.edu/docs/available/etd-10182010-105045/unrestricted/Ravulathesis.pdf (Data Download Date:09.09.2014)

10. http://www.google.com.iq/#q=Watermarking+Techniques+Spatial+Domain+Digital+Rights+Seminar++El-Gayyar (Data Download Date:09.09.2014)

11. **Cox, I. J., Miller M. L., Bloom J. A., (2000),** "*Watermarking Applications and Their Properties*", In Information Technology: Coding and Computing, International Conference, pp. 6.

12. **Alaryani H., Youssef A., (2005),** "*A Novel Audio Watermarking Technique Based on Low Frequency Components*", Proceedings of the 7th IEEE International Symposium on Multimedia (ISM'05), pp. 6.

13. **Zhang X., Yin X., Yu Z., (2008),** "*Histogram Specification-Based Audio Watermarking Technology Against Filtering Attacks in Time Domain*", Electronic Commerce and Security, International Symposium, IEEE, pp. 951-956.

14. **Lie W. N., Chang L. C., (2011),** "*Robust and High-Quality Time-Domain Audio Watermarking Subject to Psychoacoustic Masking*", Circuits and Systems, ISCAS 2001. The 2001 IEEE International Symposium, vol. 2, pp. 45-48.

15. **Shahriar M. R., Sangjin C., Ui-pil C., (2012),** "*Time-Domain Audio Watermarking Using Multiple Marking Spaces*", Informatics Electronics & Vision (ICIEV), International Conference, IEEE, pp. 974-979.

16. **Martinez-Noriega R., Nakano M., Yamaguchi K., (2010),** "*Self-Synchronous Time-Domain Audio Watermarking Based on Coded-Watermarks*", Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 6th International Conference, IEEE, pp. 135-138.

17. **Bassia P., Pitas I., Nikolaidis N., (2001),** "*Robust Audio Watermarking in The Time Domain*", Multimedia IEEE Transactions, vol. 3(2), pp. 232-241.

18. **Tsai H. H., Cheng J. S., (2005),** "*Adaptive Signal-Dependent Audio Watermarking Based on Human Auditory System and Neural Networks*", Applied Intelligence, vol. 23(3), pp. 191-206.

19. **Al-Haj A., Twal C., Mohammad A., (2010),** "*Hybrid DWT-SVD Audio Watermarking*", Digital Information Management (ICDIM), 5th International Conference on. IEEE, pp. 525-529.

20. **Tianchi L., Guangming Y., Qi W., (2011),** "*A Multiple Audio Watermarking Algorithm Based on Shear Resisting DWT and LSB*", In Networked Computing (INC), The 7th International Conference, IEEE, pp. 78-83.

21. **Al-Yaman, M. S., Al-Taee M. A., Alshammas H. A., (2012),** "*Audio-Watermarking Based Ownership Verification System Using Enhanced DWT-SVD Technique*", Systems Signals and Devices (SSD), 9th International Multi-Conference, IEEE, pp. 1-5.

22. **Elshazly A. R., Fouad M. M., Nasr M. E., (2012),** "*Secure and Robust High Quality DWT Domain Audio Watermarking Algorithm With Binary Image*", Computer Engineering & Systems (ICCES), 17th International Conference, IEEE, pp. 207-212.

23. **Wu S., Huang J., Huang D., Shi Y. Q., (2005),** "*Efficiently Self-Synchronized Audio Watermarking for Assured Audio Data Transmission*", Broadcasting IEEE Transactions, vol. 51(1), pp. 69-76.

24. **Fallahpour M., Megías D., (2011),** "*High Capacity Audio Watermarking Using the High Frequency Band of the Wavelet Domain*", Multimedia tools and Applications, vol. 52(2-3), pp. 485-498.

25. **Peng H., Wang J., Zhang Z., (2013),** "*Audio Watermarking Scheme Robust Against Desynchronization Attacks Based on Kernel Clustering*", Multimedia tools and applications, vol. 62(3), pp. 681-699.

26. **Chen S. T., Wu G. D., Huang H. N., (2010),** "*Wavelet-Domain Audio Watermarking Scheme Using Optimisation-Based Quantisation*", IET Signal Processing, vol. 4(6), pp. 720-727.

27. **Wu S., Huang J., Huang D., Shi Y. Q., (2005),** "*Efficiently Self-Synchronized Audio Watermarking for Assure Audio Data Transmission*", Broadcast IEEE Transactions, vol. 51(1), pp. 69–76.

28. **Lie W. N., Chang L. C., (2006),** "*Robust and High-Quality Time Domain Audio Watermarking Based on Low-Frequency Amplitude Modification*", Multimedia IEEE. Transactions, vol. 8(1), pp. 46–59.

29. **Wang X. Y., Zhao H., (2006),** "*A Novel Synchronization Invariant Audio Watermarking Scheme Based on DWT and DCT*", Signal Processing IEEE Transactions, vol.54(12), pp.4835-4840.

30. **Dai H. L., He D., (2009),** "*An Efficient and Robust Zero-Watermarking Scheme for Audio Based on DWT and DCT*", Microelectronics & Electronics, Prime Asia 2009, Asia Pacific Conference on Postgraduate Research, IEEE, pp. 233-236.

31. **Chen N., Zhu J., (2007),** "*Robust Speech Watermarking Algorithm*", Electronics Letters, vol. 43(24), pp. 1393-1395.

32. **Ren K., Li H., (2011),** "*Large Capacity Digital Audio Watermarking Algorithm Based on DWT and DCT*", Mechatronic Science Electric Engineering and Computer (MEC), International Conference, IEEE, pp. 1765-1768.

33. **Dhar P. K., Echizen I., (2011),** "*Robust FFT Based Watermarking Scheme for Copyright Protection of Digital Audio Data*", Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 17th International Conference, IEEE, pp. 181-184.

34. **Kang X., Yang R., Huang J., (2011),** "*Geometric Invariant Audio Watermarking Based on An LCM Feature*", Multimedia IEEE Transactions, vol. 13(2), pp. 181-190.

35. **Liu J. X., Lu Z. M., Pan J. S., (2008),** "*A Robust Audio Watermarking Algorithm Based on Dct and Vector Quantization*", Intelligent Systems Design and Applications (ISDA'08), 8th International Conference, IEEE, vol. 3, pp. 541-544.

36. **Yongqi W., Yang Y., (2008),** "*A Synchronous Audio Watermarking Algorithm Based on Chaotic Encryption in DCT Domain*", International Symposium on Information Science and Engineering, vol. 2, pp. 371-374.

37. **Guo Q., Zhao Y., Cheng P., Wang F., (2012),** "*An Audio Digital Watermarking Algorithm Against A/D and D/A Conversions Based on DCT Domain*", Consumer Electronics Communications and Networks (CECNet), 2nd International Conference, IEEE, pp. 871-876.

38. **Xiong-Hua H., Wei-Zhen J., Xing-Xing J., (2010),** "*Robust Audio Watermarking Based Non-Uniform DCT*", Intelligent Computing and Intelligent Systems (ICIS), IEEE International Conference, vol. 1 pp. 585-588.

39. http://www.irma-international.org/viewtitle/4685/
(Data Download Date:09.09.2014)

APPENDICES A

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: MARAHA, Heyam.

Date and Place of Birth: 8 June 1976, Baghdad.

Marital Status: Married.

Phone: 535 640 6198.

Email: eheyam@ymail.com

EDUCATION

Degree	Institution	Year of Graduation
M.Sc.	Çankaya Univ., Mathematics and Computer Science	2014
B.Sc.	Mosul Univ., Computer Science	1998
High School	Al-Talaye	1994

WORK EXPERIENCE

Year	Place	Enrollment
2002- Present	Kirkuk Univ. Department of Computer Science	Laboratory Trainer
2001	Mosul University	Laboratory Trainer

FOREIN LANGUAGES

English.

HOBBIES

Travel, Books.