A HIGH CAPACITY AUDIO STEGANOGRAPHY MODEL BASED ON FRACTAL CODING

AHMED HUSSAIN ALI

UNIVERSITI KEBANGSAAN MALAYSIA

A HIGH CAPACITY AUDIO STEGANOGRAPHY MODEL BASED ON FRACTAL CODING

AHMED HUSSAIN ALI

THESIS SUBMITTED IN FULFILMENT FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF INFORMATION SCIENCE AND TECHNOLOGY
UNIVERSITI KEBANGSAAN MALAYSIA
BANGI

2018

MODEL STEGANOGRAFI AUDIO BERKAPASITI TINGGI BERASASKAN KOD FRAKTAL

AHMED HUSSAIN ALI

TESIS YANG DIKEMUKAKAN UNTUK MEMPEROLEHI
IJAZAH DOKTOR FALSAFAH

FAKULTI TEKNOLOGI DAN SAINS MAKLUMAT
UNIVERSITI KEBANGSAAN MALAYSIA
BANGI

2018

DECLARATION

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

27 April 2018

AHMED HUSSAIN ALI
P72284

ACKNOWLEDGEMENT

Foremost, I would like to express my sincere gratitude towards our God (Allah), who is the all-powerful and all-knowing creator; Allah has bestowed countless blessings upon us. Some examples of which are, Allah has endowed us with the gifts of sight and hearing, the intellect, health, wealth, and family. Allah has even subjected everything in the universe for us: the sun, the moon, the heavens and the earth, and many countless things, as the Qur'an states, "If you tried to number Allah's blessings, you could never count them." (Soorat Al-Maa'idah, 16:18).

Secondly, I am extremely thankful to Allah, for connecting me with the best possible supervisor Dr. **Mohd Rosmadi Mokhtar**, whose unlimited guidance, suggestions, encouragement, inspirational character and motivations have enabled me to do my utmost best during these wonderful years of my study. I am honored and consider myself fortunate to have worked under his supervision.

Also, I would like to express my high appreciation to my co-supervisor Dr. **Loay Edwar George**, whose guidance, advice and ideas have improved my research many-fold.

I am very grateful to my Ph.D. studentship sponsorship by the Iraqi government represented by "the Ministry of Higher Education and Scientific Research" which makes this study in this field possible and helped provide such an educational environment to work in. Furthermore, I would also like to recognize with much thankfulness the role of all those employees involved in Iraqi Cultural Attache in Malaysia throughout my Ph.D. studies.

I would like to thank my parents, **Hussain and Khalida** for making me who I am today. My success symbolizes and reflects the support and love I have received from both of them.

I would like to express my appreciation to my wife **Zaynab Layth Abduljabbar** and my children **Mustafa, Meena, and Maria** whose support, encouragement, motivations, and understanding enabled me to work on my thesis activities.

I would like to thank my lab mate students, **Abdo Ali Alwasibi, Nader Salama and Nadeem Alharbawi** in UKM, Cyber Security research group for their help, friendship, and creating a pleasant working environment throughout my years in UKM.

ABSTRACT

Securing confidential information over an unsecured communication channel or environment while preserving privacy has become an essential aspect in today's digital world. Cryptography and information hiding algorithms like steganography and watermarking are the most commonly adopted solutions for the issues of data confidentiality. While cryptography distorts the secret information into unreadable texts or files, steganography conceals the secret information in different covers of different types such as text, image, audio, video, and network protocol. For that reason, the main goal of steganography is to generate stego files with three opposing features in terms of hiding capacity, transparency, and robustness. However, achieving all these features at the same time is a challenging process since they are contradictory. Therefore, the purpose of this study is to improve audio steganography performance by considering the related requirements. This is done by adopting the Design Science Research Methodology guidelines as the main research methodology. In achieving this, four objectives have been placed. The first objective is to identify an applicable compression technique that can produce a high compression ratio and acceptable audio file. The second objective of this study is to increase the hiding capacity and preserve the transparency of the stego audio by developing a cover-secret mapping algorithm based on fractal coding. The third objective is to maintain the tradeoff between robustness and transparency by developing an embedding algorithm in the time and transform domains based on the least significant bit and uniform coefficient modulation, respectively with the chaotic function. Lastly, the final objective is to propose a new audio steganography model that simultaneously adjusts the requirements based on the two proposed algorithms. The performance of the proposed model is evaluated by conducting a group of experiments using the GTZAN dataset. Objective evaluation is used to assess the achieved hiding capacity, transparency, and robustness under signal processing attacks. Statistical steganalysis is also used to evaluate the statistical transparency and the robustness of the generated stego audio using histogram and fourth first moment steganalysis. The empirical findings show that the fractal coding can be used as a compression technique and applied directly on the audio file with a compression ratio of up to 90% with an acceptable quality of the compressed audio file PSNR of 39 dB on average. The hiding capacity is improved by 30% in the time and transform domains in comparison with similar techniques. At the same time, the transparency in terms of SNR is preserved at around 70 and 50 dB in the time and transform domains, respectively. Furthermore, the proposed model has been shown to withstand brute-force attack with an ample amount of hiding possibilities. Moreover, the results also expose the robustness against the following attacks: histogram error ratio around 0.1 and statistical analysis using fourth first moment with a maximum difference of less than 0.01% in the time and transform domains, additive white Gaussian noise with 70 dB, echo addition with decay of 5% delay in 50ms, and cropping with 3% of the stego audio size. These results signify that the proposed model and the two algorithms have notably enhanced the hiding capacity of audio steganography with an acceptable transparency and robustness rate.

ABSTRAK

Mendapatkan maklumat sulit melalui saluran komunikasi atau persekitaran yang tidak selamat di samping memelihara privasi telah menjadi satu aspek penting di dalam dunia digital hari ini. Kriptografi dan algoritma menyembunyikan maklumat seperti steganografi dan tera air adalah penyelesaian yang paling umum digunakan untuk masalah kerahsiaan data. Walaupun kriptografi memesongkan maklumat rahsia supaya menjadi teks atau fail yang tidak boleh dibaca, steganografi menyembunyikan maklumat rahsia itu di dalam beberapa penutup berbagai jenis yang berbeza seperti teks, imej, audio, video, dan protokol rangkaian. Atas sebab itu, matlamat utama steganografi adalah untuk menjanakan fail stego dengan tiga ciri berlawanan dari segi keupayaan penyembunyian, ketelusan, dan keteguhan. Walau bagaimanapun, mencapai semua ciri ini pada masa yang sama adalah proses yang mencabar kerana mereka bercanggah. Oleh itu, tujuan kajian ini adalah untuk meningkatkan prestasi steganografi audio dengan mempertimbangkan keperluan-keperluan yang berkaitan. Ini dilakukan dengan menggunakan garis panduan Metodologi Penyelidikan Sains Reka Bentuk sebagai metodologi penyelidikan utama. Dalam mencapai matlamat ini, empat objektif telah ditempatkan. Objektif pertama ialah mengenal pasti suatu teknik mampatan yang boleh menghasilkan nisbah mampatan tinggi dan fail audio yang boleh diterima. Objektif kedua kajian ini adalah untuk meningkatkan keupayaan penyembunyian dan memelihara ketelusan audio stego tersebut dengan membangunkan algoritma pemetaan penutup-rahsia berdasarkan kod fraktal. Objektif ketiga adalah untuk mengekalkan penukaran antara kekukuhan dan ketelusan dengan membangunkan suatu algoritma pembenaman pada domain masa dan jelmaan berdasarkan bit paling kurang signifikan dan modulasi koefisien seragam, masing-masing dengan fungsi kacau bilau. Akhir sekali, matlamat terakhir adalah untuk mencadangkan suatu model steganografi audio baru yang serentak menyesuaikan keperluan berdasarkan dua algoritma yang dicadangkan. Prestasi model yang dicadangkan dinilai dengan menjalankan sekumpulan eksperimen menggunakan set data GTZAN. Penilaian objektif digunakan untuk menilai keupayaan penyembunyian yang dicapai, ketelusan, dan keteguhan menghadapi serangan pemprosesan isyarat. Analisis stego statistik juga digunakan untuk menilai ketelusan statistik dan keteguhan audio stego yang dijanakan menggunakan histogram dan analisis stego momen pertama yang keempat. Hasil kajian empirikal menunjukkan bahawa pengekodan fraktal boleh digunakan sebagai teknik mampatan dan digunakan secara langsung pada fail audio dengan nisbah mampatan sehingga 90% dengan kualiti yang boleh diterima bagi PSNR fail audio termampat sebanyak 39 dB secara purata. Kapasiti penyembunyian bertambah baik sebanyak 30% di domain masa dan jelmaan berbanding dengan teknik yang serupa. Pada masa yang sama, ketelusan dari segi SNR terpelihara di sekitar 70 dan 50 dB di domain masa dan jelmaan, masing-masing. Selain itu, model yang dicadangkan telah ditunjukkan sebagai boleh menahan serangan daya-kasar dengan kemungkinan penyembunyian yang mencukupi. Tambahan lagi, hasilnya juga menunjukkan kekukuhan terhadap serangan berikut: nisbah ralat histogram sekitar 0.1 dan analisis statistik menggunakan momen pertama keempat dengan perbezaan maksimum kurang dari 0.01% di domain masa dan jelmaan, hingar Gaussan tambahan 70 dB, tambahan gema dengan susut kelewatan 5% dalam 50ms, dan potongan sebanyak 3% daripada saiz audio stego. Hasil-hasil ini menandakan bahawa model yang dicadangkan dan kedua algoritma ini telah meningkatkan kapasiti penyembunyian steganografi audio dengan kadar ketelusan dan keteguhan yang boleh diterima.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BSD	Bark Spectral Distortion
DCT	Discrete Cosine Transform
DR	Difference Ratio
DSP	Digital Signal Processing
DSRM	Design Science Research Methodology
DWT	Discrete Wavelet Transform
DWPT	Discrete Wavelet Packet Transform
EWM	Efficient Wavelet Masking
FAC	Fractal Audio Coding
FC	Fractal Coding
FFT	Fast Fourier Transform
FPs	Fractal Parameters
HAS	Human Auditory System
HC	Hiding Capacity
HER	Histogram Error Rate
HVS	Human Visual System
ICT	Information and Communication Technology
IFS	Iteration Function System
IWT	Integer Wavelet Transform
LSB	Least Significant Bit
LWT	Lifting Wavelet Transform
MOS	Mean Opinion Scores
MP3	MPEG-1 Audio Layer-3
MSE	Mean Square Error
MSB	Most Significant Bit

NC	Normalized Correlation
OVSF	Orthogonal Variable Spreading FactorS
PCM	Pulse Code Modulation
PESQ	Perceptual Evaluation Sound Quality
PIFS	Partition Iteration Function System
PRNG	Pseudo Random Number Generator
PSNR	Peak Signal to Noise Ratio
PVD	Pixel Value Differencing
QIM	Quantization Index Modulation
RMS	Root Mean Square
SDG	Subjective Difference Grade
SNR	Signal to Noise Ratio
SS	Spread Spectrum
SPCC	Squared Pearson Correlation Coefficient
SSA	Shift Spectrum Algorithm
UCM	Uniform Coefficients Modulation
VoIP	Voice Over Internet Protocol
VQ	Vector Quantization
WBM	Weighted Block Matching

CHAPTER I

INTRODUCTION

1.1 RESEARCH OVERVIEW

The rapid advancements in digital communication technology have produced an exponential boost in the use of the Internet for various commercial, governmental and social interactions that involve the transmission of a variety of confidential data and multimedia objects. The difficulties of ensuring the privacy of individuals have become increasingly challenging. The degrees to which individuals appreciate confidentiality differ from one person to another (Ghebleh & Kanso 2014). Securing the content and the privacy of sensitive information as well as personal transactions over open networks has become essential but increasingly important. Therefore, the area of information and multimedia security research has attracted more interest, and its scope of applications expands significantly (Nissar & Mir 2010; Mat Kiah et al. 2011).

The objectives of the information security are confidentiality, integrity, and authenticity of the sensitive information, and are also known as CIA. The importance of the information and the expected attack by unauthorized parties define the required procedures and the security requirements since not all the information has the same level the importance. Highly secret data as in governments, military, and banking require highly secure procedures and assign different levels of priorities to the authorized parties, while in the academic and scientific application the security is an emphasis on preventing the unauthorized use of resources (Von Solms & Van Niekerk 2013).

Many approaches have been investigated and developed to protect personal confidentiality. Information security is the practice of providing secure transmission of important data that consists of two main techniques, cryptography and information hiding (Baig et al. 2016). Cryptography renders the secret data meaningless and unreadable to the attackers whereas steganography conceals the existence of them. Cryptography is the earlier approach for information security, and steganography has since emerged. Encryption leads to scrambling the secret data inside the encrypted file and making the file suspicious to the adversary while steganography hides the secret data in an innocent file making them not observable (Zaidan et al. 2010).

Information hiding can be divided into two classes, namely, watermarking and steganography (Djebbar et al. 2012). Watermarking achieves copyright protection or ownership by embedding watermarks inside the media. Steganography hides and transmits confidential data and their existence simultaneously. Hiding capacity, transparency, and robustness are the goals of the steganography technique (Cheddad et al. 2010; Zaidan et al. 2016; Kaur et al. 2017).

This chapter provides a general introduction to the research area, some background on the research with a problem statement and research question. Moreover, the research objectives are identified based on the research problem followed by the motivation and the scope of this study. This chapter is closed by highlighting the organization of the thesis.

1.2 RESEARCH BACKGROUND

The enormous volume of sensitive and confidential information is transferred over public and untrusted communication channels. Many attempts are exerted by the intruder to break the security of these data. As a result, various security mechanisms have been applied to improve data confidentiality and privacy. Data hiding or Steganography techniques are one of the promising solutions in this area (Nguyen et al. 2015).

Various steganography techniques involve using distinct media such as images, text, audio, video, and network protocols. The Human Visual System (HVS) is less sensitive to the image alteration while the Human Auditory System (HAS) is more sensitive even to a little distortion in the audio cover files, the fact that has made the audio steganography techniques less comparative (Fridrich 2009; Garcia-Hernandez et al. 2013; Hemalatha et al. 2016). Therefore, proposing new audio steganography is considered a challenge for researchers (El-Khamy et al. 2016).

In the same way, this is considered a motivation for the researchers to develop new audio steganography techniques (Trivedi et al. 2017). Moreover, the steganalysis techniques, which means the techniques that are proposed to discover the secret object from the cover media, are targeted to the image steganography schemes more than for audio steganography approaches (Djebbar et al. 2012; El-Khamy et al. 2016).

The evaluation performance of any steganography technique, in general, is controlled by three primary requirements which are transparency, high embedding capacity, and robustness (Cheddad et al. 2010; Shahadi et al. 2016). Transparency means the similarity or the minimum degradation between the stego and the original cover. Hiding capacity reflects the relation of the size of the secret file to that of the cover file. Robustness indicates the resistance of stego files to various attacks and capability of retrieving the secret message with minimum error. Robustness is the most important requirement in watermarking techniques since the aim is copyright protection so, the watermark should be robust against different types of attack while in steganography techniques, transparency and hiding capacity are important since the aim is achieving covert communications (Ballesteros L & Moreno A 2012; Renza et al. 2017).

These requirements are mostly related to each other in that they are contradictory (Shahadi et al. 2014; Tang et al. 2016). Increasing one requirement leads to decreasing the other two requirements. Attaining these requirements simultaneously for one scheme remains a challenging task. Transparency is controlled by the degradation introduced on the cover media during the embedding process for the secret data. Increasing the hiding capacity is a primary requirement in steganography

technique. However, this cannot be attained because of the cost of transparency (Ballesteros L & Moreno A 2012; Tang et al. 2016). Robustness is an important requirement in the watermarking techniques. It means that the watermark should resist the unintentional attack such as signal processing operations and malicious attack of removal or degradation (Lei et al. 2012; Lin et al. 2015).

On the other hand, data hiding techniques can be classified into several domains according to the applied domain. Several works (Lee et al. 2013; Ali et al. 2016) have classified the steganography techniques into three domains namely temporal or time, transform and compressed domains while Balgurgi & Jagtap (2013); El-Khamy et al. (2016) classified them into temporal, frequency and transform domains. Each domain has advantages and weaknesses that differ from the others as shown in more detail in Section 2.3.3. The initial domain adopted in data hiding techniques was the time domain while the frequency domain came later in order to resolve the main drawbacks of the time domain which is the robustness. The new trend is directed to adopting data compression techniques in order to increase the hiding capacity which is the main requirement in steganography techniques along with preserving the transparency of the cover file (Lin et al. 2015; Malik et al. 2016; Tang et al. 2016).

1.3 PROBLEM STATEMENT

Due to the importance of audio steganography in military, governments, intelligence agencies and other organizations and persons that want to achieve covert communications for securing the exchange of the sensitive or private information, extensive research has been carried out on this aspect (Djebbar et al. 2012; Kar & Mulkey 2015; Chowdhury et al. 2016; El-Khamy et al. 2016; Hemalatha et al. 2016; Kanhe & Aghila 2016; El-Khamy et al. 2017; Renza et al. 2017).

The steganography system performance in terms of hiding capacity and transparency is the target of researchers (Ballesteros L & Moreno A 2012; Hemalatha et al. 2016). Various audio steganography schemes in literature adopt two domains in the embedding and extraction process which are time and transform domains. The

reason behind adopting these domains is to enhance the performance of the audio steganography system through increasing the hiding capacity and the transparency. Although each domain has its weaknesses as discussed in Chapter 2, Section 2.3.3, they are still adopted by the researchers in audio steganography systems (Rekik et al. 2012; El-Khamy et al. 2016).

Steganography techniques should also be robust against some unintentional attacks such as signal processing. The challenge in steganography techniques is represented by the tradeoff between these requirements since each one is contradicting the others (Fridrich 2009; Ballesteros L & Moreno A 2012; Shahadi et al. 2016). For instance, increasing hiding capacity leads to increasing the degradation of the stego fidelity and consequently decreasing the transparency. On the other hand, increasing the robustness level results in decreasing the hiding capacity and the transparency. Fulfilling these requirements in one particular scheme is considered the challenge in this area (Shahadi et al. 2014; Tang et al. 2016).

In relation to the hiding capacity, one way that can increase it is to reduce the size of the secret file. The reduction in the size can be achieved by adopting the data compression. Thus, incorporating the data compression in the audio steganography can increase the hiding capacity same as in image and text steganography (Begum & Venkataramani 2012; Lin et al. 2015; Malik et al. 2016; Tang et al. 2016).

However, several compression techniques were proposed in the literature which can be used. They are classified into lossy and lossless techniques. Lossy compression techniques are widely used since they achieve high hiding capacity and quality of the reconstructed files (Sayood 2006; Mammeri et al. 2012). Lossy techniques such as Discrete Wavelet transform (DWT), Discrete Cosine Transform (DCT), Vector Quantization (VQ) and Fractal Coding (FC) are examples of this type of compression techniques. Each of them provides different compression ratio and reconstructed file quality. Thus, it is necessary to identify which technique can give high compression ratio and acceptable audio quality. Based on the literature, the issues in audio steganography can be focused in three directions:

The first issue is related to the hiding capacity and transparency. Hiding capacity means the amount of the secret data that can be hidden in the cover while transparency refers to the perceptual similarity between the cover and the stego file. More details about these two requirements are discussed in Chapter 2, Section 2.3.4. The issue is represented by the tradeoff between the hiding capacity and transparency. The hiding capacity has an inverse relationship with the transparency (Shahadi et al. 2016; Wu et al. 2016). When hiding large secret data is required within a specific size of cover audio, this inevitably causes distortion in the stego file which attracts the eavesdropper's attention and thus the hidden data becomes detectable. The achieved hiding capacity in the related works in the time domain is 70% with the transparency of the stego audio file of 52 dB while in transform domain the hiding capacity is 50% with transparency 39 dB of the stego audio. Increasing the hiding capacity more than the achieved in time and transform domain leads to decreasing the fidelity less than 52 and 39 dB, respectively since it needs to modify more cover samples in order to hide the additional secret data. So it is essential to adopt techniques to reduce the size of the secret file instead of modifying more cover samples to preserve the transparency of the stego file. Maintaining the tradeoff between the hiding capacity and stego transparency is considered the first obstacle when designing audio-based steganography approaches (Garcia-Hernandez et al. 2013; Ma et al. 2015).

The second issue is an inverse relationship between robustness which means the resistance against attacks and hiding capacity on one hand and robustness with transparency, on the other hand, that is hard to achieve in one steganography system (Bender et al. 1996; Djebbar et al. 2011; Djebbar et al. 2012). Increasing the robustness of the hidden data requires adopting techniques for cover samples modification in a way that affects the transparency of the stego file, so it requires adopting a technique that can improve the robustness without affecting the transparency of the stego file. On the other hand, increasing hiding capacity more than 70 and 50 % in time and transform domain, respectively results in decreasing the robustness level in terms of steganalysis and signal processing attacks since more data to be hidden means more stego artifacts and then vulnerability to attacks.

The third issue in the audio steganography is related to adjusting all the steganography requirements simultaneously (Tang et al. 2016). Adjusting the requirements means the ability to preserve the tradeoff between them by increasing the hiding capacity more than 70 and 50% of the cover size in time and transform domains, respectively along with enhancing or maintaining the transparency of 52 and 39 dB in time and transform domain, respectively and resisting against some of steganalysis types and attacks such as signal processing operations. Adjusting these requirements is demanded in designing and developing any successful steganography or watermarking schemes (Bhat et al. 2010; Lei et al. 2012).

In the last ten years, the term adaptive steganography has slowly emerged, which means adopting former methods before or during the embedding process to enhance the performance of the steganographic system. Such methods are used to either specify the appropriate positions in the cover for embedding or compress the secret or cover data to reduce the size.

Data compression techniques are utilized with text steganography techniques as in (Begum & Venkataramani 2012; Malik et al. 2016) and image steganography in (Tsai & Tsai 2011; Ma et al. 2015); however, compression techniques have not been widely investigated in audio steganography.

Therefore, the interest of this study is to solve some of the above issues represented by the tradeoff between steganography requirements by proposing an audio steganography model for enhancing hiding capacity while maintaining transparency and robustness. The details of the proposed model are discussed in Chapter 4. The objective measurements are used to evaluate the performance of the proposed audio steganography model as they are presented in Chapter 5.

1.4 RESEARCH QUESTIONS

This research addresses the following research questions in order to organize the direction of this study:

1. What is the applicable compression technique that can produce high compression ratio and acceptable audio file?

2. What is the technique that can increase the hiding capacity more than 70% and 50% without affecting or preserving the transparency of the stego audio of 52 and 39 dB in time and transform domain, respectively?

3. Which algorithm can maintain transparency of 52 and 39 dB in time and transform domain, respectively and resist steganalysis and signal processing operation attacks when the hiding capacity is increased?

4. What is the possibility of achieving high hiding capacity, transparency and acceptable robustness in a particular steganography model simultaneously?

1.5 RESEARCH OBJECTIVES

This study is dedicated to solving the described problem and answering the research question defined in the previous section. The aim of this study is to improve the performance of the audio steganography by proposing a model that can enhance hiding capacity while preserving transparency and robustness at the same time. In addition, two prototypes system are also developed in this study. The objectives of this study are:

1. To identify an applicable compression technique that can produce high compression ratio and acceptable audio file.

2. To develop a cover-secret mapping algorithm to increase the hiding capacity and maintain the transparency of the stego file.

3. To develop an embedding algorithm to maintain the tradeoff between transparency and robustness along with the achieved high capacity.

4. To propose a new audio steganography model that simultaneously adjusts the requirements based on the two proposed algorithms.

1.6 RESEARCH METHODOLOGY

To accomplish the development of a comprehensible methodology, this study adopts the design science research methodology (DSRM) model that was proposed by (Peffers et al. 2007) and is used in the scientific research on information systems in general. This model is customized accordingly and adopted as a research methodology in this study. The DSRM model is composed of six phases which is conducted during this research as shown in Figure 1.1. These phases are:

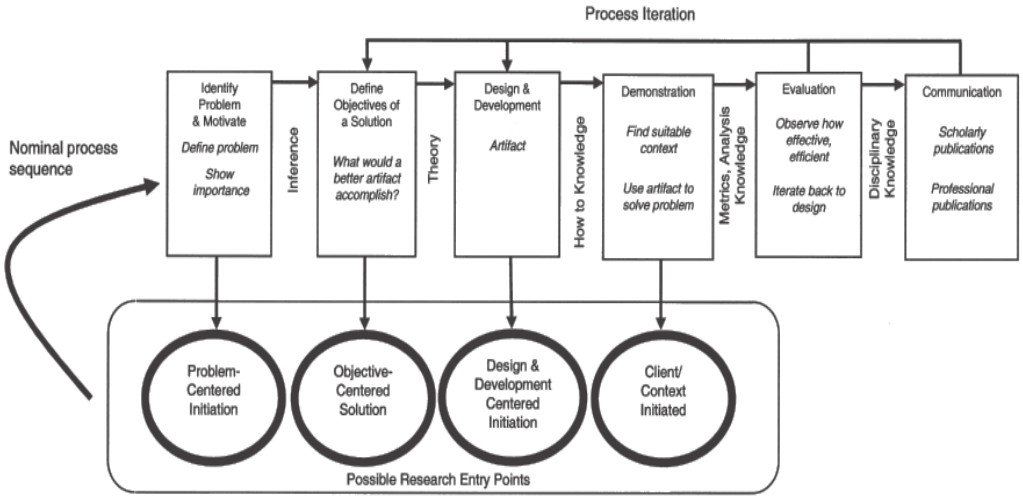

Figure 1.1 DSRM Model (Peffers et al. 2007)

- Problem identification and motivation: define the research problem specifically and highlight the significance of the solution for the problems of audio steganography.

- Define the objectives: determine the objectives of the study based on the area of study and as determined from the defined problems.

- Design and development: develop an algorithm or design model and its prototype as an artifact for the identified problems in audio steganography in order to create an effective solution.

- Demonstration: this stage includes a prototype and experiment that is used in the artifact to solve the problem identified in audio steganography.

- Evaluation: compare the result of the proposed algorithms and model and measure how this artifact supports the solution in the demonstration.

- Communication: this phase transfers the importance of the problem, the utilization of the artifact and the proper design to the researchers and other relevant audiences.

Chapter 3 discusses each of these phases in more detail with the proposed conceptual model.

1.7 RESEARCH MOTIVATION AND SCOPE

With the growth of the information and communication technology (ICT) in the last decade, several information hiding techniques have been proposed. Transmission, storage representation, retrieval and security of information are the challenges faced (Baig et al. 2016). Information hiding is the complementary approach of cryptography since many countries restrict the use of encryption in their communications. Steganography is one of the information hiding branches which is used in many applications, for instance, covert communication (Ballesteros L & Moreno A 2012; Djebbar et al. 2012).

Digital audio steganography has been established as a vital discipline in information security. That is, in part, due to the strong interest in the research community. The motivation behind developing audio steganography techniques is due to the growing need for various organizations to transmit and communicate securely as military or intelligence agencies as well as companies and organizations that provide public services to protect the confidentiality of information (El-Khamy et al. 2016).

Extensive steganography techniques were proposed in the literature using text and image as a cover compared to the audio and video files (Garcia-Hernandez et al. 2013; Atawneh & Sumari 2015). This is because of the sensitivity of the Human Auditory System (HAS) compared to Human Visual System (HVS) and this is considered another motivation for proposing audio steganography. Moreover, audio

steganography may also be used to embed medical images or information that is sent to various recipients such as doctors-in-charge of the corresponding patient (Hemalatha et al. 2016).

The accessibility and popularity of audio files make them an obvious choice for carrying hidden information. Furthermore, most steganalysis efforts are more focused on digital images leaving audio steganalysis relatively unexplored (El-Khamy et al. 2016).

In this study, a digital audio steganography model is proposed in order to improve the embedding efficiency in terms of hiding capacity and maintain transparency and robustness. The cover and the secret files used in this study are uncompressed audio samples. The model exhibited under this proposed research can be used to hide audio files in other audio files in the type of WAV with an uncompressed audio.

For the purpose of the evaluation process in this study, objective performance measures are selected to assess the performance of the proposed model such as Mean Square Error (MSE), Signal to Noise Ratio (SNR), Peak to Signal to Noise Ratio (PSNR), Normal Correlation (NC) and Hiding Capacity (HC). In addition, the robustness is evaluated under some signal processing attacks such as Additive White Gaussian Noise (AWGN), Cropping, and Echo addition. Moreover, statistical steganalysis is also tested to show the perceptual transparency against this kind of steganalysis (see section 5.2.1 for more detail).

1.8 THESIS ORGANIZATION

This thesis is organized into six chapters. Chapter 1 presents an essential introduction to the study. Chapter 2 presents a brief description of the terminology of steganography, classification, domains, and requirements. This chapter also discusses audio based steganography with human auditory system and domains. A review of the audio steganography techniques in the time and transform domains with the analysis and their limitations are also discussed which leads to a formulation of the

research questions and objectives of this study. Related concepts that can be adopted in the proposed conceptual model are presented in this chapter. Finally, this chapter is concluded with the discussion and chapter summary.

Chapter 3 describes the adopted research methodology for this study starting with the identification of the problem phase to the communication phase in order to achieve the objectives of this study, in addition to the conceptual model and the methods adopted in the proposed model.

Chapter 4 explains in detail the proposed model and its phases, which consist of four main phases; encoding, embedding, extraction and decoding phases. Each phase contains subprocesses. In addition, the demonstration section in this chapter includes the development of the proposed model and fractal audio coding into a prototype with their modules.

Chapter 5 is dedicated to conducting the experiments, dataset that is used to select the audio file and the analysis of the results for the experiments. The evaluations are performed using the common objective measurements and statistical steganalysis. This chapter is composed of five experiments for fractal coding, transparency, hiding capacity, robustness and statistical steganalysis. Results analysis for the proposed model and comparing them to the proposed schemes of the related works in time and transform domains are also discussed in this chapter.

Chapter 6 concludes the thesis by describing how the research objectives have been achieved, giving an overview of the contributions of the study to knowledge and practice perspectives, limitations of the research and suggestions for further research. The organization of thesis is illustrated in Figure 1.1.

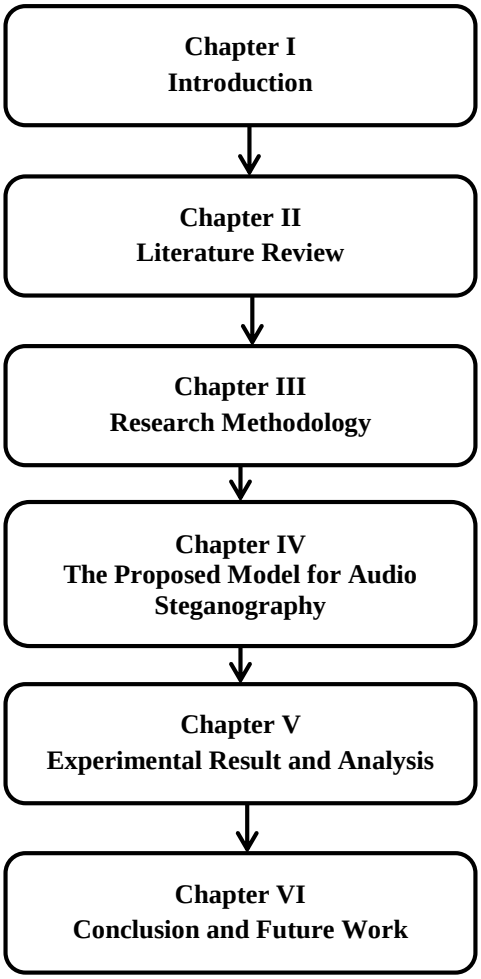

Figure 1.2 Thesis Organization

CHAPTER II

LITERATURE REVIEW

2.1 INTRODUCTION

Steganography is the science that establishes a covert communication channel between two parties and the existence of this channel should be difficult to know by the adversary (Bandyopadhyay et al. 2008).

This chapter gives an introduction to information security and its techniques by highlighting the main difference between steganography and cryptography. This chapter also elaborates on the terminology of steganography systems, cover classification, and the important requirements. Then, a brief introduction to steganalysis and its types is presented. This is followed by information on audio-based steganography and its domains adopted by the techniques applied. After that, the related works on audio steganography are presented in two domains: time and transform domains. In addition, analysis and limitation of these related works are also provided in this section. Related concepts that are adopted in the conceptual model with their techniques are presented in the next section. Finally, this chapter ends with two sections which are discussion and summary of this chapter. Figure 2.1 presents an overview of how the literature review is organized.

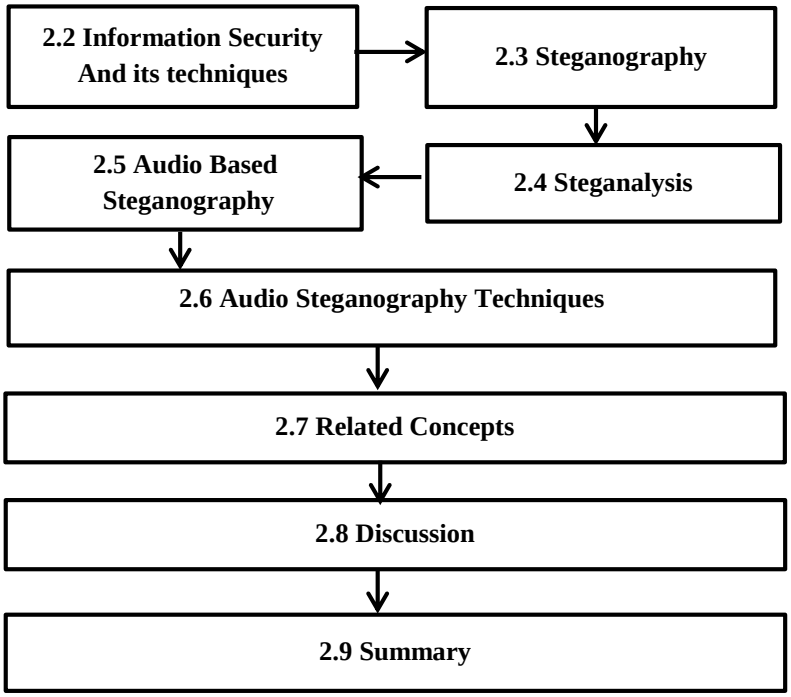

Figure 2.1 Structure of the literature review

2.2 INFORMATION SECURITY AND ITS TECHNIQUES

Information security is facing challenges nowadays with the increasing use of public or private digital data transmission over public unsecured communication system such as the Internet. This has led to a search for a tool that transmits information in a secure manner so that an intruder other than the sender and the intended receiver cannot know or even sense the presence of a hidden message, which is known as information hiding (Provos & Honeyman 2003).

For decades people have struggled to develop pioneering methods for secret communication. Generally, computer and network security have some requirements that should be met in order to ensure secure systems. Thus, in order to determine the performance of security technology, three key concepts should be considered (Von Solms & Van Niekerk 2013):

1. Confidentiality deals with protecting, detecting, and deterring the unauthorized disclosure of information. This means that the secret message should be read only by the intended recipient. This is exactly the objective of confidentiality.

2. Integrity deals with preventing, detecting, and deterring the unauthorized modification of information. An integrity attack is potentially more dangerous than a confidentiality attack.

3. Availability relates to preventing, detecting, or deterring attempted access to critical information.

Figure 2.2 shows the security concepts.

Figure 2.2 General Security Objectives

The information security has two main techniques used to secure the important data which are information hiding or data hiding and cryptography. Further, data hiding techniques can be classified into steganography and watermark as depicted in Figure 2.3. Information security techniques are interlinked. Steganography and watermark are similar in their objectives and caused confusion for the reader. Therefore, it is necessary to discuss these techniques and their aims in addition to cryptography (Cheddad et al. 2010).

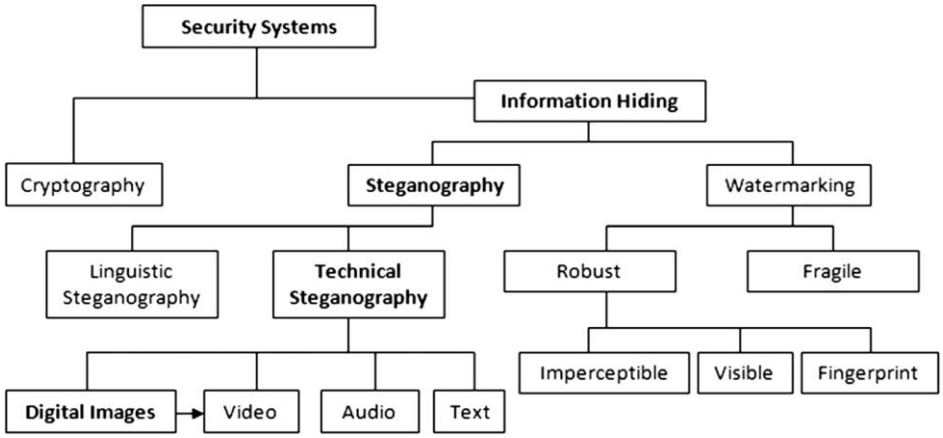

Figure 2.3 Information Security Techniques (Cheddad et al. 2010)

The aim of steganography is accomplish covert communication through embedding and extracting secret data in the cover file while watermarking is used for copyright protection or keeping the ownership by hiding the watermark in the cover file (Cheddad 2009). Cryptography is a technique utilized for data protection by scrambling the confidential data into a meaningless form (Zaidan et al. 2010). Table 2.1 gives more details about the similarity and differences between the information security techniques.

Table 2.1 Comparison of the information security techniques

Criteria	Steganography	Watermark	Cryptography
Objective	Secret Communication	Copyright/Ownership protection	Data protection
Fail when	Secret data is detected	Watermark is removed	Cipher text is De-cipherd
Carrier/ Secret	Any digital media	Any digital media	Plain text
Visibility	Invisible/Inaudible	Visible/Invisible	Visible
Requirement	Transparency and hiding capacity	Robustness	Robustness
Result	Stego-file	Watermarked-file	Cipher-text file

2.3 STEGANOGRAPHY

The term steganography relates to the fifth century through data hidden on the back of wax writing tablets or tattooed on the scalps of slaves (Johnson & Jajodia 1998). The origin of the word steganography derives from two Greek words: steganos means covert or secret, and graphy meaning writing (Nedeljko 2004). It also means the art and science of hiding secret data in an innocent looking dummy container in such a way that the existence of the embedded data is imperceptible and undetectable. Thus, steganography is the process of hiding secret data within publicly accessible information (Ali et al. 2016).

With the progress in the ability of computers, the Internet and with the development of digital signal processing, DSP, information, and coding theory, steganography has become digital. Earlier steganography approaches used images as cover media and later they were extended to adopt other multimedia like audio and

video (Hariri Mehdi et al. 2011). Many techniques have been developed for information hiding and interest is directed towards multimedia files like images or videos as a cover signal whereas fewer techniques have been developed for audio file especially with the high rate data embedding. This is most likely due to the HAS which is more sensitive compared to the HVS (Shirali-Shahreza & Manzuri-Shalmani 2008).

The main goal of steganography is the confidentiality of embedded data. Steganography hides the existence of embedded data. Therefore, unauthorized people do not sense the existence of the secret data. From a confidentiality standpoint, steganography provides a higher level of information protection than cryptography since the cryptography techniques generate an encrypted file which is unreadable and scrambled but the existence of the secret message does not hide while the steganography hides the existence of the secret message so the stego file does not give any attention.

2.3.1 Terminology

Steganography systems, as described by (Katzenbeisser & Petitcolas 2002) have a general principle which is based on a simple scenario known as the prisoner's problem. The scenario is summarized by the idea of two parties, the sender, and receiver who need a secure way to communicate by sending a secret message using an innocent cover so that a third party, the adversary, may not discover the existence of such a message. So the sender and receiver resort to using steganography to avoid detection while the adversary or the intruder works at steganalysis since she/he needs to find out whether or not the two parties communications include secretly embedded messages.

The purpose of this scenario is to prevent the intruder from observing or noticing the presence of a hidden message. On the other hand, the receiver extracts the embedded message as long as he knows the embedding method and the stego key used in the embedding process. Only the transmitter and the intended recipient should have the stego key. The security of steganography systems must be based on the

assumption that attackers have full knowledge of the steganography system design, the embedding and extracting algorithm. Therefore, most steganography systems prompt users to provide a stego key when they try to embed information in a cover file. Therefore, stego keys are usually carefully chosen and should be as strong as possible in order to prevent attackers from breaking the steganography systems using all possible stego keys (Cox et al. 2007). Figure 2.4 illustrates the general structure of steganography systems.

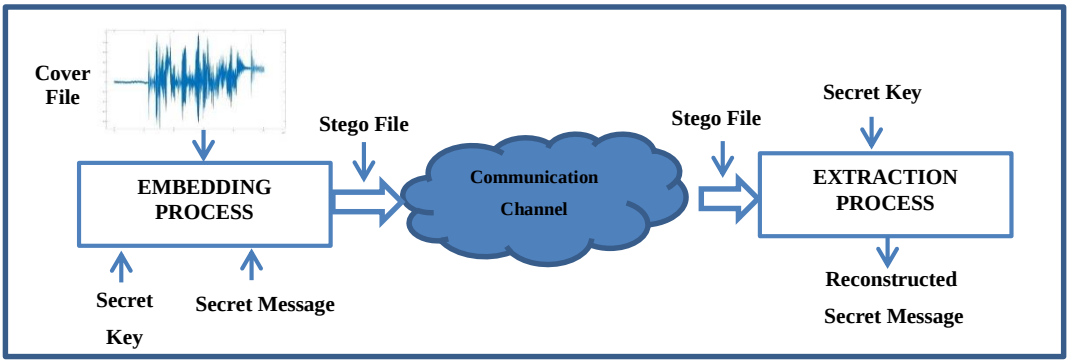

Figure 2.4 General Steganography Terminology

The main terminologies used in any steganography system are (Al-Ani et al. 2010; Djebbar et al. 2012):

1. The cover object is the cover media that is used for embedding the secret message.

2. The secret message is the information that is to be hidden in a suitable cover object producing a stego object. Note that, in this thesis, the terms message, hidden message, and secret message are used interchangeably.

3. The process of hiding information is called the embedding process which is the way or the idea that is usually used to embed the secret message in the cover object. It is also called the steganography system, steganography technique, or steganography scheme while extracting and reconstructing the secret message from the stego object is called the extraction process.

4. The secret key is a random key agreed upon between the participants, and it is usually used to control the hiding process/algorithm so as to restrict detection or recovery of the embedded data to authorized parties only.

2.3.2 Classification Based on the Cover

Steganography is used to hide confidential information in un-secure channels so that it can be transmitted and received safely to the authorized party using the file as a container or cover which is called cover file. In this era, many types of digital files are used to conceal the secret data; however, it is not possible to use all types of files or data as cover files for steganography since each cover file must have a sufficient redundant area to be replaced by the secret message (Katzenbeisser & Petitcolas 2000). Multimedia files are frequently used because of their diffusion around the Internet and thousands of them are shared daily between users (Djebbar et al. 2012). In addition, data can also be hidden or embedded using network protocols (Ali et al. 2016). In general, the cover files represent the cover for hidden data and their sizes are determined by the secret data size that is supposed to be embedded. So, cover files are the fundamental component in steganographic systems. This section exhibits a review of the covers that are used as cover files for embedding secret data in various steganography techniques which are generally classified into four categories (text, image, audio, video, and network) as shown in Figure 2.5.

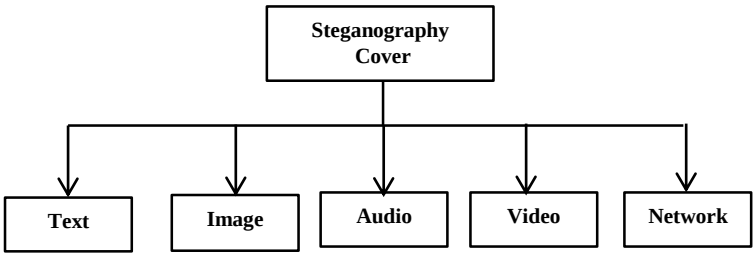

Figure 2.5 Steganography Cover

Text steganography is the process of encoding a secret message into a text file. This can be applied in different ways using the properties of the sentences like altering text format or depending on the number of words. It is the earliest method that is used

in steganography. This type of steganography is difficult to implement with secret data of a huge size due to the low redundancy of the information in the text file and thereby low hiding capacity. However, recently several text steganography articles have been published (Roy & Venkateswaran 2013; Yadav & Batham 2015; Malik et al. 2016; BAAWI et al. 2017).

On the other hand, the most common media file type used in steganography is image files. The frequent use of image files between users through the Internet makes it the most chosen for hiding secret information, in addition to their low sensitivity towards the HVS and the redundancy of information inside the image file (Cheddad et al. 2010). The hiding process is carried out through a slight alteration of the visible properties of the image file in order to reduce the awareness of the presence of the secret data. The earliest methods that were used in embedding involved using the Least Significant Bit or LSB, echo hiding and spread spectrum. These techniques can be achieved with disparate degrees of success for different kinds of image files (Thenmozhi & Chandrasekaran 2012; Hemalatha et al. 2013).

The third category is video steganography. The video is generally a combination of audio and image files. Therefore, most image steganography techniques can be applied to video files. A large amount of information can be concealed in the video file and changes might not be observed because of a continuous flow of information. This is the main advantage of using video steganography (Ramalingam & Isa 2014; Suresh & Pathasarathy 2016).

The audio steganography approach uses the audio file as a cover file in embedding which is not simple and can be considered a challenge due to the sensitivity of the HAS. Thus, human ear senses the variations in an audio file over a range of power greater than one billion to one and range of frequencies greater than one thousand to one. However, it has holes and a large dynamic range can contribute to data hiding (Bender et al. 1996). Audio files make convenient cover files for hiding because of their high level of redundancy and high data transmission rate in comparison with other multimedia files. Several techniques are discussed later.

With the growth of network protocol and services, the researchers direct their interest towards adopting techniques for the real-time cover media to increase security and resist difficulties from data eavesdropping, which is called Network Steganography (Lubacz et al. 2014). Voice over Internet Protocol (VoIP) service has taken most of the attention because of the additional security that is provided because these techniques are real time and do not give the warden adequate time to discover the secret data. Moreover, stream data is a suitable cover because of the dynamic and variable length for hiding the data which is relatively larger than static files (Mazurczyk 2013; Ali et al. 2016).

2.3.3 Data Hiding Domains

In general and whatever the cover file, data hiding techniques can be classified into three different domains: time or spatial, transform and compressed (Chang & Nguyen 2014; Ali et al. 2016; Sun 2016). Each of them has some pros and cons. Figure 2.6 presents these domains.

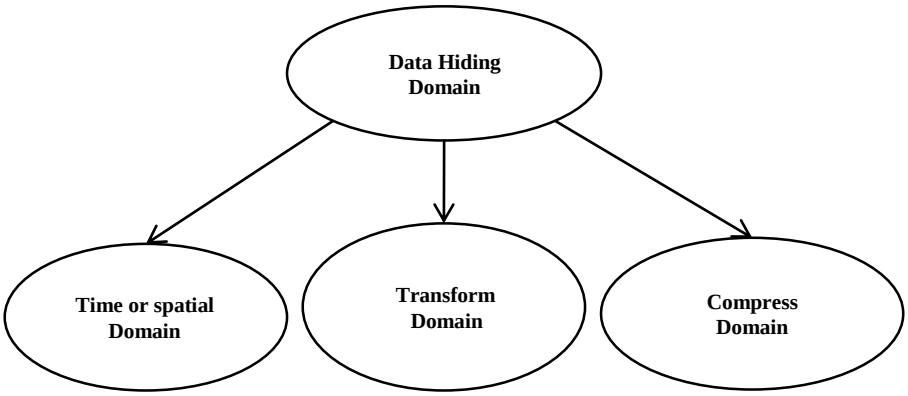

Figure 2.6 Data hiding Domains

In the time or spatial domain, the secret message is embedded directly into the cover file because it is simple to implement and offers high hiding capacity but it has low robustness and perceptibility such as conventional LSB (Divya & Reddy 2012). The techniques adopt this domain include methods have features of less computation complexity and they can be implemented in real-time applications. Figure 2.7 shows the embedding process in this domain.

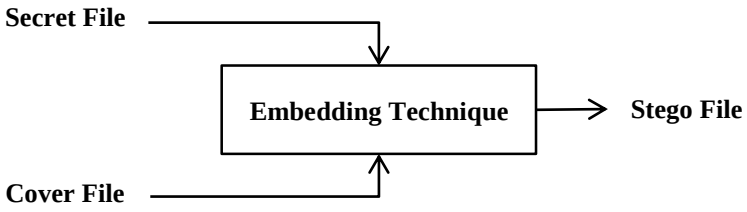

Figure 2.7 Data hiding in time domain

In the transform domain, the cover file is transformed into another form that is used in the embedding process which overcomes the robustness that tolerates noises and some signal processing operations in addition to the issue of transparency that is found in the time domain. On the other hand, it suffers from computational complexity and low hiding capacity like Fast Fourier Transform (FFT), Discrete Cosine Transform (DCT), Spread Spectrum (SS) and Discrete Wavelet Transform (DWT) (George & Mahmood 2010). These techniques are difficult to adapt for real-time applications due to latency results from complexity. Figure 2.8 shows the data hiding process in the transform domain.

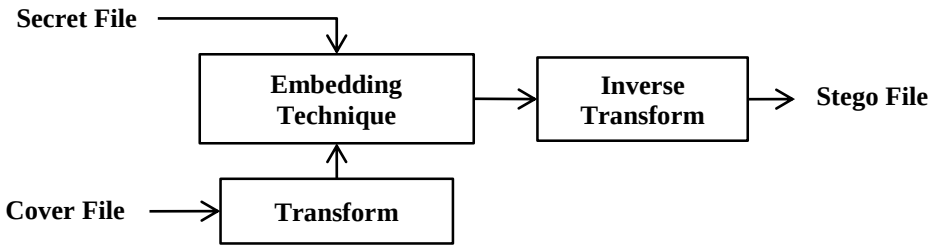

Figure 2.8 Data hiding in transform domain

In the compressed domain, several techniques have been developed and adopted in image steganography (Shie & Jiang 2012; Lee et al. 2013; Chang & Nguyen 2014). In this domain, the cover image or the secret image is compressed using different compression techniques to develop steganography techniques and produce high capacity and compression ratio. Vector Quantization (VQ) is a data compression technique that consists of three components: encoder, decoder, and codebook. The codebook must be shared between sender and receiver. VQ is normally used for hiding secret data in the compressed cover file by dividing the secret image into blocks. Each block is then mapped to the most similar codebook and the

codebook index is used instead of the block samples. Figure 2.9 shows the embedding process in the compressed domain. However, the proposed techniques in audio steganography did not discuss this domain as shown in Section 2.5.2.

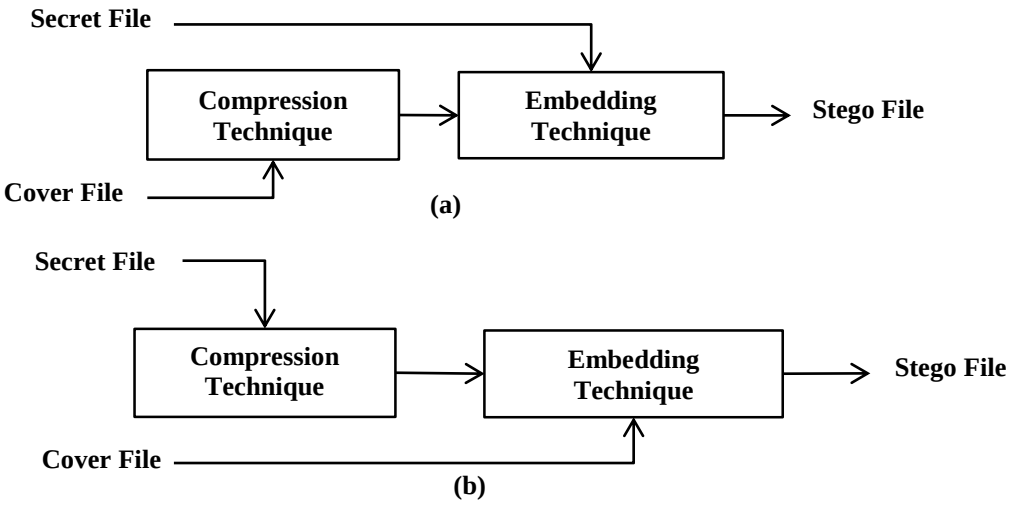

Figure 2.9 Data hiding in the compressed domain, (a) compress the cover file, (b) compress the secret file.

2.3.4 Requirements of the Steganographic System

The evaluation of any steganographic system depends on several requirements or factors. These requirements should be considered in the designing and assessment of the steganography system. Basically, steganographic systems have three fundamental characteristics which must be investigated which are transparency, hiding capacity and robustness (Fridrich 2009; Ballesteros L & Moreno A 2012). These requirements are mostly related and contradictory to each other. Attaining them for one scheme remains a challenging task (Shahadi et al. 2014; Tang et al. 2016).

a. Transparency

Transparency, undetectability or imperceptibility is the ability to embed data without perceptually degrading the original cover. It could be expressed as a measure of how similar perceptually and statistically the stego is to the original cover file. On the other hand, the term security is usually equivalent to transparency. Steganographic systems can be considered secure if it is impossible for attackers to detect the presence of

hidden data in the stego files by using any accessible means (Cox et al. 2007). In this study, the term transparency is used.

Before going through the discussion of the evaluation of this requirement, it is necessary to distinguish between two common terms which are interchangeably used in steganography articles when referring to stego transparency, which are fidelity and quality. On the other hand, in the signal processing field, the perceptible measurements of the signals are classified into these two types.

Fidelity means the perceptual similarity between two files before and after processing, while the quality is a measure of the goodness of an audio or image as perceived by the human eye/ear. For example, audio with noise or hissing is naturally considered to be of low quality. A stego file has high fidelity when it is indistinguishable from the original cover file; however, it may have a slightly lower quality. So, fidelity is used when talking about transparency.

Evaluating the fidelity of a stego file is a significant requirement in the performance of the steganography algorithm (El-Khamy et al. 2016; Hemalatha et al. 2016). Generally, there are two ways to measure stego quality: objective and subjective measurement methods (Mat Kiah et al. 2011). The objective methods are based on measurements automatically computed using the stego data, while subjective methods are based on human observer judgment. In practice, subjective evaluation is usually time consuming and expensive. The goal of the objective method is to develop quantitative measures that can assess the fidelity of the stego file. Three objective metrics are used in the assessment of stego fidelity, MSE, SNR, and PSNR. These metrics are discussed in more detail in Section 5.2.1a. The transparency of the cover file using SNR of more the 20 dB and PSNR with more than 30 dB indicates acceptable (Al-Haj 2014; Mosleh et al. 2016; Renza & Lemus 2018).

b. Hiding Capacity

Generally, hiding capacity is the amount of secret information that can be successfully hidden in the cover without introducing any perceptual distortion for the cover.

However, there are two overlapping terms in the data hiding which are hiding capacity and payload or embedding rate. Hiding capacity denotes the ratio of secret data size to the size of the original audio file and it is measured by hundred percentages, while payload represents the number of secret bits that can be hidden during a unit of time and it is measured by bit per second (bps). More details are given in Chapter 5, Section 5.2.1b (Ali et al. 2016; Zaidan et al. 2016).

Steganographic systems are mainly used for secret communication and aim to maximize the steganographic capacity and maintain the perceptual fidelity of the cover (Mat Kiah et al. 2011). However, the steganographic capacity tends to be restricted by the size of cover files. Therefore, developing a steganography technique should take into consideration how to increase the amount of secret data that can be hidden without affecting the properties of stego files (Malik et al. 2016; Tang et al. 2016). Hiding capacity more than 50% and embedding rate more than 50 bps are considered high (Ballesteros L & Moreno A 2012; Lei et al. 2012; Hemalatha et al. 2016).

c. Robustness

The significance of this factor differs between the steganography and watermarking systems. Robustness is the main requirement in the watermarking systems whereas it is considered secondary in steganography techniques (Petitcolas et al. 1999; Ballesteros L & Moreno A 2012). The system is robust if the secret data is still detectable or recoverable after intentional attacks to reveal the secret message and signal processing operations. Thus, robustness represents the ability of secret data to withstand such kinds of attack. Moreover, it entails the survivability of this watermark against all kinds of signal processing and transformations (e.g. scaling, rotating, sharpening, filtering, adding noise, and blurring).

On the other hand, robustness is not an issue or a top priority for steganographic systems. The design of most steganographic systems does not consider robustness as a fundamental requirement since the majority of these systems assume the passive warden scenario (Cox et al. 2007; Fridrich 2009). Hence, steganographic

systems are either not robust against modifications or have limited robustness against some of the signal processing. However, watermarking systems must be robust and resist any kind of transformations or manipulations that may attempt to remove the watermark. Sections 5.2.1c and 5.2.1d show the metrics that are used in the evaluation of robustness.

d. The Tradeoff Between Requirements

As mentioned in Chapter 1, Section 1.3, the aim of steganography is to increase steganographic capacity and maintain the transparency. However, these requirements have an inverse relationship. The hiding capacity and transparency are contradictory (Ballesteros L & Moreno A 2012). The more data is hidden in the cover (high hiding capacity), the more distortion is caused to it.

On the other hand, it is difficult to simultaneously enhance robustness and high hiding capacity in particular steganography systems (Bender et al. 1996; Djebbar et al. 2012). Consequently, any steganographic system should achieve a better tradeoff among these requirements. Figure 2.10 shows the contradictory relationship between the data hiding requirements.

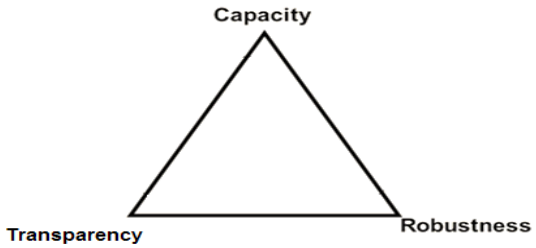

Figure 2.10 Magic triangle of the data hiding requirement (Sun 2016)

The challenge in designing and developing steganography systems is the tradeoff between the requirements and how to adjust or maintain them simultaneously (Bhat et al. 2010; Lei et al. 2012).

2.4 STEGANALYSIS

Steganography is the art of hiding a secret object in cover media, while Steganalysis is the art of discovering the secret object from the cover media (Abdelfattah & Mahmood 2013). With an increased emphasis on security, both steganography and steganalysis have recently drawn research attention since the when steganography techniques are increased, interesting in developing steganalysis techniques are increased as well. While it is relatively easy to embed a secret message in media such as an image, audio or video, the detection of an embedded message i.e., steganalysis is challenging because of the many different methods used in steganography and the continuous evolution of new steganography algorithms. The purpose of steganalysis is to identify if the cover (image, text, audio or video) has been manipulated by embedding a secret message using some embedding technique. The different cover types and many possible embedding techniques introduce great complexity in designing a reliable steganalyzer (Nissar & Mir 2010).

The steganalysis techniques are classified by the researchers into two types of the categories. Abdelfattah and Mahmood (Abdelfattah & Mahmood 2013) classified steganalysis techniques into specific and universal steganalysis. The specific steganalysis techniques are designed for a targeted embedding technique. Thus, they yield very accurate decisions when they are used against a particular steganographic technique. In universal techniques, dependency on the behavior of the individual embedding techniques is removed by determining and collecting a set of distinguishing statistics that are sensitive to a wide variety of embedding operations.

The second classification discussed by Nissar and Mir (Nissar & Mir 2010) is based on whether the signature or the statistics steganalysis technique is being used and furthermore concerns specific and universal approaches.

Signature steganalysis utilizes the alterations of the media properties due to the hiding process that may introduce some form of degradation or unusual characteristics and patterns. These patterns and characteristics may act as signatures that broadcast the existence of an embedded message.

Statistical steganalysis, as the name implies, analyses the underlying statistics of the cover file to detect the embedded secret information. Statistical steganalysis is considered more powerful than signature steganalysis because it depends on mathematical techniques which are more sensitive than visual perception.

Although significant research efforts are reported in the domain of image steganalysis, fewer efforts are reported in the field of audio steganalysis. This might be attributed to the nature of audio file which is different from the image file, the fact that poses challenges in obtaining the statistical features of audio (Kraetzer & Dittmann 2007; Meghanathan & Nayak 2010; El-Khamy et al. 2016).

2.5 AUDIO BASED STEGANOGRAPHY

Audio steganography is the scheme of concealing confidential data in a cover of audio file so that no one other than the sender and intended receiver is able to notice that secret data is hidden inside (Antony et al. 2012). Embedding secret data in the cover audio file is more challenging than using images since the HAS is more sensitive to small changes in the audio samples in comparison to the sensitivity of HVS to the pixel modification (Hemalatha et al. 2016). On the other hand, in comparison to other types of file, audio files are larger in size than other cover files, there is a high level of redundancy and a high data transmission rate and these factors make it more suitable as a cover file (Nosrati et al. 2012).

2.5.1 Human Auditory System

Since the dynamic range of the HAS is wider in comparison to the HVS, hiding data in audio is more challenging in comparison with hiding data in images or video (Nedeljko et al., 2004; Zhang et al., 2007). Digital audio technologies like the development of MPEG rely on the detailed knowledge of the HAS. The relevant information is not only limited by the ability of the ear to hear frequencies in a band between 20 Hz and 20 kHz and the dynamic range of over 96 dB, but the physical phenomenon of different frequencies and their corresponding process in HAS is also important for considering in depth the correlation between hearing sensations and

acoustic stimuli (Michael et al., 2003). One phenomenon is that the sensitivity of the human ear to changes in louder audio output is greater than when compared to quieter audio output (Andres et al., 2002; Yong et al., 2009).

The HAS operates over a wide dynamic range. When using digital images as cover files the difficulty of the human eye to distinguish colors is taken advantage of. Similarly, when using digital audio one can count on the different level of sensitivity of the human ear when it comes to sounds of low and high intensity. Usually, higher sounds are perceived with better fidelity than lower sounds, thus it is recommended for hiding the secret message without being noticed by the human ear (Athawale 2010).

2.5.2 Audio Steganography Domains

Although the data hiding technique is classified into three domains as shown in Section 2.3.3, audio steganography techniques are mainly classified into two domains, time and transform domains, depending on the related works in the literature shown in Section 2.6 (Shahadi et al. 2014; Hemalatha et al. 2016). However, some authors classify the transform domain into frequency and wavelet domain (Shahadi et al. 2015; El-Khamy et al. 2016).

In the time domain, the cover sample values are modified to embed the secret data. These techniques provide a very simple way to hide secret data, easy to implement, high hiding capacity and transparency. However, these techniques suffer from weakness in terms of withstanding signal processing like compression, filtering and noise attacks. LSB technique is the best example of a time domain technique. Other audio steganography techniques that come under this category are echo hiding and silent interval (Djebbar et al. 2011). The robustness of the methods in this domain is weak since any change of the values of the audio samples, the secret message is destroyed and cannot retrieve. With low payload capacity, these techniques may tolerate the noise addition and robustness (Djebbar et al. 2012).

When the cover object is transformed from the time to the frequency domain, the transformed coefficients are used to conceal the secret data. The embedding technique, in this case, can be classified under transform domain techniques. Techniques adopting this domain have the advantage that these techniques can withstand signal processing operations so that the secret information can be retrieved even if the stego object undergoes signal processing operations. Audio steganography approaches that applied transform domain give better results than time domain techniques in terms of transparency and robustness, however, the achieved capacity is low (Shahadi et al. 2014; El-Khamy et al. 2016; Hemalatha et al. 2016).Certain characteristics of the HAS must be exploited for hiding information effectively (Al-Haj 2014). DCT, DWT, and SS are examples for a transform domain (Djebbar et al. 2012).

2.6 AUDIO STEGANOGRAPHY TECHNIQUES

This section is dedicated to presenting the proposed schemes in related works in audio steganography. Most of these techniques focus on improving the important requirements of the audio data hiding which are the hiding capacity, transparency, and robustness. Reviewing the recent solution helps to avoid the drawbacks of these solutions and find the vulnerabilities in these techniques.

The aim of this study is to improve the overall performance of audio steganography. For this reason, it is crucial to have an overview of efforts that have been made to achieve this aim. Therefore, the next two subsections discuss the schemes in the related works that have contributed to enhancing the requirements mentioned earlier in Section 2.3.4. The discussion can be separated into two main parts based on the domains adopted by the proposed schemes which are time and transform domains as shown in Section 2.5.2.

2.6.1 Techniques Adopting the Time Domain

As mentioned in Section 2.5.2, the techniques that adopt this domain embed the secret data directly into the cover samples without any transformation. This provides

simplicity in the embedding process, high hiding capacity and fewer cover sample distortions. However, these techniques are considered as less robust when compared to others since any modification in audio samples, the secret message cannot retrieve (Ballesteros L & Moreno A 2013; Hemalatha et al. 2016). Table 2.2 sums up these techniques. LSB has been widely adopted in literature to propose audio hiding techniques as it is simple to implement and it provides higher transparency. Techniques that contribute to enhancing the performance of audio steganographic systems in this particular domain are reviewed in this section and the summary of these related works is presented in Table 2.2.

(Delforouzi & Pooyan 2009) proposed a method that composed of compression, encryption and LSB insertion techniques. The objective is to enhance the security, stego quality, and the payload. The method first compresses the secret message using WINRAR or ZIP lossless software then encrypts the compressed message using a pseudo-random number then the resulting data is adaptively hidden in the LSB of the audio samples. Less computational, blind recovery, SNR 52 dB on average and 203.6 Kbps on average payload is the output of the proposed method. The hiding capacity is, on average, 28.8% of the cover size since the used audio files are mono, 16 bits per sample with a 44,1kHz sample rate. However, the hiding capacity of the proposed method is low.

A method that uses two approaches, noise gate software logic with 8th standard LSB embedding was developed by (Ahmed et al. 2010) to improve the hiding capacity and the robustness of the audio data hiding. The first approach is used to select samples for embedding in order to avoid hiding in the silent periods as it can be detected by HAS and the second is 8^{th} LSB technique for embedding in the cover audio file. The subjective and objective test show increases in the SNR and SDG compares with standard LSB. However, there is no test for the robustness and the authors did not mention the hiding capacity of the method.

Bit modification algorithm is proposed by (Gopalan & Shi 2010) to make a tradeoff between data robustness, payload, and transparency. The quality of the stego file depends on the modified bit. The results show that the payload or the hiding

capacity is 3000 bits per second when the cover audio is 8 kHz and 8 bps, which means about 4.4 % of the cover file size. The fidelity is measured objectively by SNR and Bark Spectral Distortion (BSD) which is 39 dB and 0.1, respectively. The Bit Error Rate (BER) is less than 1. Moreover, it is also noted from the result that the hiding capacity is low and they did not mention the value of the added noise with the result of the BER.

(Kekre et al. 2010) proposed two approaches for embedding using multiple LSB in the audio file to improve the hiding capacity. These techniques use different numbers of LSB bits depending on the most significant bit (MSB) of the cover audio sample. One approach checks one MSB to check and embed in 6 or 7 LSBs and the other checks two MSBs to embed in 4, 5, 6, and 7 LSBs. The authors used a 16-bit cover sample for embedding their approaches. The results showed that the obtained hiding capacity is between 35% and 70% of the cover file size. The signal to noise ratio (SNR) of the stego file is 52 dB on average. Moreover, no discussion of the security of the proposed approaches is presented.

AES Encryption with a random bit and sample selection technique is proposed by (Asad et al. 2011). This technique is used to enhance the security of conventional LSB method. Firstly, two and three MSBs of the cover samples are checked to select the bit and sample for embedding, respectively. However, the technique hides one bit per 8 samples so the hiding capacity is low at about 12.5% of the cover file size and there is no mention of the fidelity of the proposed technique in terms of SNR. The selection criterion is weak since it cannot be considered as random depending on the MSB.

(Darsana & Vijayan 2011) proposed a scheme using two methods, modified LSB using 4^{th} bit and flip the others and Pixel Value Differencing (PVD). The scheme aims to reduce the hiding distortion of the cover, increase the robustness and increase the hiding capacity. The test results show that the SNR is preserved around 40 dB while the hiding capacity of the PVD is higher at about 20% of the cover size. The hiding capacity is still low; moreover, no test for the security of the scheme is discussed.

A genetic with multiple LSB approach is presented by (Zamani et al. 2011) to solve two problems of LSB substitution technique for audio steganography. The low robustness against attacks which try to reveal the hidden message is the first issue, while the second is the low robustness against distortions signal processing. The proposed algorithm embeds the secret bits in selected samples in deeper layers of the cover samples and alters other bits to decrease the error. The hiding capacity is 2 bits per byte; it means 25% of the cover size since the samples rate is 44.1 kHz and 16 bps. The results show that the SNR of the obtained stego file is between 73 and 98 dB depending on the size of the secret message and cover file. However, there is no result shown for the robustness of the proposed approach.

An audio in audio steganography method is proposed by (Sakthisudhan et al. 2012). In this method, the secret audio is encrypted using logical operation between the secret audio bits and pseudo sequence and the encrypted sequence is embedded in the audio cover samples using a standard LSB technique. The limitation of this study is that the size of the cover audio file should be eight times larger than the secret audio file thus the hiding capacity of this proposed method is 12.5% of the cover file size. There is no result for the fidelity of the stego file.

(Rahim et al. 2014) proposed a technique to increase the hiding capacity with minimum cover modification. The technique hides the secret data in dual bits in the 4^{th} and 1^{st} LSB and makes a modification to the other bits of the same sample. The achieved hiding capacity in this technique is between 2.87 and 4.76% of the cover file size with SNR between 58.2 and 60.5 dB. However, the hiding capacity is low and security is not discussed.

(Bhattacharjee et al. 2015) presented a technique for enhancing data confidentiality and portability. The technique integrates a fixed length coding lossless compression to reduce the size of the secret text file size and 2 LSB (1^{st} and 4^{th} with modification of the other bits) embedding approaches. The results exhibit a hiding capacity of 7.26% of the cover file size with high stego fidelity SNR around 60 dB. However, hiding capacity is low and they did not discuss the robustness against intentional attack.

An audio steganography technique is proposed by (Datta et al. 2015) to enhance the hiding capacity and robustness that hides text in audio cover. It used a modular operator to increase the capacity and security. It used the hexadecimal equivalent of the secret string and modulo operator for hiding. The secret text is converted to ASCII then groups every 4 bits forming a hexadecimal that is compared with the remainder of the division of amplitude samples of the cover audio by 16 and adjusts the amplitude of the cover samples so that the remainder is equal to the hexadecimal digit. This technique reduces the number of samples required to hide secret data since every 4 bits is converted to a hexadecimal value and hide one time instead of in four audio samples, so the hiding capacity increases but it is not mentioned explicitly. However, the quality of the stego file compared with two other techniques mentioned in the article is less in terms of SNR value and the selected samples for embedding are done in a sequential manner; this leads to being exposed by the adversary. No test result for robustness was conducted.

(Bazyar & Sudirman 2015) proposed an embedding technique for increasing hiding capacity. The technique used variable and multiple LSB bits up to 7 LSBs depending on the two MSBs of the cover sample. The results showed that the obtained hiding capacity is between 35% and 55%, and the SNR of the stego file is 62 dB on average using 4 LSBs. Security against intentional attack was not discussed for this technique.

(Kar & Mulkey 2015) proposed a scheme based on multi-threshold error criterion with chaotic model distribution to enhance the transparency, the security of the audio steganography and evade detection by statistical analysis. The hiding capacity of this scheme depends on the threshold criterion. Increasing the threshold implies more bits to embed in a higher position and less error bit rate; however, there is a limitation since increasing threshold more can introduce noticeable distortion in the stego audio file. They used only subjective metrics in their tests, namely Mean Opinion Scores (MOS). The results present that the proposed scheme produces better MOS for stego audio files and yields large capacity with similar stego quality when compared with similar schemes. The higher hiding capacity of the scheme is 27.9% of the cover file size. There is no result for fidelity of the stego files.

(Chowdhury et al. 2016) proposed an approach to enhance the security of the conventional LSB based on dual encryption methodology. Firstly, the secret text data is encrypted using a pattern matching algorithm and the second factor is the encrypted position for embedding based on the binary sequence of the cover audio data. They compared their results with the results of the conventional LSB in terms of SNR and MSE and the results show slight decreases in the SNR of the cover file while ensuring the integrity and the quality of the secret message. However, the approach did not have a secret key and the authors did not mention the hiding capacity of the proposed approach.

Table 2.2 Summary of the related works in time domain

Author/year	Method/Technique	Secret data	Finding(s)	limitations
(Delforouzi & Pooyan 2009)	Compression software, Encryption, and LSB techniques	Message	▪ Enhance the security ▪ Stego quality SNR 52 dB on average ▪ Payload 203.6 Kbps on average and the hiding capacity is on average 28.8%	▪ Hiding capacity is low.
(Ahmed et al. 2010)	Noise gate software logic with 8th standard LSB embedding	Message	▪ Enhance the perceptibility of the cover in terms of SNR and SDG ▪ Enhance the robustness	▪ There is no test for the robustness ▪ The hiding capacity did not mention
(Gopalan & Shi 2010)	Bit modification algorithm	Message	▪ Payload is 3K bits per second, about 4.4 % hiding capacity ▪ The fidelity is 0,1 BSD, SNR is 39 dB ▪ Robustness, BER is less than 1	▪ They did not mention the value of the added noise with the result of the BER
(Kekre et al. 2010)	Multiple LSB	Audio	▪ Hiding capacity is between 35% and 70% ▪ SNR of the stego file is 52 dB on average	No discussion for the security of the proposed approaches is presented
(Asad et al. 2011)	AES Encryption, LSB with random bit and sample selection	Message	▪ Enhance the security of conventional LSB	▪ No results mentioned about the fidelity of the stego file in terms of SNR ▪ The selection criterion is weak

to be continued…

…continuation

				since it cannot be considered as random depending on the MSB
(Darsana & Vijayan 2011)	Two methods, modified LSB and Pixel Value Differencing (PVD)	Text	▪ Enhance the robustness ▪ The SNR is preserved around 40 dB ▪ The hiding capacity of the PVD is about 20% of the cover file.	▪ The hiding capacity still low ▪ Not test for the security of the scheme is discussed
(Sakthisudhan et al. 2012)	LSB, Encryption using pseudo sequence with logical operation	Audio	▪ The hiding capacity is 12.5% of the cover size	▪ There is no mention about the fidelity of the stego file
(Zamani et al. 2011)	A genetic with multiple LSB approach	Message	▪ Robustness ▪ Hiding capacity is 2 bits per sample, 25% of the cover size ▪ PSNR of the obtained stego file is between 73 to 98 dB depending on the size of the secret message and cover file	▪ No result for robustness test
(Rahim et al. 2014)	Dual LSB 1^{th} and 4^{th} with modification other bits for less distortion	Text	▪ The achieved hiding capacity is between 2.87 and 4.76% ▪ SNR between 58.2 and 60.5 dB	▪ Low hiding capacity ▪ The security was not discussed
(Bhattacharjee et al. 2015)	Fixed length coding lossless compression and 2 LSB embedding approaches	Text	▪ Hiding capacity 7.26% ▪ SNR around 60 dB	▪ Hiding capacity is low ▪ They did not mention about the security of the technique
(Datta et al. 2015)	Modulo operator	Text	▪ Increase the hiding capacity and robustness	▪ Hiding capacity not mentioned explicitly ▪ Less security since the cover samples for embedding are selected in a sequential manner ▪ No test result for robustness

to be continued…

…continuation

(Bazyar & Sudirman 2015)	Variable and multiple LSB	Message	▪ Hiding capacity between 35% and 55% ▪ SNR 62 dB on average	▪ Security weak.
(Kar & Mulkey 2015)	Multi-threshold, error criterion, chaotic model distribution and LSB	Text	▪ Enhance the transparency and evade the detection by statistical analysis ▪ Enhance security of the audio steganography	▪ The hiding capacity is 27.9% ▪ There is no result for the objective test as SNR or PSNR.
(Chowdhury et al. 2016)	Dual encryption, One for encrypt secret key and second for position for hiding, LSB	Text	▪ Enhance the security ▪ The average SNR is 52 dB.	▪ Approach did not have secret key ▪ Not mentioned about the hiding capacity of the proposed approach

2.6.2 Techniques Adopting the Transform Domain

DWT and DCT are the common techniques adopted in the transform domain because of their capability to increase the transparency of the stego file and robustness of steganography systems (Ballesteros L & Moreno A 2012; Shahadi et al. 2014; El-Khamy et al. 2016; Kanhe & Aghila 2016; Renza et al. 2017). Several other methods are proposed in this particular domain. Table 2.3 summarizes the related works in the transform domain in terms of method or the techniques that were adopted, the secret message used, findings and limitations.

(Pooyan & Delforouzi 2007) proposed a method based on LSB with a 5 level lifting wavelet transform and hearing threshold to increase the hiding capacity and stego quality. The method also uses pseudo number sequence to enhance the security of the proposed method. The achieved payload is 239793.75 bps and the hiding capacity is 33.9%. SNR and MOS are used to evaluate the stego quality which is 37.8 dB and 4.8 on average, respectively. However, there are no test results for the robustness of the proposed method.

(Shirali-Shahreza & Manzuri-Shalmani 2007) introduced a method using adaptive lifting Int2Int wavelet scheme and LSB. The aim of the proposed method is to increase the hiding capacity, quality of the stego audio and low error rate.The

number of secret bits to be hidden depends on the value of the coefficient and they are then embedded in the LSB of the details coefficients. Results show payload up to 200 kbps using Haar wavelet, 28.3% of the cover hiding capacity, 0% error rate with a payload of less than 100 kbps and SNR is 42 dB on average.

(Delforouzi & Pooyan 2008) presented LSB-based audio hiding using a wavelet domain method. It uses one level Haar wavelet transform and the pseudo number generated with a hearing threshold to increase the payload, transparency of stego audio and the security. The payload of this method is 296352 bps, 42% of the cover file size hiding capacity with SNR equal to 40.2 dB and MOS equal to 4.8 stego fidelity.

(Shahadi & Jidin 2011) proposed an algorithm based on discrete wavelet packet transform (DWPT) with adaptive hiding based on LSB. In this algorithm, the strength of cover samples and the matching of bit blocks are the two factors that affect the hiding process. The results showed that the payload is 300 kbps; embedding capacity is up to 42% of the cover file size with SNR of 50 dB stego quality. Robustness against secret message revealing was not discussed.

(Sheikhan et al. 2011) proposed a method for hiding information in DWT using the LSB substitution technique. The cover signal is divided into several sub-bands using DWT. The sub-band with lower than or equal energy to the hearing threshold is used in the embedding process. The result exhibits stego quality in SNR as 76 dB and MOS as more than 4.7, and the hiding capacity is 34.5 % of the cover file size on average. Robustness against secret message revealing was not discussed.

(Rekik et al. 2012) proposed a new method using DWT and FFT to enhance the security of speech communication. The method is composed of hiding the secret speech parameters in the high-frequency regions of the magnitude spectrum of the cover speech. The experiments used NOIZES database which consists of 15 male and 15 female audio signals with a sample rate of 25 kHz. The proposed method shows that the stego file renders the steganalysis. The results show that the quality of the stego file in terms of SNR and PESQ is 22 dB and 3.5, respectively. There is no

mention of the hiding capacity and no experiments conducted to show the robustness of the proposed method against steganalysis and intentional attack.

(Ballesteros L & Moreno A 2012) proposed a model named Efficient Wavelet Masking (EWM) by adopting efficient sorting of the wavelet coefficients of the secret signal and indirect LSB substitution. This model provides higher statistical transparency compared to LSB, Frequency Masking (FM), Spread Spectrum (SS) and Shift Spectrum Algorithm (SSA). The model is assessed by statistical analysis using the fourth first moment of average (μ), variance (σ), skewness (s), and kurtosis (k) in the time, frequency and wavelet domain. The results show the model is stable since the maximum differences of the statistics of the cover and stego file are still less than 10%. The best hiding capacity with higher statistical transparency is between 25 and 33% of the cover size. There is neither a result for objective measurements such as SNR and PSNR nor was robustness against attack and signal processing conducted for the proposed method.

(Ahmed & George 2013) proposed a method for hiding secret data into voiced areas of the audio file. The proposed method adopts Haar DWT and Uniform coefficients Modulation (UCM) based on the uniform scalar quantization in the embedding and extraction process. The achieved hiding capacity is low of 5% and the fidelity of the stego file PSNR is 42 dB. The embedding and extraction process is simple, However, there is a lack in robustness (security) since the secret data embeds into coefficients of the cover samples in a sequential way and the method did not evaluate using steganalysis techniques.

(Shahadi et al. 2014) presented an approach based on an Int2Int lifting wavelet transform, LSB and encryption techniques to increase the stego transparency, robustness, and security of the audio steganography system. The achieved hiding capacity is 25% of the cover file size and stego quality is above 45 dB. Moreover, robustness against AWGN up to 60 dB with NC equal to 0.9618 and Difference Ratio (DR) of the statistical steganalysis less than 0.7 for the first fourth moments are the results of the conducted experiments for the proposed approach. However, the achieved hiding capacity is low.

(Shahadi et al. 2014) proposed a scheme for high hiding capacity, transparency and security using an Int2Int lifting wavelet, Weighted Block Matching (WBM) and adaptive position embedding techniques. The obtained payload and hiding capacity are around 340 kbps and 48%, respectively. Moreover, the transparency in terms of SNR is above 35 dB with AWGN robustness of up to 60 with NC=0.9944.

(Shivdas 2014) proposed a method in DCT domain based on sample comparison and encryption using a pseudo number and XOR operator. This method is used to hide text or audio in the audio cover. The hiding capacity is up to 25% of the cover file size and the SNR of the stego file is within 50 dB. The hiding capacity is low and a robustness test was not conducted.

(Verma et al. 2014) proposed a scheme based on sample comparison in the DWT domain using Haar wavelets and encryption using a XOR operator for high capacity and stego quality. The audio files used in the test were 8 bps, mono, and 11.25 kHz sampling rate. The findings for this article showed an embedding rate of up to 4 kbps that is above 25% of the cover file size and stego quality SNR above 35 dB. Robustness testing was not discussed.

Haar DWT and a dynamic encryption algorithm were used in the presented technique by (Hemalatha et al. 2015) to enhance the robustness and the security of the audio steganographic system. The technique hides encrypted text using the number of secret characters in the detailed coefficients of the audio cover. The proposed technique was tested using objective measurements such as SNR and SPCC. The results show a hiding capacity of 25% of the cover file size, SNR equal to 40.34 dB and robustness against noise addition, cropping, low and high pass filtering with 35 SNR on average. However, the technique has low hiding capacity and security since it uses a number of characters as a secret key for encryption. In addition, there is a limitation of the types and the number of the used characters.

(El-Khamy et al. 2016) proposed a scheme for concealing encrypted images using RSA in audio cover by sample comparison in a DWT domain and coefficients selected using the pseudo number. This scheme is used to enhance security and

robustness. The experimental results show the embedding rate is 5698 bps and 41.73 SNR stego quality. Moreover, the results demonstrate that the proposed scheme is robust against some of the signal processing attacks as AWGN noise, MP3 compression and echo addition. However, the obtained results from the proposed technique in terms of hiding capacity and quality of the reconstructed image file after attacks are low; less than around 1% and 15.9 dB on average, respectively.

Integer Wavelet Transform (IWT) is applied to the audio cover and secret signal is the technique proposed by (Hemalatha et al. 2016) to provide high capacity and security. Only secret approximation coefficients are taken to reduce the size of the secret data and compared with the cover approximation coefficients to generate a key and this key is encoded by arithmetic encoding and hidden in the 2^{nd} and 3^{rd} LSBs of detailed cover coefficients. The audio files used in the experiments were 8 bps and 88 kHz sample rate (by email with the author). SNR and SPCC are used in the evaluation and the results present a hiding capacity of 50% of the cover file and 39.3 dB SNR with 0.9978 SPCC stego quality. The proposed technique is robust against common attacks with 29.2 SNR on average. However, the quality of the reconstructed image is low at about 23.4 dB since the technique ignores half of the wavelet secret coefficients and no secret key is used.

(Kanhe & Aghila 2016) proposed a method to enhance the hiding capacity, robustness, and transparency based on voiced and unvoiced frames and DCT coefficients magnitude modification. The results show that the payload is 0.7 kbps with high stego quality SNR of 51.4 dB and it is robust against re-quantization, re-sampling, and AWGN attacks. On another hand, the hiding capacity is quite low.

(El-Khamy et al. 2017) proposed image in audio hiding scheme for improving the hiding capacity, security and the robustness of the audio steganography using two levels integer wavelet transform, wavelet coefficients modification, XOR, and chaotic map techniques. Payload with 21845 bps, a hiding capacity of 25% of the cover file and SNR 44.6 dB stego quality are the obtained results from the proposed scheme. Moreover, the scheme is robust against AWGN up to 20 dB, MP3 compression with 64 kbps and cropping up to 100000 sample attacks. The hiding capacity is still low.

(Renza et al. 2017) proposed a scheme based on Quantization Index Modulation (QIM) in a wavelet domain and Orthogonal Variable Spreading Factor (OVSF) codes to conceal text in the audio cover. A high transparency stego file and increasing the security of the secret message are the targets of this scheme. The author selected 20 audio files from the website www.freesound.org for their experiments and Haar wavelet was chosen to decompose the audio signal. PSNR and Histogram Error Rate (HER) were used in the evaluation process. The findings are PSNR higher than 100 dB, HER at 10^{-16} and the security of the scheme depends totally on the secret key. The author did not discuss the hiding capacity percentage of the proposed scheme in the article.

Table 2.3 Summary of the related works in transform domain

Author/year	Method/Technique	Secret data	Finding(s)	limitations
(Pooyan & Delforouzi 2007)	LSB with 5 level lifting wavelet transform and hearing threshold	Message	▪ The achieved payload is 239793.75 bps ▪ The hiding capacity is 33.9%. ▪ SNR and MOS are 37 dB and 4.8 on average, respectively	▪ There are no test results about the robustness of the proposed method
(Shirali-Shahreza & Manzuri-Shalmani 2007)	Adaptive lifting Int2Int wavelet scheme and LSB	Message	▪ Payload up to 200 kbps ▪ Hiding capacity 28.3% ▪ 0% error rate with payload less than 100 kbps ▪ SNR around 42 dB	-
(Delforouzi & Pooyan 2008)	Haar wavelet transform, LSB and pseudo number generated with hearing threshold	Message	▪ Payload 296352 bps ▪ Hiding capacity 42% ▪ Stego fidelity SNR equal to 40.2 dB and MOS equal to 4.8	-
(Shahadi & Jidin 2011)	Wavelet packet transform (DWPT) with adaptive hiding based on LSB technique	Text, image, and audio	▪ Payload is 300 kbps ▪ Hiding capacity is up to 42% ▪ Stego quality with SNR of 50 dB	▪ Security against secret message reveal did not discuss
(Sheikhan et al. 2011)	Discrete wavelet transform (DWT) and LSB technique	Audio	▪ Stego quality in SNR is 76 dB and MOS is more than 4.7 ▪ Hiding capacity is 34.5 % of the cover file size	▪ Robustness against secret message revealing did not discuss

to be continued…

…continuation

(Rekik et al. 2012)	Discrete wavelet transforms (DWT) and the fast Fourier transform (FFT)	Speech	▪ The quality of the stego file SNR and PESQ are 22 dB and 3.5, respectively	▪ There is no mentioned about the hiding capacity ▪ No experiments conducted to show the robustness of the proposed method against steganalysis and intentional attack.
(Ballesteros L & Moreno A 2012)	Efficient Wavelet Masking (EWM) indirect LSB substitution	Speech	▪ Robust against statistical analysis with the maximum difference between the cover and stego file is still less than 10%. ▪ The hiding capacity with higher statistical transparency is between 25 and 33%	▪ There is neither result of objective measurement as SNR and PSNR nor robustness against attack and signal processing is conducted
(Ahmed & George 2013)	Haar DWT, UCM and voices and unvoiced blocks		• The achieved hiding capacity is low of 5% • The fidelity of the stego file PSNR is 42 dB.	• There is lack of robustness (security) since the secret data embeds into coefficients of the cover samples in a sequential way • The method did not evaluate using steganalysis techniques
(Shahadi et al. 2014)	Int2Int lifting wavelet transform, LSB and encryption techniques	Text, image, and audio	▪ Hiding capacity is 25% ▪ Stego quality is above 45 dB ▪ Robustness against adaptive noise up to 60 dB (AWGN) ▪ Difference Ratio (DR) of the statistical steganalysis less than 0.7 for the first fourth moments ▪ Simple and symmetry	▪ Low hiding capacity
(Shahadi et al. 2014)	Int2Int lifting wavelet, Weighted Block Matching (WBM) and adaptive position embedding techniques	Text, image, and audio	▪ Payload and hiding capacity is around 340 kbps and 48%, respectively. ▪ The transparency in terms of SNR is above 35 dB	-

to be continued…

…continuation

			▪ AWGN robustness up to 60 with NC=0.9944.	
(Shivdas 2014)	Sample comparison in DCT domain and encryption using pseudo number and XOR operator	Text and audio	▪ The hiding capacity is up to 25% of the cover file size ▪ The SNR of the stego file is within 50 dB.	▪ The hiding capacity is low ▪ Robustness test did not conduct.
(Verma et al. 2014)	Samples comparison in DWT domain using Haar wavelet and encryption using XOR operator	Text	▪ Embedding rate up to 4 kbps that is above 25% of the cover file size ▪ Stego quality SNR above 35 dB.	▪ Robustness test did not discuss.
(Hemalatha et al. 2015)	Haar DWT, coefficients replacement and dynamic encryption algorithm	Text	▪ Hiding capacity 25% of the cover file size ▪ SNR equal to 40.34 dB ▪ Robust against noise addition, cropping, low and high pass filtering with 35 SNR on average	▪ Low hiding capacity ▪ Weakness in the security since using number of characters as a secret key for encryption ▪ There is a limitation of the types and the number of the used characters.
(El-Khamy et al. 2016)	RSA Encryption and sample comparison in DWT domain with random coefficients selection using pseudo number	Image	▪ Embedding rate is 5698 bps ▪ 41.73 SNR stego quality ▪ Robustness against AWGN, MP3 compression and echo addition	▪ Hiding capacity and quality of the reconstructed image file after attacks are low; less than around 1% and 15.9 dB on average, respectively
(Hemalatha et al. 2016)	Integer Wavelet Transform (IWT), Key generation, arithmetic encoding and LSB techniques	Audio	▪ Hiding capacity 50% ▪ Stego quality SNR and SPCC 39.3 dB and 0.9978 respectively	▪ The reconstructed secret quality is low about 23.4 ▪ No secret key is used.
(Kanhe & Aghila 2016)	Voiced and unvoiced frames and DCT coefficients magnitude modification	Message	▪ Payload is 0.7 kbps ▪ High stego quality SNR 51.4 dB ▪ Robust against re-quantization, re-sampling, and AWGN attacks	▪ Hiding capacity is quite low
(El-Khamy et al. 2017)	IWT, wavelet coefficients modification, XOR and chaotic map techniques	Image	▪ . Payload is 21845 bps and hiding capacity 25% of the cover file ▪ SNR 44.6 dB stego quality ▪ Robust against AWGN up to 20 dB,	▪ The hiding capacity is low compared with recent techniques

to be continued…

…continuation

(Renza et al. 2017)	QIM in wavelet domain and Orthogonal Variable Spreading Factor (OVSF)	Text	MP3 compression with 64 kbps and cropping up to 100000 samples attacks. ▪ PSNR higher than 100 dB ▪ HER is 10^{-16} ▪ The security of the scheme depends totally on the secret key	▪ The author did not discuss the hiding capacity percentage of the proposed scheme in the article

2.6.3 Analysis and Limitation of Related Works

Different schemes in time and transform domains have been proposed in the literature as shown in the two sections above. The analysis and the limitations have been extensively discussed. As a summary, there are several major points that can be highlighted in Table 2.4.

Table 2.4 Summary of the critical analysis of the related works

Categories	Time Domain	Transform Domain
Main technique	LSB, PVD and Modulo operator	DCT, DWT and LWT with the LSB, sample comparison, UCM and QIM
Main contribution	Increase the hiding capacity and maintain the transparency	Enhance the robustness and transparency
Hiding capacity	70%	50%
Transparency	52 dB	39-50 dB
Robustness	Limitation in security and did not evaluate using steganalysis	Few related studies are evaluated using signal processing operation and steganalysis

As shown in the table above, the analysis can be directed in two ways. The first one is related to the proposed schemes that were implemented in the time domain. The common techniques were used in this domain is LSB, PVD and modulo operator. These techniques are adopted in order to enhance the hiding capacity and maintain the transparency. Although the achieved hiding capacity using these techniques is 70% of

the cover file size and the transparency is 52 dB, increasing the hiding capacity more than the achieved ratio while preserving the transparency demands further research. On the other hand, there is a weakness in the security aspect since the proposed schemes neither adopt secret key nor use randomizing techniques in the embedding and extraction process. Thus, the secret message may be vulnerable to be discovered by the attacker. Furthermore, the proposed schemes are not evaluated according the common audio steganalysis.

The second way is associated to the performance of the proposed schemes in the transform domain. Regarding the signal processing transformation, DCT, DWT and LWT are the common transforms used along with the LSB, QIM and UCM techniques. Adopting these transforms is to maintain the tradeoff between the robustness and transparency. The achieved hiding capacity of the schemes in this domain is less than the achieved in the time domain which is 50%. However, the robustness is enhanced so that the schemes in this domain withstand some of the unintentional attack such as signal processing operation. Moreover, the transparency is preserved by 39-50 dB, on average. On the other hand, from the security point of view, randomization is adopted by these schemes to enhance the security by selecting only the cover samples for embedding. However, the selection of the secret bits is still in a sequential manner. Some schemes in this domain were evaluated using signal processing operations and audio steganalysis techniques while others are not.

As a result, these techniques any steganography technique that is to be designed or developed should, therefore, deliberate on the primary requirements which are hiding capacity, transparency, and robustness and maintaining the tradeoff between them and evaluating it using most of the evaluation criteria discussed in the literature.

2.7 RELATED CONCEPTS

This section discusses the related concepts that are adopted in the proposed conceptual model in order to achieve the research objectives and improve the performance of the audio steganography system. The discussion in the next subsections is focused on four

main concepts which are data compression, signal processing, embedding and randomization.

2.7.1 Data Compression

Data compression algorithms are used to reduce the number of bits required to represent any file such as an image or a video sequence or music. It is also the art or science of representing information in a compact form. The reason for the need for data compression is that more and more of the information that is generated and used is in a digital form consisting of numbers represented by bytes of data (Tang et al. 2016).

There are two algorithms pertaining to compression techniques: compression, and reconstruction. A compression algorithm takes an input and generates a compressed file that requires fewer bits while a reconstruction algorithm operates on the compressed file to generate the reconstructed file. These algorithms are shown schematically in Figure 2.11. The compression and reconstruction algorithms together are called the compression algorithm.

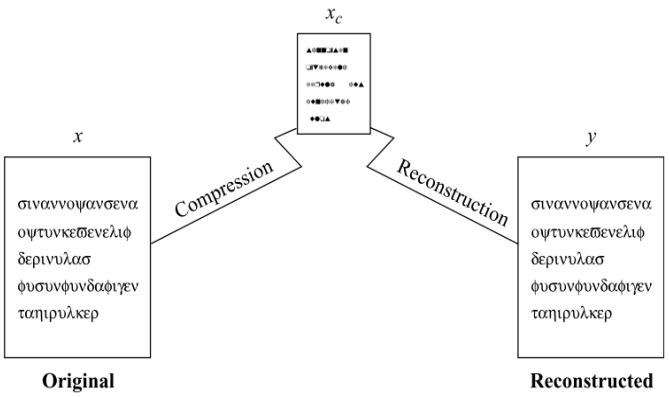

Figure 2.11 Compression and reconstruction (Sayood 2006)

Based on the requirements of a reconstruction file, data compression schemes can be divided into two classes: lossless and lossy compression schemes (Sayood 2012; Ibaida et al. 2014).

For lossless compression techniques, the original data can be recovered fully from the compressed data but at the cost of low or moderate compression ratio. Lossless compression is generally used for applications that cannot tolerate any difference between the original and reconstructed data such as text and medical image compression. Huffman Coding, Run Length Encoding, and Arithmetic Encoding are common examples of lossless compression techniques (Kodituwakku & Amarasinghe 2010).

Lossy compression techniques involve some loss of information and data that have been compressed using lossy techniques generally be recovered or reconstructed with approximation form. The benefit of lossy compression is the higher compression ratio that can generally be obtained compared to lossless compression and less encoding/decoding time (Sayood 2006; Mammeri et al. 2012). There is a tradeoff between the quality of the reconstructed file and compression ratio and this can be assessed by measurement performance. DCT, DWT, VQ, and FC are examples of lossy compression (Lin et al. 2015; Chang et al. 2016; Sheltami et al. 2016; Tang et al. 2016).

According to the discussion above, lossy compression technique can be adopted in the conceptual model since it can produce high compression ratio and lead to increase the hiding capacity which is the second objective of this study.

a. Comparing Lossy Compression Techniques

Nowadays, there is an interest in lossy compression techniques that can be noted in many applications such as Wireless Sensor Network (WSN), information hiding (George & Ahmad 2010; Lin et al. 2015; Pan et al. 2015; Thepade & Thube 2015; Chang et al. 2016; Hemalatha et al. 2016; Tang et al. 2016) and the medical field in wireless body sensor networks (Ibaida et al. 2014) since these applications require a high compression ratio for the data and a good reconstructed signal as the size of the data grows with the limit in bandwidth of the communication channel.

DCT, DWT, VQ, and FC are the common compression algorithms or techniques in the literature (Mammeri et al. 2012; Chang et al. 2016; Tang et al. 2016). Related to the literature review, and in line with our knowledge, there are two articles in wireless sensor networks (WSNs) (Mammeri et al. 2012; Sheltami et al. 2016) and some others (Wu et al. 2005; Wang et al. 2010; Wang et al. 2013; Ibaida et al. 2014; Das & Ghoshal 2015; Jaferzadeh et al. 2016) present the features of these algorithms. Table 2.5 summarizes the features in terms of compression ratio, reconstructed file quality, transform or block-based, the ability to be applicable in real time hardware environments, requested encoding and decoding time, needing an additional requirement for encoding and decoding and computational complexity.

Table 2.5 Comparison of the data compression techniques

Features	DCT	DWT	VQ	FC
Compression ratio	Moderate	Low	Acceptable	High
Reconstructed file quality	Low	High	High	High
Transform/block based	Transform	Transform	Block	Block
Real-time hardware environments	Applicable	Applicable	-	Not Applicable
Encoding/Decoding time	Same Moderate	Same Fast	Long encoder Simple and fast decoder	Long encoder Simple and fast decoder
Additional requirement	No	No	Codebook	No
Computational complexity	High	Moderate	Increase with long codebook	Moderate

As seen in Table 2.5, it is clear that DCT and DWT are the nominated lossy techniques for an application that requires a moderate compression ratio and implementation in real-time environments. Meanwhile, when a high compression ratio is required regardless of the compression process time, VQ and FC is a suitable choice.

According to the analysis above, FC can be used since it can produce high compression ratio and high reconstructed file quality and it does not require codebook in the encoding and decoding process as VQ.