الجمهورية الجزائرية الديمقراطية الشعبية
Democratic and popular republic of Algeria

وزارة التعليم العالي و البحث العلمي
Ministry of higher education and scientific research

Ziane Achour University of Djelfa
Faculty of Exact and Computer Sciences
Department of Mathematics and Computer Science

جامعة زيان عاشور بالجلفة
كلية العلوم الدقيقة والإعلام الآلي
قسم الرياضيات والإعلام الآلي

Doctoral Thesis

Domain: Mathematics&Computer Science

Field: Computer Science

Specialty: Mathematical Optimization for Signal Processing and Communication Networks

By:

Mr. **Ahmed MERRAD**

Implementation of a biometric speech watermarking based on wavelet transform

Defended on: February 5, 2019

Jury Members:

Mr. Abdelhalim MAYOUF	**Professor**	**univ of Djelfa**	**President**
Mr. Slami SAADI	**MC-A**	**univ of Djelfa**	**Supervisor**
Mr. Mecheri KIOUS	**MC-A**	**univ of Laghouat (UATL)**	**Examiner**
Mr. Nadji HERMAS	**MC-A**	**univ of Djelfa**	**Examiner**
Mr. Farid MESELMI	**Professor**	**univ of Djelfa**	**Examiner**
Mr. Ali BENZIANE	**MC-B**	**univ of djelfa**	**Invited**

Academic Year of submission: **2018/2019**

Dedicates

This thesis work is dedicated to:

My beloved parents, my dear aunt “Om Mohammed”, my brothers and my sister, all of my family, my friends and anyone he likes me. I hope that God will preserve them and give them health and wellness.

Acknowledgement

First and foremost, I must acknowledge my limitless thanks to Allah, for blessing me with the power and health to continue this work successfully.

I would like to express my sincerest gratitude to my supervisor Dr. **Slami SAADI** for his patience with me and understanding during this process. I am totally sure that this thesis would have never become truth, without His guidance.

Also I would like to thank to all members of the doctoral training committee for give us an opportunity to exploit our capabilities.

Special thanks to the Jury-members for agreeing to read the thesis and to participate in the defence of this thesis.

I extend my sincere thanks to all members of the faculty of exact and computer sciences.

Finally, thanks to all people who contributed, from near or far, for the success of this work.

Abstract

In the thesis we proposed and implemented two blind and robust schemes. The first scheme for speech and audio watermarking, we used the discrete wavelet transform (DWT) after framing the signal, and then we applied the discrete cosine transform (DCT) on each frame. For correlation purpose, sub-sampling is performed to decompose the frame into two segments. The embedded of watermark bit is in norm value. For security concern, Arnold transform is employed on the watermark image in order to save detection security. The fully blind detection is accomplished without using the original speech/audio signal and the insertion parameter is not required. The second scheme proposed using discrete wavelet transform (DWT) and discrete cosine transform (DCT) after sub-sampling the signal. The insertion of biometric watermark bits are randomly in high energy parts of DCT coefficients. Experimental assessment shows a good tradeoff between security, capacity, imperceptibility and robustness against various signal processing attacks for both audio and speech signals. The comparisons with other published schemes in recent few years demonstrate preference of our proposed schemes.

Key words: Biometric, speech, Watermarking, Wavelet transform, DCT

Résumé

Dans cette thèse, nous avons proposé et mis en œuvre deux systèmes aveugles et robustes. Dans le premier système de tatouage des signaux audio et parole, nous avons utilisé la transformée des ondelettes discrète (DWT) après la division du signal en portions, puis nous avons appliqué la transformée discrète de cosinus (DCT) sur chaque portion. À des fins de corrélation, un sous-échantillonnage est effectué pour décomposer la base en deux segments. Le bit de tatouageembarqué est en valeur normale. Pour des raisons de sécurité, la transformation d'Arnold est utilisée sur l'image de tatouage afin de préserver la sécurité de la détection. La détection aveugle est réalisée sans utilization du signal audio/parole originale et le paramètre d'insertion n'est pas requis. Le deuxième système proposé utilise une transformée en ondelettes discrète (DWT) et une transformée en cosinus discrète (DCT) après sous-échantillonnage du signal. L'insertion de bits de tatouage biométrique se fait d'une manière aléatoire dans les parties à haute énergie des coefficients DCT. L'évaluation expérimentale montre un bon compromis entre sécurité, capacité, imperceptibilité et robustesse contre diverses attaques de traitement de signal pour les signaux audio et parole. Les comparaisons avec d'autres travaux publiés au cours des dernières années démontrent la préférence des systèmes proposés.

Mot-clefs: Biométrie; parole; Tatouage; Transformé en ondelettes; DCT

ملخص

في هذه الأطروحة اقترحنا و نفذنا مخططين صامتين وصلبين. المخطط الأول من اجل تزويد ملف كلام أو صوت بوسم (العلامة المائية)، حيث استخدمنا (DWT) بعد تقسيم الإشارة إلى مقاطع صغيرة ثم طبقنا (DCT) على كل مقطع. بعد ذلك استعملنا تقنية (sub-smpling) للحصول على شعايين تكون قيمهما متقاربة. من اجل الأمان وظفنا تحويل أرنولد. استخراج الوسم في هذا المخطط يكون بدون استعمال الملف الأصلي ولا معامل الإزاحة.أما في المخطط الثاني فإننا استخدمنا (DWT) و (DCT) بعد تجزئة الإشارة إلى قسمين باستعمال تقنية (sub-sampling) ، التزويد بالوسم البيومتري يكون عشوائيا في الجزء الأول لقيم (DCT) أين تكون هناك القيم الكبرى للإشارة. أيضا في هذا المخطط لا نحتاج لا الإشارة الأصلية ولا معامل الإزاحة من اجل استخراج الوسم. التجارب أثبتت و جود توافق بين الشفافية، السعة، الأمان و الصلابة ضد الهجمات المتعددة من اجل كل من ملف الصوت أو الكلام.المقارنات مع طرق أخرى نشرت في السنوات الأخيرة أثبتت أفضلية مخططاتنا.

الكلمات المفتاحية: بيومتري، كلام، الوسم، تحويل المويجات.

Table of contents

List of tables

List of figures

Abbreviations

Bps: Bit Per Second.

DVD: Digital Versatile Disc.

CD: Compact Disc.

ID: IDentification

TV: TeleVision.

MRI: Magnetic Resonance Imaging.

VHF: Very High Frequency.

DCT: Discrete Cosine Transform.

DWT: Discrete Wavelet Transform.

cA: Approximation Coefficients.

cD: Details Coefficients .

IDWT: Inverse DWT.

SVD: Singular Value Decomposition.

Mod: Modulo.

QIM: Quantization Index Modulation.

FFT: Fast Fourier Transform.

LPT: Log-Polar Transformation.

EO: Exponential Operation.

LO: Logarithm Operation.

DA/AD: Analog Discrete/ Discrete Analog.

UDWT: Undecimated Discrete Wavelet Transform .

LWT: Lifting Wavelet Transform.

LQIM: Logarithmic Quantization Index Modulation.

DWPT: Discrete Wavelet Packet Transformation.

QRD: QR decomposition.

SAPSO: Self-Adaptive Particle Swarm Optimization .

QWT: Quaternion Wavelet Transform.

VDVM: Variable-Dimensional Vector Modulation.

ISVD: Inverse SVD.

PCA: Principal Component Analysis.

LSB: Least Significant Bit.

IDCT: Inverse DCT.

SNR: Signal-to-Noise Ratio.

SSNR: Segmental Signal-to-Noise Ratio.

MOS: Mean Opinion Score.

BER: Bit Error Rates.

NC: Normalized Correlation.

AWGN: Add White Gaussian Noise.

SQAM: Sound Quality Assessment Material.

UZAD: University Ziane Achour, Djelfa.

IFPI: International Federation of the Phonographic Industry .

Hz: Hertz.

RAM: Random Access Memory.

Inf: Infinity.

CCCD: Coefficients Cross-Correlation Degree.

AMM: Adaptive Mean Modulation.

HVS: Human Visual System.

HAS: Human Auditory System.

Δ: Quantization step.

General introduction

Currently, digital data becomes an essential component in today's individuals, companies and governments and exceed over the analog data. The success of digital data over analog data is principally due to advantages like: speed of transmission, compact storage, copying without losing the quality and editing plainly. May possibly the advantages alter to disadvantages, from where unlawful using like to illegal copying, manipulation to avoid the content from its original to not authenticate form, speedy distributing using internet networks without permission. Eliminated the unauthorized utilisations of digital content practiced with three arts: steganography, cryptography and watermarking.

Cryptography is the art of coding the messages, where sender convert plaintext to cipher text by using encryption key and in other side, only the intended people could have access to the information and decrypt cipher text to plain text by using the same key [1,2]. Steganography is the art and science of writing secret messages in such a way that no one apart from the intended receiver and sender knows of the existence of the message [3]. Cryptography only protects the contents of a message, but steganography protects the content of messages and the communication parties [4]. Digital watermarking is a technique for insertion additional information directly into host signals; also watermarking techniques are usually one-to-many whereas steganography is a technique that establishes a covered information channel in point-to-point connections [5].

Today, the number of papers written concerning digital watermarking has grown due to watermarking is a technique which provides solution for many important applications. Our interest is on speech and audio watermarking.

watermarking of speech and audio signals is further challenging compared to the watermarking of images or video sequences, due to the broad dynamic range of the human auditory system (HAS) in comparison with human visual system (HVS) .

In this thesis we proposed and implemented two blind schemes for audio and speech watermarking. The schemes work on discrete wavelet domain which the Discrete Wavelet Transform has historically shown its suitability for watermarking applications. The two algorithms proposed to satisfy the requirements of audio and speech watermarking, and also to get better compromise between the three important requirements (robustness, imperceptibility and data payload) and superiority than other schemes presented in recent years.

The thesis is divided into five chapters and organized as follow:

First chapter gives a general idea of digital watermarking concerning definition, basic framework, requirements, applications and classification of digital watermarking also introduced the watermarking in biometric systems.

Second chapter divides into two main parts, the first part includes the popular techniques used in digital watermarking such transformation techniques, Algebraic techniques, encryption technique (Arnold transform). The second part contains a great number of proposed schemes for audio and speech watermarking based on DWT, DCT or based on hybrid DWT and DCT.

Third chapter provides three parts, the first parts gives the first proposed scheme which based on DWT, DCT, sub-sampling, Norm space and Arnold scrambling. The second part presents algorithm based on hybrid DWT/DCT and sub-sampling. The last part includes the metrics and measure to evaluate the performance of our proposed algorithms.

Fourth and Fifth chapters gives the results of proposed schemes from side of tables, graphs, curves, comparisons and all necessary analysis and discussion. The last part of the thesis displays the overall summary of our findings, in addition the perspectives of the future works.

Chapter I

Digital watermarking

I.1. Introduction

Currently, distributing, sharing, producing, editing, recording and archiving the digital content is easy to do in a short period and with high quality, due to the spread of the internet, personal computers and manipulation software, The digital content could be digital audio, speech, image, video, text or any form of digital information, hence, advanced technique for protected and efficient access to information is required, manifested in digital watermarking technique.

Recently digital watermarking has become one of the popular research areas, due to it can offers a new way to solve problems related in the information security. In other words, the digital watermarking has the capability to protect digital content against unauthorized uses.

On the other hand, applications of digital watermarking are several and cover a wide number of fields such as: copyright protection, information carrier, broadcast monitoring, fingerprinting, authenticity data, medical safety, and so on. Digital watermarking also should satisfy some properties like robustness, imperceptibility, capacity etc.

This chapter provides a brief introduction to digital watermarking. Definition of digital watermark and digital watermarking then gives the framework of basic digital watermarking systems and we will know what are the requirements of audio and speech watermarking followed by the various applications are using digital watermarking, then we introduce the different classifications of the digital watermarking, moreover, we give the idea about watermarking in biometric systems. Finally, conclusion summarized the important points in this chapter.

I.2. Digital watermarking background

I.2.1. Digital watermark

A digital watermark is a digital distinguishing piece of information that is merged to a noise tolerant signal as digital speech and audio, and can be a stream of bits that it is intended to protect [6].

I.2.2. Digital watermarking

Digital watermarking is a technique allows merging secret binary information silently within digital data. Deliberation is required in the embedding to reduce undesirable modifications in digital content. The watermark embedding is done without changing the file format or file size. In audio and speech watermarking the perceived sound quality is maintained as well [7]. In addition the embedded information can be extracted using suitable techniques without problems [6].

I.3. Framework of basic digital watermarking systems

Every digital watermarking system divides into two distinct processes: an embedding process and detection process which are depicted in Fig1. The embedding process uses the digital content as host signal, the watermark bits and key to produce the watermarked data. The detection process takes the (possibly modified) watermarked data, the key and optionally original data and extracts the watermark [8].

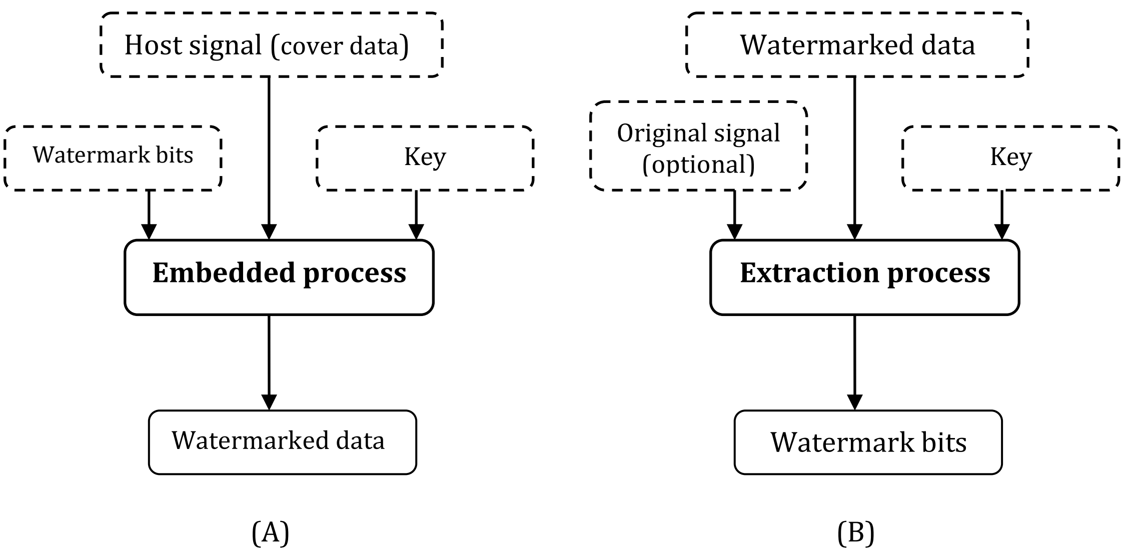


Figure 1: General model for digital watermarking (A): embedding process (B): Detection process

I.4. Requirements of audio and speech watermarking

The audio and speech watermarking systems are generally desired to satisfy some requirements, like robustness, security, capacity, imperceptibility and speed. However, designing a watermarking system excels in all of the requirements is impossible [9]. Therefore, Watermark scheme properties depend extremely on the application for which the watermark is designed, for example, it is no necessary to create a robust watermarking scheme for secure information carrier application.

In the below part, we will examine those properties in detail.

I.4.1. Imperceptibility

In some applications, the watermark embedding process must not influence the perceptual quality of original audio/speech signal. The difference between the original audio/speech and watermarked audio/speech version can hardly be distinguished by the human ears.

In addition, there are two approaches used to assess the perceptual quality of audio: subjective evaluation test and objective evaluation test.

I.4.2. Robustness

For the watermarks schemes that are not specially designed to be fragile, Robustness is an important postulate . The embedded watermark data should not be removed or eliminated during normal usage or by unauthorized distributors using common signal processing operations and attacks. Namely, the extraction process can detect the digital watermark from the attacked watermarked signal version. There are a many expected attacks on audio/speech signals for Examples noise addition (AWGN), re-sampling, re-quantization, random samples cropping etc [10].

I.4.3.Capacity

The quantity of bits that can be embedded into a host signal within a unit of time is defined as capacity or payload. In digital audio/speech watermarking system is the numbers of bits that can be embedded into the audio/speech signal in a one-second audio/speech fraction, expressed in bit per second (bit/s or bps). Necessity of data payload varies, depending on the watermarking applications and the embedding watermarking scheme [9, 11, 12, and 13].

I.4.4.Security

The property of security is indispensable in all watermarking systems. The security implies that the watermark can only be detectable by the authorized person [13]. Otherwise the attackers may possibly detect the watermark. Then, they are able to modify the watermark without much impairment on the digital data quality. In this case, secret keys (usually pseudorandom sequences) and/or scrambling operations can be adopted to add randomness into the embedding and extraction processes, so that the digital watermarking system is self-secured [10].

I.4.5.Speeds

The required speed of watermark system depends on the application at hand, For example, in Broadcast monitoring applications, embedding and detection must be through real time [11], but in the purpose of copyrights protection, no trouble too much about the embedding time, as long as it is not weird. On the contrary, the detection phase is expected to take as short time as possible [13].

I.4.6.Blind detection

Watermark detectors scheme can be classified into informed and blind, according to whether the original signal needs to be available to the watermark detection process or not. An informed detector, also famous as a non-blind detector, uses the original signal in a detection process whereas blind detectors do not use the original signal for watermark detection. Although non-blind schemes are more robust in detecting watermarks, the multimedia industry appears to favour the blind schemes due to their practicality [14].

I.4.7.Trade-off

The robustness, imperceptibility and capacity are three disharmonious important properties of a watermarking scheme [11]. From fig.2 observed that there exists a trade-off between them, for instance, increasing the capacity typically introduces additional distortion into data content, and, also, decreasing capacity decreases robustness. Consequently, a trade-off between them must be achieved [15].

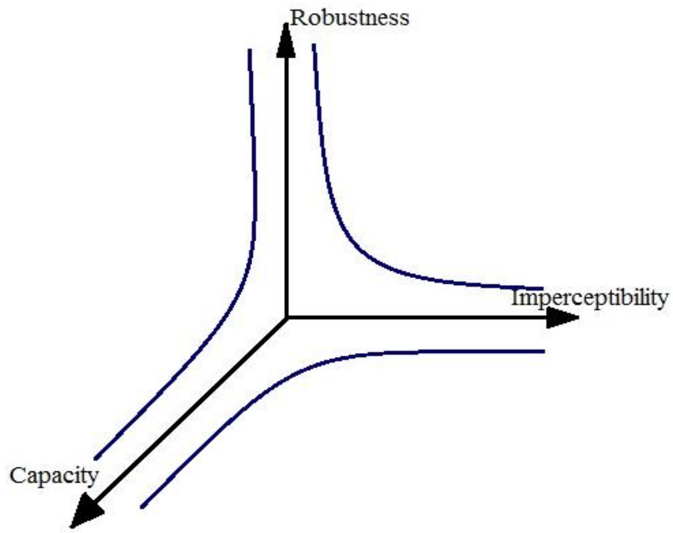


Figure 2: Trade-off among robustness, imperceptibility and capacity

I.5. Watermarking applications

Digital watermarking is one of the important technologies due to the using in a broad range of applications like copyright protection, copy protection and device control, fingerprinting, data authentication and tampering verification, broadcast monitoring, secure information carrier and medical applications.

I.5.1.Copyright protection

Copyright protection is the most important application of digital watermarking due to the exploration of digital watermarking was driven by the desire for copyrights protection [10].

The underlying strategy idea is to integrate a watermark by the authors or originators containing their own intellectual property signature such a logo, message ... into the original multimedia data and delivers it as usual. By doing this, and in dispute case the rightful owner can demonstrate the ownership by extracting the embedded watermark [16].

In this application the watermarking should be very robust and secure to survive common signal processing modifications and intentional attacks [17].

On the other hand, since ownership protection applications is not necessary for the watermark to be very long, the data payload for this application does not have to be high [10,17].

I.5.2.Copy protection and device control

It is possible for playback and recording devices to react to embedded signals [18]. Digital watermarks can be embedded within a digital data to enable copy control devices, [16] in this combination, the recording devices might prevent recording action of a signal if it detects a watermark that indicates recording is prohibited [18]. In such a system, watermarks containing copy control information and identification bits are embedded in the DVD audio track repeatedly. During playback, if the detected watermarks do not match those of specific disc, then the playback will be halted [19].

I.5.3.Fingerprinting

Customers are buying different data types, such as images, video, and audio over the Internet or on CDs/DVDs, but some customers can make illegal copies or redistribute them. In this case additional data embedded by a watermark in the fingerprinting applications can increase the data security and discover the source of the leak. To recognize those who make unlawful action, an automated agent scanning system can be used to track down the traitor [16]. For example, watermarks carrying hidden dissimilar serial or ID numbers are embedded in different copies of movie CDs or DVDs before distributing them to a large number of recipients [17].

On the other hand, biometrics technology, such as fingerprint, iris, and speech recognition, plays an essential role in today's personal identification systems. Digital watermarking of fingerprint images can be applied to protect the fingerprint images against malicious attacks, can discover trickery fingerprint images, and can make secure transmission .Fingerprinting in digital watermarking is usually used as the process of embedding the identity to an image in such a way that it is difficult to remove [20].

The algorithms implemented in fingerprinting applications require high robustness against intentional attacks and signal processing modifications also the embedding capacity required [17].

I.5.4.Data authentication and tampering verification

Experts of information technology advice the people in this "Do not completely trust what you see in digital form". For the individuals who want to recognize whether the digital content is trust worthy, fragile watermarking techniques provide a possible solution [21].

The digital watermark can be used to confirm that the digital content has not been tampered. Any such modification on the data destroys or changes the integrated watermark. To prove the authenticity it should be extract the watermark bits without errors [16].

The watermarking for data authentication and tampering verification requires the fragility of watermark, the embedding capacity has to be high and the detection must be performed without the original host signal [17].

For example if say Bob wants to transmit a digital file to Alice. He embeds a fragile watermark in the file and delivers it to Alice by a channel which could be the Internet. Before Alice receives the file, John happens to obtain the watermarked file. He modifies the content of the file and sends it to Alice afterwards. When Alice receives the corrupted file she has no idea as to whether the content is trusty. She therefore verifies if the received file contains the watermark. Due to the fragile nature of the watermark, it has weak resistance against tampering; Alice is unable to find any watermark. She knows immediately that the file she has received had been tampered with [21].

I.5.5.Broadcast monitoring

Several companies and individuals like advertisers, owners of copyrighted works and performers are interested in the field of broadcast monitoring.

Designed by advertisers to ensure that they receive all of the air time they purchase for radio/TV station. Used by owners of copyrighted works to make sure their works are not unlawfully re-broadcasted by other impermissible stations. Designed by performers to assemble the royalties from radio or TV stations once broadcasting their works.

It is costly and prone to error to employ an individual to monitor the broadcast by listening, watching or recording the broadcast. Watermarks however, can be embedded to the digital content before broadcasting. Then the Computer systems can be used to monitor broadcasting by examines the existence of watermarks from the broadcasted content [22].

I.5.6.Secure information carrier

The watermarking techniques can offer an ideal solution for transferring digital content from one place to another place in a safe mode [23]. In this application the embedded watermark is expected to have a high capacity, the robustness against intentional attacks is not necessary and the decoding algorithm should be without using original signal [17].

I.5.7.Medical applications

Medical field is another important application in watermarking. The watermarking in this field can collect all information of one patient in one data, which it grantee impossibility to mix between two patients because the mix leads to disaster.

For example information about patients such as names, personifications and their diagnosis can be embedded within medical images of patients. The medical images could be X-ray image or MRI image. In transmission case to guarantee security, it can be use the medical images with the information of patients as watermark and embedded it within other data.

However, the medical image watermarking requires great prudence when embedding additional data within the medical images because the additional information must not affect the image quality [24].

I.5.8.Air traffic control

Digital watermarking also can be applicative for air traffic control. In an air traffic control environment, there are several aircrafts communicating with the controller in a single very high frequency (VHF) channel. The aviator of an aircraft starts the communication by indicating the aircraft call sign. Generally, the aircraft registration number serves as the call sign. There is possible for confusion if two flights on the same VHF channel at any time have similar sounding call signs. By hiding exclusive information about an aircraft in the voice message, any doubt over aircraft identification is prohibited. Digital watermarking of speech is thus used to supply automatic identification of the aircraft [17].

Great number of digital watermarking applications mentioned above, and its importance establish importance of this technology in nowadays through it is can lead us to a safe technology.

I.6. Classification of digital watermarking techniques

Watermarking techniques on general can be classified into several categories as shown in Fig .3. The watermarking can touch different *types of digital data* such as image, audio, speech, video and text document. Appending to *working domain*, the watermark system could be embedded the watermark bits in spatial and transform domain. *From blindness* side, system watermarking is categorized into two ways including blind or non-blind extraction as defined in section I.4.6. According to *the human perception*, the watermarks can be divided into two types: perceptible or imperceptible watermark. Perceptible watermarks can be appear to eyewitness in images and video watermarking, however, an audible sound in any instant time of digital audio, speech and video. Imperceptible as defined before in section I.4.1. Appending to *robustness* can be classified into robust, fragile or semi-fragile watermark. Robustness of watermark as defined in section I.4.2. A fragile watermark is a watermark that is sensitive to any manipulation, generally applied in data authentication purpose. In a temperate approach, a semi-fragile watermark is marginally robust and can be sensitive to some attacks.

The watermarking techniques can also be classified *into reversible and irreversible* techniques; reversible watermarking approach allows deleting the whole watermark and obtaining the exact host signal from the watermarked signal. However, design a reversible watermark scheme implicates a few losses of robustness and security. Non-reversible watermarking usually introduces a slight but irreversible degradation in the original signal. The adaptation reversible Watermarking system must only in applications where need total restoration of the host signal such in medical application.

I.7. Watermarking in biometric systems

Biometric watermarking is an idea allows doing hybridization between biometric technologies and watermarking. Objective of this approach to employ biometric templates such a digital fingerprint as "watermark" to be embedded in classical robust watermarking applications like copyright protection in order to enable biometric recognition after the extraction of the watermark. Therefore, the capacity and imperceptibility are required in these digital watermarking systems, the robustness against unintentional and malicious cover data manipulations is necessary [25].

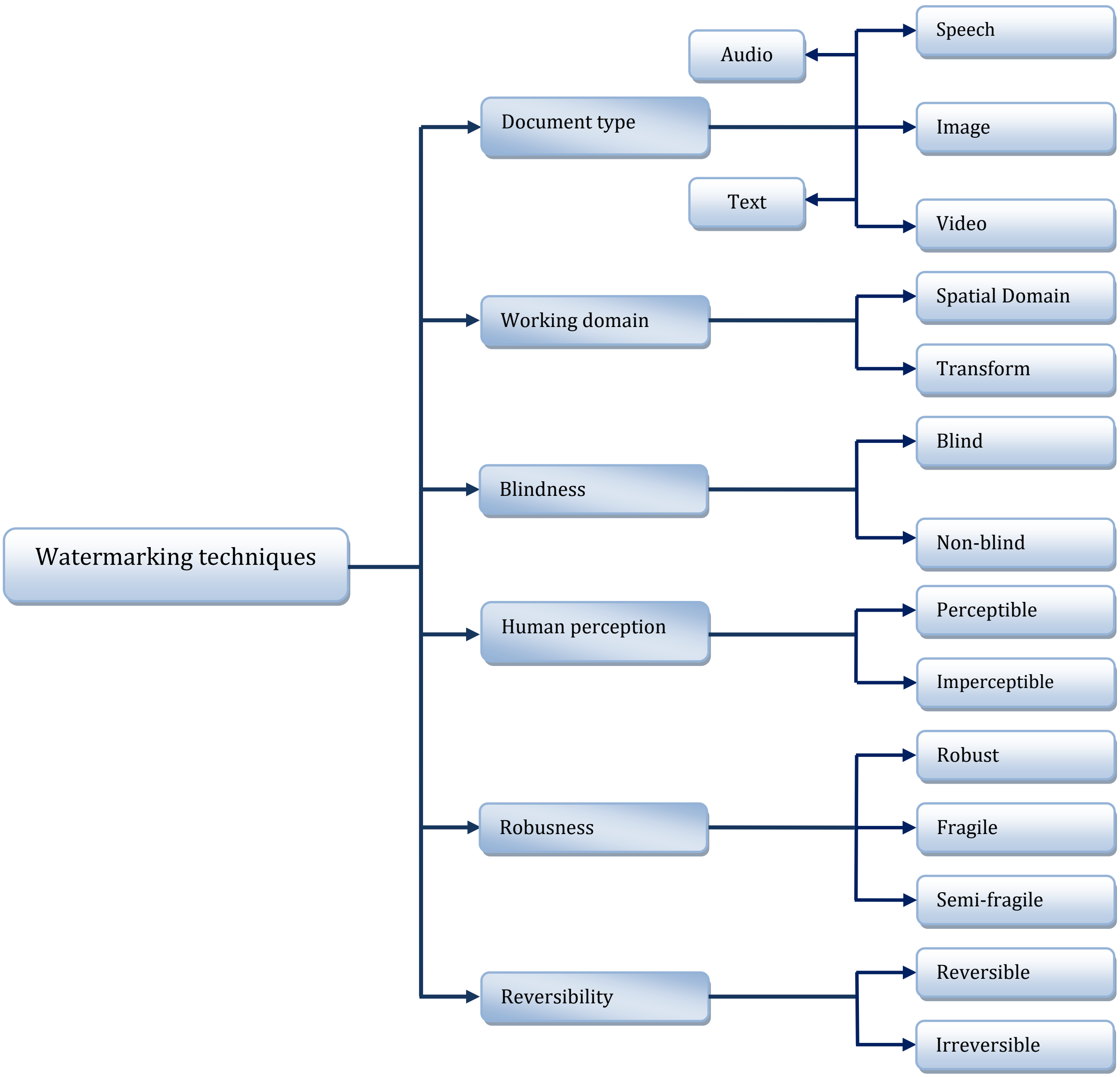


Figure 3: Watermarking techniques classifications

I.8. Conclusion

In this chapter we tried to give the better definition of digital watermark and digital watermarking, then we presented the general framework of basic digital watermarking systems also we gave a figure for understanding the system of watermarking without complexity. In other section we showed the requirements of audio and speech watermarking. For known importance of digital watermarking we gave a many field can employ it. Also we presented diverse classifications of digital watermarking. In last element, the chapter introduced the watermarking in the biometric systems.

Chapter II

Techniques and state of the art

II.1. Introduction

Through the digital watermarking can offer solution for security of multimedia data, there is a lot and different of schemes designed for it and became an interesting research field. In the audio and speech watermarking case a various techniques has been used to apply the watermark. This chapter divide into two main parts; in the first part we'll give some used techniques, such the transformation techniques (DWT, DCT), Algebraic techniques (SVD, QR, NORM), Arnold transform and QIM. The second part includes great amount of proposed algorithms in the few recent years for speech and audio watermarking, which based on transformation approaches. Also we introduce time domain aspect briefly.

II.2.Techniques

II.2.1.Transformation techniques

II.2.1.1.Discrete cosine transform (DCT)

The DCT is a recognized transform capable to illustrate fragments of an audio signal in terms of summing up of cosine functions in diverse frequencies. One of the major important obvious features of DCT transform is energy storage in a small number of samples. This feature is used to decrease curvature of the original signal in speech watermarking process [26-27]. The discrete cosine transform is a scheme for converting a signal into fundamental frequency components. The DCT definition of a 1-D sequence of length N is:

$$c(u) = a(u) \sum_{x=0}^{N-1} f(x) \cos\left(\frac{\pi(2x+1)u}{2N}\right) \quad (1)$$

For $u = 0,1,2,\dots,N-1$

Where, *x(n)* is the original signal and N is the number of samples. In analogous way, the inverse transform is expressed as:

$$f(x) = \sum_{x=0}^{N-1} a(u)\mathrm{c(u)}\cos(\frac{\pi(2x+1)u}{2N}) \quad (2)$$

For $u = 0,1,2,\dots,N-1$

In both equations, *a(u)* is defined as:

$$a(u) = \begin{cases} \frac{1}{\sqrt{N}} u = 0 \\ \sqrt{\frac{2}{N}} u \neq 0 \end{cases} \quad (3)$$

The characteristics of this algorithm are strong, well hidden and resistant to a variety of signal deformation resistance. The digital watermark in the DCT transform domain has important ability of lossy compression resistance. The disadvantage is its immense amount of calculations [28].

II.2.1.2.Discrete wavelet transform (DWT)

The DWT is a novel transform that gives a time-frequency representation of a signal [29]. It was developed to overcome the small variations of the signal with time that are not well covered by Fourier transform in frequency domain. It can as well be practical to analyze non stationary signals [29]. And it is used in a large scale for signal processing purposes [30-31]. DWT decomposes an input signal ***S*** into two sets of coefficients, at the heart of DWT is a pair of filters: low pass and high pass, the approximation coefficients cA1 (low frequencies) are produced by passing the signal throughout low pass filter, the details coefficients cD1 (high frequencies) are produced by passing the signal throughout high pass filter, followed by down-sampling.

Depending on the purpose and the length of the signal, the signal is decomposed on multi-level discrete wavelets [32], where the next decomposition level splits the approximation coefficients cA1 in two parts using the same scheme, replacing S by cA1, and producing cA2 and cD2. Fig.4 illustrates 2 phases DWT decomposition:

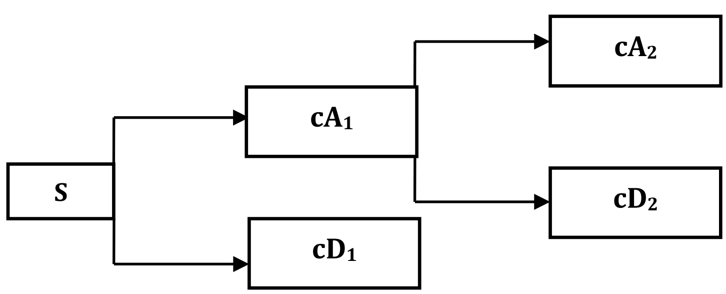


Figure 4: 2-levels DWT decomposition

Inverse DWT process reconstructs or synthesizes the original signal by assembling those components back without loss of information [33], the up-sampling operator is used to recompose the samples eliminated by down-sampling. Fig.5:

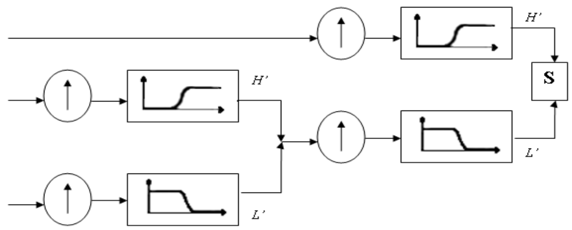


Figure 5: Rebuilding a decomposed signal with IDWT

II.2.2.Algebraic techniques

II.2.2.1.Singular value decomposion

The Singular Value Decomposition SVD is a numerical technique in linear algebra, the SVD of a matrix $A_{N\times N}$ is the factorization of A into the product of three matrices A = USV^T as shown in equation below :

$$\begin{bmatrix} A_{1,1} & \cdots & A_{1,n} \\ A_{2,1} & & A_{2,n} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,n} \end{bmatrix} = \begin{bmatrix} U_{1,1} & \cdots & U_{1,n} \\ U_{2,1} & & U_{2,n} \\ \vdots & \ddots & \vdots \\ U_{n,1} & \cdots & U_{n,n} \end{bmatrix} \times \begin{bmatrix} S_{1,1} & \cdots & 0 \\ 0 & & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & S_{n,n} \end{bmatrix} \times \begin{bmatrix} V_{1,1} & \cdots & V_{1,n} \\ V_{2,1} & & V_{2,n} \\ \vdots & \ddots & \vdots \\ V_{n,1} & \cdots & V_{n,n} \end{bmatrix}^T \quad (4)$$

Where the U and V are orthogonal and the matrix S is diagonal matrix with positive elements, and superscript T denotes matrix transposition. The diagonal elements of S are called the singular values (SVs) of A and are assumed to be arranged in decreasing order $S_{i,i}>S_{i+1,i+1}$. The columns of U, denoted by U_i, are called the left singular vectors, while the columns of V, denoted by V_i, are called the right singular vectors of A.

The SVD has several interesting characteristics: the sizes of the matrices for SVD transformation are not fixed, and the matrices need not be square, changing SVs slightly does not influence the quality of the signal much, the SVs are invariant under common signal processing operations, and the SVs suit intrinsic algebraic properties [11].

The SVD transform has been used in many audio and speech watermarking algorithms [34, 35]. The algorithms varied in the way the singular values were used in the watermarking process.

II.2.2.2.QR decomposion

QR factorization is another numerical technique in linear algebra, QR decomposition is an elementary operation, which decomposes a matrix into an orthogonal and a triangular matrices. Let A be a m×n real matrix. This matrix can be decomposed using the QR as follows:

$$A = Q \times R \qquad (5)$$

Where Q is m×n orthogonal matrix ($Q^T \cdot Q = I$) and R is n×n upper triangular matrix.

The R matrix can be used for scheming robust watermarking method due to the elements of R matrix do not change notably when a perturbation is added to matrix A [36]. For that there are authors used QR decomposion to designing image watermarking schemes [37, 38] and audio watermarking schemes [36, 39].

II.2.2.3. Norm space

Norm space is an important numerical analysis in the linear algebra. To define the norm we suppose that $A=\{a_i, 1 \le i \le N\}$ is a 1×N vector, σ is the norm of A, after that we can get that:

$$\sigma = \|A\| = \sqrt{\sum_{i=1}^{n} a_i^2} \qquad (6)$$

$$A = \sigma u^T \qquad (7)$$

Where $u = \frac{A^T}{\|A\|}$ is a n×1 vector

In the watermarking methods the embedding of the watermark bit is in the norm space, so to get a modified norm σ_w, it can reconstruct A_w with σ_w, which is called inverse norm,

$$A_w = \sigma_w u^T \qquad (8)$$

The embedding in the norm space can be spread the watermark information throughout the vector of the norm which can gives the watermarking algorithms high robustness as demonstrated in [40,41].

II.2.3.Arnold transform

The KxK binary watermark image W is transformed into W' by Arnold transformation to reduce the autocorrelation coefficient of image and next the privacy of watermark is reinforce [42]. Arnold transformation is cyclic and while it is iterated occasionally the original signal will be reached. The Arnold scrambling algorithm [43] has the characteristic of ease and periodicity, so it is used usually to offer an extra level of safety all along through digital watermarking. Arnold Transform is well recognized as cat look transforms and is just appropriate for N × N dimension signals. It is defined as:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} \equiv \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} mod\ N \qquad (9)$$

Where mod N is modulo N (Euclidian division rest), (x, y) are the coordinates of original watermark and (x', y') is the coordinates of scrambled watermark. N is the height or size of the signal which is to be processed. Arnold Transform is periodic in nature. The decryption of signal depends on the scrambling key which can be employed as secret key and defines the number of times it has been scrambled.

II.2.4.Quantization index modulation (QIM)

Quantization Index Modulation (QIM) is popular and a simplest method employed in several audio watermarking algorithms to embed and extract the watermark bits. For example [44]: an implantation of QIM as follows: suppose the original sample is x, the quantization step is Δ, the quantization function is q(x,Δ), w represents the watermark bit to be embedded (0 or 1), then the watermarked sample y is denoted as:

$$y = q(x, \Delta) + \frac{\Delta}{4} \times (2 \times w - 1) \qquad (10)$$

The quantization function is defined as below:

$$q(x, \Delta) = \left[\frac{x}{\Delta}\right] \times \Delta \qquad (11)$$

Where [x] is the rounding function which rounds to the nearby integer of x. In Figure 6, firstly the sample x is quantized to the q(x,Δ) or black circle. If the to be embedded watermark bit is 1, then the $\Delta/4$ is added to the quantized sample value which shifts the sample up to the white circle. If not, $\Delta/4$ is subtracted from the quantized sample value, which moves the sample down to the cross (x).

At the decoder part, the difference between the received sample and its quantized value is computed. If it is between (0, Δ/4), then the extracted watermark bit is "1". If the difference lies between (-d/4, 0), then the embedded watermark bit is "0". Otherwise, the received signal is not watermarked. This can be illustrated with bellow equations:

$$w = 1, if\ 0 < y - q(y, \Delta) \leq \frac{\Delta}{4} \qquad (12)$$

$$w = 0, if\ -\frac{\Delta}{4} \leq y - q(y, \Delta) < 0 \qquad (13)$$

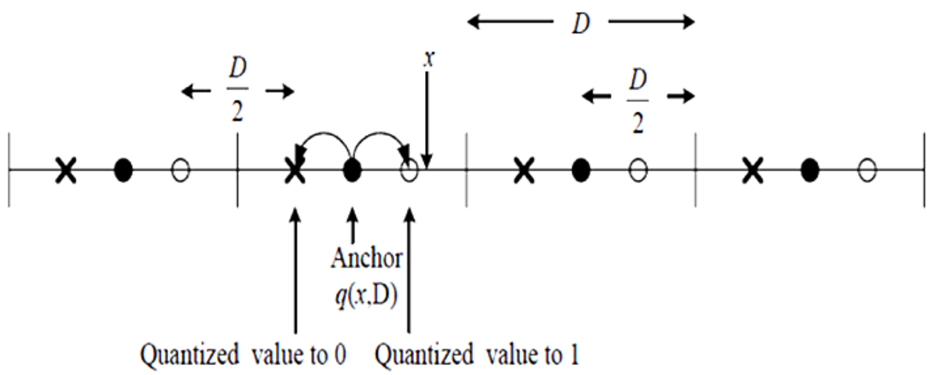


Figure 6: QIM illustration

II.3. State of the art

II.3.1. Methods based on Transformation domain

The embedding in frequency domain makes the watermark more robust than the time domain because it offers to embed the watermark bits in fundamental frequencies of the signal. It can be represents the signal in frequency by computing with mathematical transformations like fast Fourier transform (FFT), discrete cosine transform (DCT) and discrete wavelet transform (DWT).

II.3.1.1.Algorithms based on DCT

Blind and robust audio watermarking scheme is given in [45], the adopted watermarking technique combined with SVD, DCT and synchronization code technique. The watermark bits embedded within high-frequency band of the SVD-DCT block blindly. Also a chaotic sequence is adopted as the synchronization code and inserted into the host signal.

For copyright protection of audio signal, the authors in [46] proposed blind singular value decomposition (SVD) based audio watermarking scheme using entropy and log-polar transformation (LPT). Firstly the original audio is divided into non overlapping segments and discrete cosine transform (DCT) is applied to each frame. Low frequency DCT coefficients are segmented into sub band and entropy of each sub band is calculated. Watermark data is embedded into the Cartesian components of the largest singular value obtained from the DCT sub band with highest entropy value of each frame by quantization.

In [47] the authors implement a blind audio watermarking methodology for robust, transparent and high capacity watermarking technique. The watermark embedding is performed by modulating the vectors in the DCT domain subject to an auditory masking constraint and the abrupt artefacts in frame boundaries are further rectified via linear interpolation over transition areas.

The authors of paper in [48] introduced a blind audio watermarking algorithm in discrete cosine transform (DCT) domain based on singular value decomposition (SVD), exponential operation (EO), and logarithm operation (LO). However, the scheme to begin with framing the original audio signal into non-overlapping segments then DCT is applied to each segment. Low frequency DCT coefficients are segmented into sub-bands and energy of each sub band is calculated. EO is performed on the sub-band with highest power of the DCT coefficients of each frame. SVD is applied to the exponential coefficients of every sub bands with highest power represented in matrix form. Watermark information bit is embedded into the largest singular value by using a quantization function.

II.3.1.2. Algorithms based on DWT

Paper in [49] introduced a DWT based audio watermarking algorithm robust against the DA/AD conversions. To oppose the magnitude distortion, the relative power relationships among different groups of the DWT coefficients in the low-frequency sub-band are utilized in watermark embedding. Additionally, the resynchronization is proposed to cope with the linear temporal scaling. The time-frequency localization features of DWT are exploited to save the computational load in the resynchronization.

Authors in [50] used undecimated discrete wavelet transform (UDWT) and invariant histogram for audio watermarking algorithm with excellent audible quality and realistic resistance against de-synchronization attack such as arbitrary cropping, time-scale change, pitch shifting, and jittering. The proposed scheme begin with performing undecimated discrete wavelet transform (UDWT) is performed on original host audio. Secondly, the invariant histogram is extracted from a chosen wavelet coefficients range in the approximation coefficients. Followed by, the bin of histogram is segmented into several groups, each group including four successive bins. For each group, one watermark bit is embedded by reassigning the number of wavelet coefficients in this group of four bins. Finally, the digital watermark is embedded into the original audio signal in UDWT domain by modifying a little set of wavelet coefficients.

Bahat and all in [51] suggested secure, robust, and blind adaptive audio watermarking scheme based on SVD in the DWT domain using synchronization code. The watermark is embedded by performing a quantization index modulation (QIM) method on the singular values in the SVD of the wavelet domain blocks.

For an imperceptible and robust audio watermarking, the paper in [52] introduced an algorithm based on the discrete wavelet transform. Whereas, to locate the most appropriate regions where the watermark bits embed imperceptibly and robustly, the host original audio signal was decomposed by performing two-level DWT, in addition the embedding was did in details coefficients.

Lifting wavelet transform (LWT) and singular value decomposition (SVD) are used in [53] by inserting the watermark in the coefficients of the LWT approximation coefficients taking advantage of both SVD and quantization index modulation (QIM). Additionally, the synchronization code technique is also integrated into the hybrid LWT–SVD audio watermarking method.

Paper in [41] introduced a blind and adaptive audio watermarking algorithm. The algorithm encrypts the binary watermark image by Arnold transform an embedded it in the vector norm of divided approximation components, after DWT of the original audio signal through quantization index modulation (QIM) with an adaptive quantization step selection scheme. Furthermore, a detailed method has been designed to seek the appropriate quantization step parameters.

A blind audio watermarking algorithm based on the vector norm and the logarithmic quantization index modulation (LQIM) in the wavelet domain is introduced in [40]. The algorithm adopted µ-Law companding to transform the vector norm of the segmented wavelet approximation components of the original audio signal. And then a binary watermark image scrambled by the chaotic sequence is embedded in the transformed domain with a uniform quantization scheme.

In [32] the proposed scheme is for embedding copyright information within audio files as a proof of their ownership. The proposed algorithm embeds the watermark bits on the elements of singular values of the Discrete Wavelet Transform (DWT) sub-bands of the audio frames.

An audio watermarking technique for copyright protection is given in [54]. The watermarking algorithm is based on Discrete Wavelet Transform (DWT) and the Singular Value Decomposition (SVD) techniques. Firstly, the input audio signal is segment into frames, followed by DWT decomposition, also the embedding method is proposed.

Non-blind, imperceptible and robust audio watermarking algorithm, based on the Discrete Wavelet Transform (DWT) and the Singular Value Decomposition (SVD) is proposed in [55]. The audio signal is sampled, quantized, and partitioned into frames by the algorithm, also A three-level DWT operation is applied on every frame, followed by a matrix formation of the frames'third detail sub-bands, on which the SVD operator is applied. The algorithm added the singular values of both the audio signal and the watermark image in order to embed watermark bits.

Authors in [29] create a new scheme for blind digital audio watermarking based on DWT and SVD. In the algorithm, an original audio signal is divide as blocks and each block is decomposed on discrete wavelet transform for two level, then SVD transform is applied on the first quarter audio approximate sub-band coefficients, obtain a diagonal matrix. The watermark information is embedded into the diagonal matrix. The drawback of this scheme is not robust against random cropping.

DWT-Arnold Transform based audio watermarking technique is suggested in [42]. Firstly the original signal is divided into many frames. Secondly, perform DWT on original audio signal. Then, embedded the transformed watermark by Arnold transform within low frequency using proposed equation.

Using the flexibility of discrete wavelet packet transformation (DWPT) to approximate the critical bands and adaptively determines suitable embedding strengths for carrying out quantization index modulation (QIM), an audio blind watermarking scheme is presented in [56]. The singular value decomposition (SVD) also employed to analyze the matrix created by the DWPT coefficients and insert watermark bits by manipulating singular values subject to perceptual criteria.

Author in [39] proposed a blind audio watermarking algorithm based on lifting wavelet transform (LWT) and QR decomposition (QRD) for audio copyright protection. The proposed method divides the original audio signal into non-overlapping frames, and then select the approximate coefficients obtained by performing two-level LWT on each frame and rearranged it into a square matrix, followed by applying QRD on each matrix. Watermark bit is embedded into the largest element of the upper triangular matrix.

A new audio watermarking algorithm based on self-adaptive particle swarm optimization (SAPSO) and quaternion wavelet transform (QWT) is suggested in [57].a synchronization sequence generated by chaotic signals is also employed in the algorithm to resist de-synchronization attack. The proposed scheme embedded the watermark by modifying the singular values of the host signal based on the MSS algorithm.

Authors of [58] introduced a flexible variable-dimensional vector modulation (VDVM) scheme to maximize the efficiency of the norm-space DWT-based blind audio watermarking. The watermarking method is carried out by modifying the vector norms drawn from the DWT coefficients in approximation coefficients. The embedding power, which is manifested as the quantization step size, has been deliberately regulated subject to the auditory masking threshold.

Approach proposed in [36] for blind audio watermarking using the QR factorization in wavelet domain. The watermark image is embedded in the R matrices of low frequency blocks DWT coefficients of audio signal. The embedding of watermark is by applying a Quantization Index Modulation (QIM) process on the determined optimal sample for every matrix R.

Blind digital speech watermarking technique for online speaker recognition systems is presented in [59]. That scheme based on Discrete Wavelet Packet Transform (DWPT) and multiplication to embed the watermark in the amplitudes of the wavelet's sub bands.

The presented system in [60] is based on wavelet Transform (DWT) for blind audio watermarking. The original audio undergoes wavelet based approach and later segments it into frames. Fibonacci numbers is used to embed the watermark bits into DWT elements.

Paper in [61] presents a blind audio watermarking algorithm in transformed domains based on SVD, DWT, and QIM. In the scheme, an original audio signal is split into blocks and each block is decomposed into two levels discrete wavelet transform, and then the approximation coefficients are decomposed by the SVD transform, obtaining a diagonal matrix. The prepared watermarking and synchronization code bit stream is embedded into the diagonal matrix using Quantization Index Modulation (QIM). Following that, we apply ISVD and IDWT to obtain the watermarked audio signal.

An adaptive audio watermarking algorithm in the wavelet domain presented in [62] to optimize the payload by strategically using some of its local features. The proposed adaptive algorithm aims to resolve the problem of over-loading and under-loading the audio signals with watermark data making the payload optimized for every individual audio signal. Some audio features are strategically extracted and the most discriminatory features are selected by Principal Component analysis (PCA) approach.

Authors in [63] outline a package synchronization scheme for blind speech watermarking in the discrete wavelet transform (DWT) domain. Following two-level DWT decomposition, watermark bits and synchronization codes are embedded within selected frames in the second-level approximation and detail sub-bands, respectively where the embedded synchronization code is used for frame alignment and as a location indicator.

II.3.1.3.Algorithms based on Hybrid DCT and DWT

A DWPT-DCT framework for blind audio watermarking is presented in [64]. However, framework jointly exploiting the discrete wavelet packet transform (DWPT) and the discrete cosine transform (DCT) to perform variable-capacity blind audio watermarking without introducing perceptible distortion. In this algorithm the quantization steps for QIM are not only perceptually determinable during watermark embedding but also retrievable during watermark extraction.

Authors in [65] proposed a blind scheme for audio watermarking using Arnold transformation with discrete wavelet and cosine transform. The 2-level DWT is performed on the input digital audio signal then the approximation components divided into frames, followed by apply DCT on each frame where scrambled watermark image by Arnold transform is embedded. To obtain the watermarked version all of segments are regrouped before apply inverse of both DWT and DCT.

The scheme presented in [27] beginning by framing the audio signal into various segments of fixed length, Do the DCT on low frequency coefficients later than apply the H-level DWT on each segment. Embedding the scrambled Watermark bits as per the Quantization Function selected. The scrambling of watermark image is done by Arnold transform. Finally, apply the inverse DCT and inverse DWT on the modified low frequency coefficients followed by the re-arrangement of modified segments into a single audio.

In [43] authors proposed new algorithm for audio watermarking using Discrete Wavelet Transform (DWT) and Discrete Cosine Transform (DCT). In addition, Arnold transform and error correction technique are utilized to progress the performance of the proposed algorithm. Watermark is embedded in the DCT blocks of the selected middle frequency sub-bands of 3-levels DWT transformed of a cover audio.

II.3.2.Spatial based techniques

In the time domain based algorithms, the insertion of watermark bits are directly in the samples of signal without using transformation techniques. Time domain based methods are simplest to implement, require less computation and can has a high capacity. On the other hand, it is easy to destroy the watermark. Three methods are associate to this category are Least Significant Bit (LSB) alteration, Echo addition and phase coding methods have been developed[4, 54, 66].

II.4. Conclusion

The chapter 2 separated into two essential parts; the part one presented the techniques used in digital watermarking specially in audio and speech watermarking like transformation techniques, numerical decomposion techniques, the transformation techniques for encryption the watermark and QIM techniques. In parts two, we tried to collect great amount of schemes presented in many papers in the few recent years related in audio and speech watermarking, in the schemes we concentred on transformation approaches, the time domain presented in a few words.

Chapter III

Proposed schemes and evaluation metrics

III.1. Introduction

This chapter include three main parts, the first and the second part gives the new proposed schemes for blind digital speech and audio signals watermarking and the third part gives the various measurements to assess the efficient of the proposed schemes.

In our first proposed scheme, various combinations are used based on DWT and DCT, appending decomposing technique called sub-sampling which it used for watermarking images in [67] and embedding in the norm space, which is a numerical analysis of the linear algebra and can improve the robustness of the algorithm, because the watermark embedded in the norm can be spread throughout all the samples [41]. We also used Arnold transform to encrypt our watermark and grantee the security.

In our second proposed scheme, we segment the speech signal into two segments using sub-sampling technique, and then apply DWT on each segment, followed by DCT to select the part with high energy when we can embed the watermark. Finally the last part provides all measurement and attacks used in the experiments to evaluate our two schemes.

III.2. Blind secured scheme for audio/speech based on DWT-DCT-subsampling-norm Space

Under watermarking terms, the watermark bits must be distributed along the whole speech/audio signal, and for that we decomposed the signal into many segments equal to the number of bits we want to embed, then we apply DWT to extract the approximation coefficients and put the watermark bits there, where the human auditory system is less sensitive. It allowed us making the watermark strong and inaudible with keeping the imperceptibility. And we also applied DCT in order to obtain two vectors having convergent values following it by sub-sampling decomposition into frames for correlation purpose. This decomposition abates a little robustness against the re-sampling attack but gives our proposed design other advantages against other attacks and allows the imperceptibility to remain very high. Extraction is blind in our proposed design, without using original signal. The decomposed speech/audio signal into segments is subjected again to DWT and DCT transforms, then the produced vectors are sub-sampled and normalized before extracting the bits used to construct the image and apply the inverse of Arnold transform using the key used in the embedding process to produce the watermark image (Arnold transform is employed to increase security). The steps below explain more the two processes in fig.7 and fig.8 (embedding and extraction respectively):

III.2.1. Embedding process

Step 1: Insert watermark image $\mathbf{WI_{NxN}}$

Step 2: For the input speech/audio signal **x** decomposed into **N×N** segments;

Step 3: Scramble watermark image $\mathbf{WI_{NxN}}$by Arnold transform using a key and restructure into one dimensional;

W={w(j),1≤ j≤J}, where J=NxN;

For each frame (F_j,1≤ j≤ NxN**) apply the steps (4~12)**

Step 4: Apply 1-level DWT with **'db1'** produces **cA1** and **cD1**

cA: represents the low frequencies (approximation coefficients);

cD: represents the high frequencies (detail coefficients);

Step 5: apply DCT on **cA1** produces vector named **V**;

Step 6: decompose the vector **V** into two (correlated) sub-vectors $\mathbf{V_1}$ and $\mathbf{V_2}$ using the following sub-sampling operations:

$$\mathbf{V}_1(\mathrm{k}) = \mathbf{V}(2\mathrm{k})(14)$$

$$\mathbf{V}_2\ (\mathrm{k}) = \mathbf{V}(2\mathrm{k}-1) \quad (15)$$

Where k=1,., length of V/2.

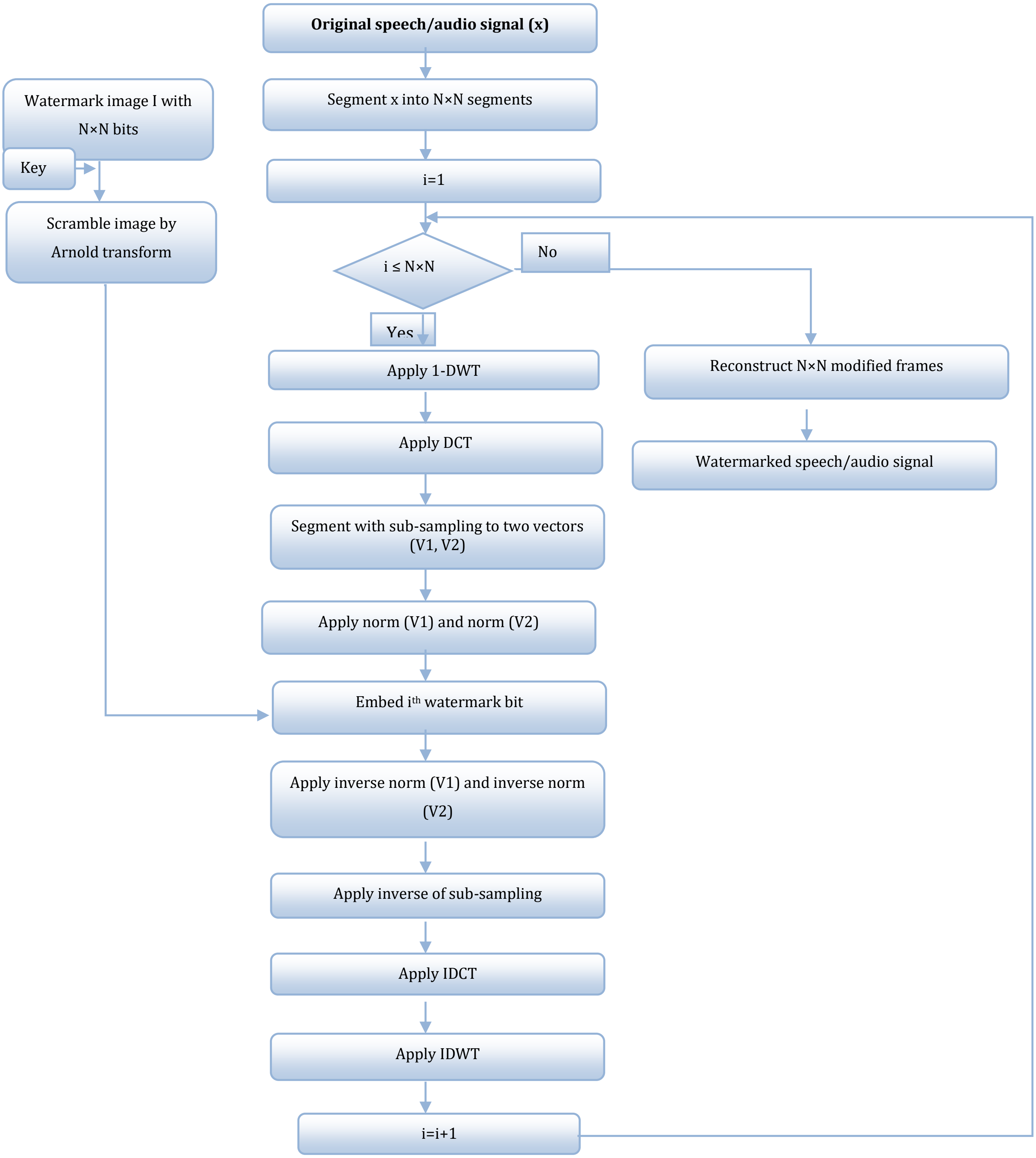


Figure 7: Watermark Embedding Process (DWT-DCT-Subsampling-Norm)

Step 7: apply the norm of $\mathbf{V_1}$ and $\mathbf{V_2}$ produces $\mathbf{nrm_{V1}}$ and $\mathbf{nrm_{V2}}$ respectively as the following formulas:

$$\begin{cases} \mathbf{nrm_{V1}} = \sigma_1 = \|V_1\| = \sqrt{\sum_{i=1}^{n} V(i)_1^2} \quad (16) \\ \boldsymbol{u_1} = \dfrac{\boldsymbol{V_1^t}}{\|\boldsymbol{V_1}\|} = \dfrac{\boldsymbol{V_1^t}}{\boldsymbol{\sigma_1}} \quad (17) \end{cases}$$

$$\begin{cases} \mathbf{nrm_{V2}} = \sigma_2 = \|V_2\| = \sqrt{\sum_{i=1}^{n} V(i)_2^2} \quad (18) \\ \boldsymbol{u_2} = \dfrac{\boldsymbol{V_2^t}}{\|\boldsymbol{V_2}\|} = \dfrac{\boldsymbol{V_2^t}}{\boldsymbol{\sigma_2}} \quad (19) \end{cases}$$

$\mathbf{V_1}$, $\mathbf{V_2}$, $\mathbf{u_1}$ and $\mathbf{u_2}$are a $1 \times n$ vectors, $\boldsymbol{\sigma_1}$ and $\boldsymbol{\sigma_2}$ are the norm of $\mathbf{V_1}$ and $\mathbf{V_2}$ respectively

Step 8: Embedding the bit

$$\mathbf{nrm} = \frac{\mathbf{nrm_{V1}} + \mathbf{nrm_{V2}}}{2} \quad (20)$$

If (**W(j)=**1)

$$\begin{cases} \mathbf{nrm_{V1}} = \text{nrm} + \Delta; \quad (21) \\ \mathbf{nrm_{V2}} = \text{nrm} - \Delta; \quad (22) \end{cases}$$

Else

$$\begin{cases} \mathbf{nrm_{V1}} = \text{nrm} - \Delta; \quad (23) \\ \mathbf{nrm_{V2}} = \text{nrm} + \Delta; \quad (24) \end{cases}$$

End

Step 9: Construct V'_1 and $\mathbf{V'_2}$ with modified norm of each segment as these formula:

$$\boldsymbol{V'_1} = \mathbf{nrm_{V1}} u_1^t \quad (25)$$

$$\boldsymbol{V'_2} = \mathbf{nrm_{V2}} u_2^t \quad (26)$$

Where $\mathbf{u_1}$ and $\mathbf{u_2}$ calculated on the step 7

Step 10: Combine the two sub-vectors $\mathbf{V'_1}$ and $\mathbf{V'_2}$ using the opposite operation in step 6 produce the vector **V'**:

$$\mathbf{V'}\,(2k) = \mathbf{V'_1}\,(k) \quad (27)$$

$$\mathbf{V'}\,(2k-1) = \mathbf{V'_2}(k) \quad (28)$$

Where k=1,..., length of V/2

Step 11: Apply **IDCT** on the modified vector **V'** produces modified approximation **cA1'**;

Step 12: Apply **IDWT** on **cA1'** and **cD1** produces modified frame;

Step 13: Reconstruct the watermarked speech/audio signal with modified frames.

III.2.2. Extracting process:

Step 1: For the input speech/audio signal **x'** decomposed into **N×N** segments;

For each frame (F_j,1≤ j≤ NxN**)**

Step 1: Apply steps (4~7) of the embedding process

Step 2: Extraction of the bit

If ($\mathbf{nrm_{V1}}$>$\mathbf{nrm_{V2}}$)

$$W(j) = 1; (29)$$

Else

$$W(j) = 0; (30)$$

End

Step 3: Construct the image with extracted bits

Step 4: Apply inverse of Arnold transform using key used in the embedding process to produce the watermark image

III.3. Blind scheme for biometric speech watermarking using DWT-DCT-sub_sampling

We can insert watermarks in high energy regions where human auditory system is less sensitive to, such as the low resolution estimation bands. Embedding watermarks in these sections permit us to raise the robustness of our watermark at small to no further impact on image quality [68]. After Discrete Wavelet Transform, most of the speech signal's energies are concentrated in the approximation coefficients and the rest of them are in details coefficients, which means are not lost.

Speech signals are decomposed into low frequency and high frequency with discrete wavelet transform. Low frequency part focuses the majority of the energy of speech signal, which is the most important component of the original signal. cA presents approximate part. High frequency component focuses the small energy of speech signal. cD presents detail part. Wavelet basis and wavelet level can be chosen according to the type of the algorithm [29]. Thus, digital watermarking is extremely flexible in design.

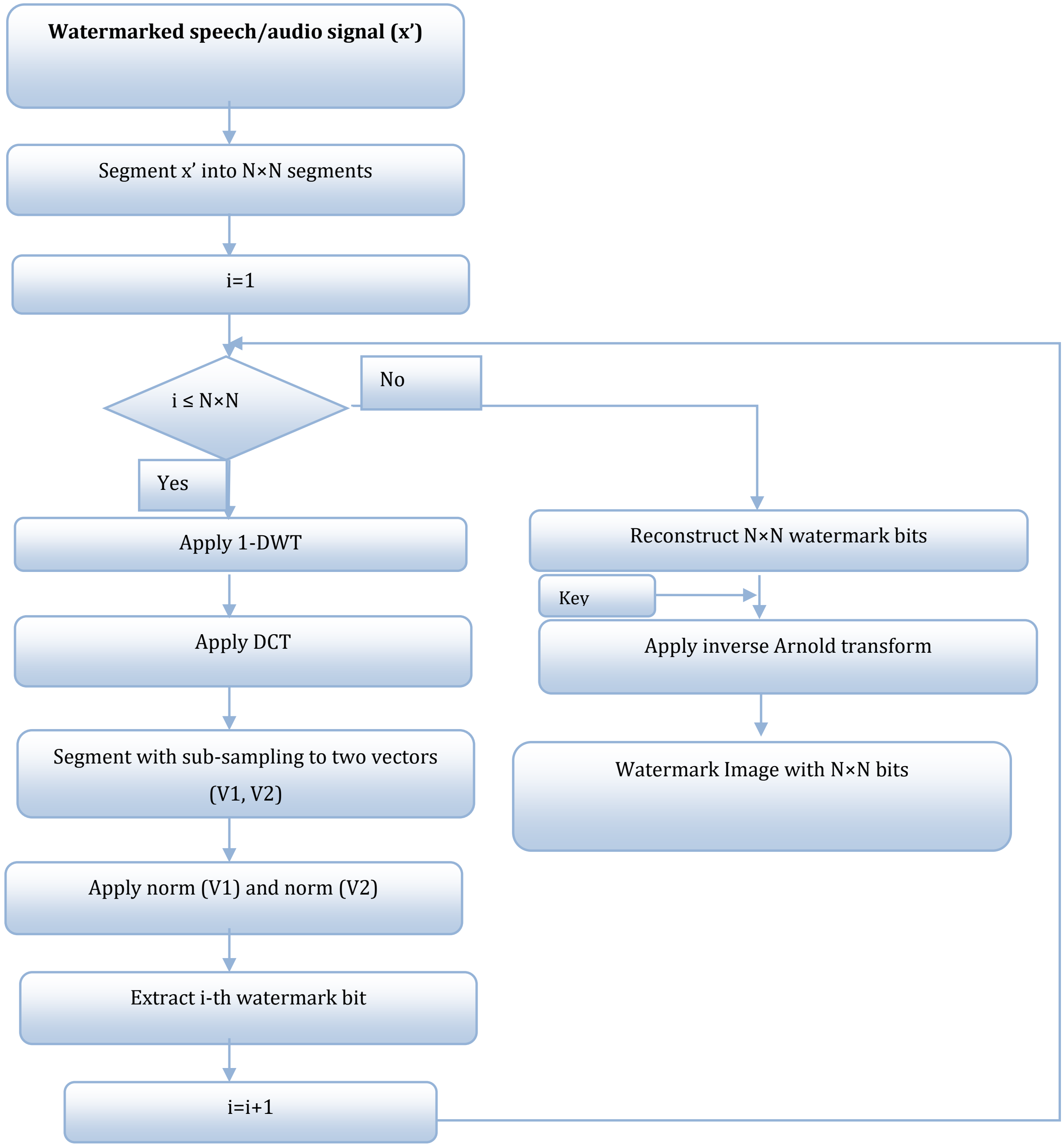


Figure 8: Watermark Extracting Process (DWT-DCT-Subsampling-Norm)

III.3.1. Embedding process

In the proposed scheme, the embedding of watermark image process, Fig.9, is described in the following steps:

Step 1: For the input speech signal, X(n)is decomposed into two segments with sub-sampling as follows:

Seg1: include samples with odd indices; **Seg2:** include samples with even indices;

Seg1={x(1),x(3),x(5),...,};Seg2={x(2),x(4),x(6),...,}; **(Note:** with respecting the arrangements;)

Step 2:applying 1-level DWT with 'db1' of each segment produces:

For seg1: cAseg1, cDseg1; For seg2: cAseg2, cDseg2;

cA: represents the low frequencies (approximation coefficients); **cD**: represents the high frequencies (detail coefficients);

Step 3: Applying DCT on cAseg1 and cAseg2 produces two vectors D1 and D2 respectively;

The insertion of watermark bits is in the DCT coefficient so we apply DCT on cAseg1 and cAseg2 to produce two vectors (D1: DCT coefficient of cAseg1; D2: DCT coefficient of cAseg2)

Step 4:

Insert the watermark image $\mathbf{W_{nxm}}$ and restructure into one dimension vector; $W_i=\{w_i(j), 1\le j\le J\}$, where J=nxm;

Step 5:

- Include a key in order to random the insertion of the watermark image;
- Generate a vector numerated from **1** to **(length of D2)/4**, (for the component with higher energy)
- Random with the introduced Key and generate an additional vector named: **rD**;

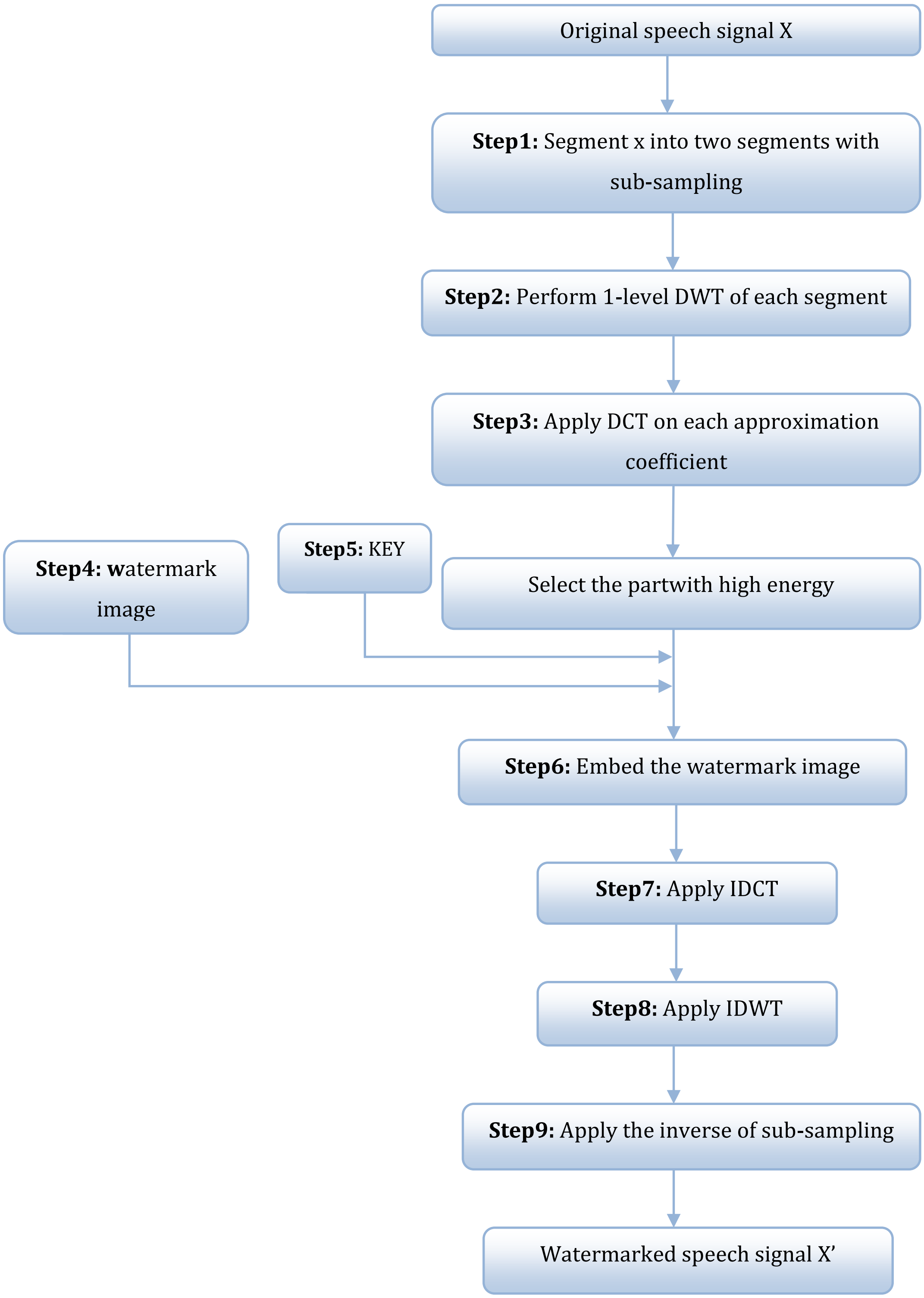


Figure 9: Watermark embedding process (DWT-DCT-Subsampling)

Step 6:

D1 and D2 are modified as follows:

For j=1 to length of watermark (J)

Let $\mathbf{K} = \mathrm{rD}(\mathrm{j})$; (31) $\mathbf{mD} = \frac{\mathrm{D1(K)+D2(K)}}{2}$; (32)

If Wi(j)=1

$$\begin{cases} \mathrm{D1(K)} = \mathrm{mD} + \Delta; & (33) \\ \mathrm{D2(K)} = \mathrm{mD} - \Delta; & (34) \end{cases}$$

Else

$$\begin{cases} \mathrm{D1(K)} = \mathrm{mD} - \Delta; & (35) \\ \mathrm{D2(K)} = \mathrm{mD} + \Delta; & (36) \end{cases}$$

End

End

Step 7:

Applying IDCT on the modified D1 and D2 to get watermarked approximation coefficients.

Step 8:

Applying IDWT on the watermarked approximation coefficients to get modified segments (mseg1, mseg2).

Step 9:

Rebuild the watermarked speech signal with the inverse of step 1 (inverse of sub-sampling);

X'={ mseg1(1),mseg2(1), mseg1(2),mseg2(2), mseg1(3),mseg2(3),...}; (X': watermarked speech signal)

III.3.2. Extraction process

The of watermark image extraction process, Fig.10, is described in the following steps:

Step 1:

We do the steps 1, 2, 3 and 5 on the watermarked speech signal X'.

Step 2:

For j=1 to length of watermark we want to detect

Let $\mathbf{K} = \mathrm{rD}(\mathrm{j})$ (37)

If$(\mathrm{D1}(\mathrm{K}) > \mathrm{D2}(\mathrm{K}))$

$W_i'(j)=1;$ (38)

Else

$W_i'(j)=0;$ (39)

end

End

To illustrate well the working of these steps, we give the following examples for Algorithm explanation:

Step 1:

For the input speech signal x(n) decomposed into two segments with sub-sampling as follow:

Seg1: include samples with odd indices; Seg2: include samples with even indices;

Seg1={x(1),x(3),x(5),...,};Seg2={x(2),x(4),x(6),...,};

Note: with respecting the arrangement;

For example:

The input signal is

x=[0.7,0.03,0.27,0.04,0.09,0.82,0,69,0.31,0.95,0.03,0.43,0.38,0.76,0.79,0.18,0.48]

From which we can get Seg1,Seg2 as follows:

Seg1=[0.03,0.04,0.82,0.31,0.03,0.38,0.79,0.48] ;

Seg2=[0.7,0.27,0.09,0.69,0.95,0.43,0.76,0.18];

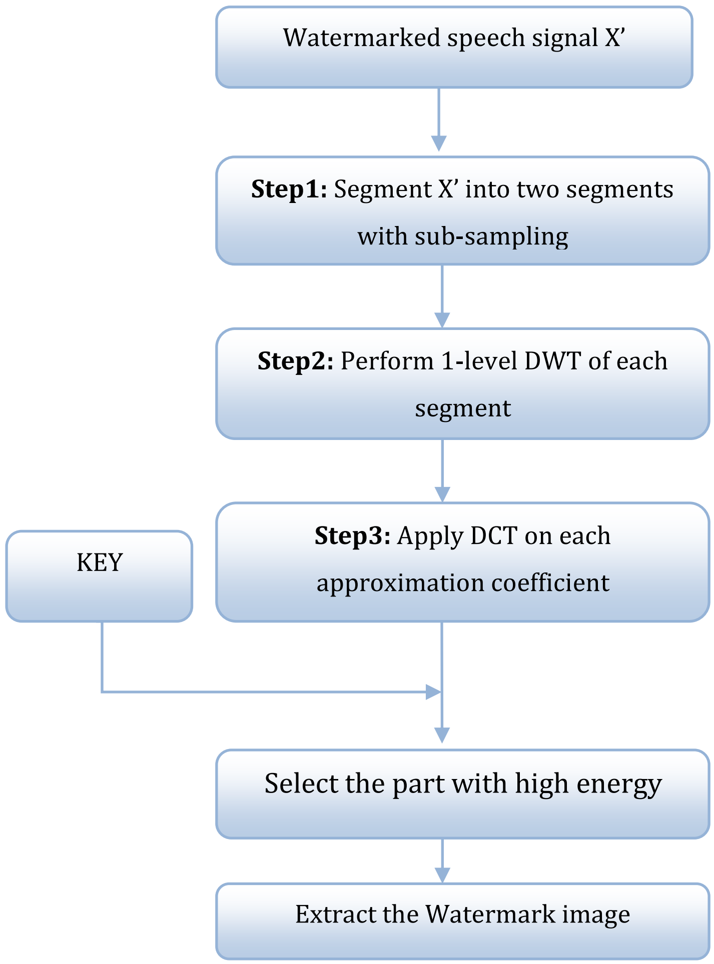


Figure 10: Watermark extracting process (DWT-DCT-Subsampling)

Step 5:

a. Including a key is to random the insertion of the watermark image;
b. Generate a vector numerated from **1** to **(length of D2)/4**; "for the component with high energy"
c. Random with the introduced Key generates an additional vector named **rD**;

For example:

We suppose that the length: D2=28, from which we construct a vector with elements from 1 to $\frac{28}{4}$ ([1 2 3 4 5 6 7])

Using the key introduced in step 5(a), we randomize this vector to produce a random vector for example: rD=[3 2 6 7 4 1 5](randomizing in function with the key value)

Step 6:

D1, D2 modified as follow:

For j=1 to length of watermark (J)

$$\text{Let}\mathbf{K} = \mathrm{rD(j)};\mathbf{mD} = \frac{\mathrm{D1(K)+D2(K)}}{2}; \qquad (40)$$

If Wi(j)=1

$$\begin{cases} \mathrm{D1(K) = mD + \Delta;} & (41) \\ \mathrm{D2(K) = mD - \Delta;} & (42) \end{cases}$$

Else

$$\begin{cases} \mathrm{D1(K) = mD - \Delta;} \\ \mathrm{D2(K) = mD + \Delta;} \end{cases}$$

End

End

<u>For example:</u>

We suppose that the watermark length is 4 (4 bits), then the values that will change (after watermarking) are 4 samples from D1 and 4 samples from D2 selected using the first 4 values of vector rD and following the example of the previous step:

- When j=1 then rD(1)=3 and the first bit is put into the sample D1(3) and D2(3) from the function condition in step 6.
- When j=2 then rD(2)=2 and the second bit is put into the sample D1(2) and D2(2) from the function condition in step 6.
- When j=3 then rD(3)=6 and the third bit is put into the sample D1(6) and D2(6) from the function condition in step 6.
- When j=4 then rD(4)=7 and the second bit is put into the sample D1(7) and D2(7) from the function condition in step 6.

III.4. Evaluation

Evaluation the performance of our watermarking proposals based on three common metrics: Imperceptibility, Robustness, Payload or capacity.

III.4.1. Imperceptibility

Imperceptibility or inaudibility means that watermark embedded into the host signal is inaudible; in this simulation as the majority of this work we use various measurements to assess the quality of the watermarked speech/audio signal. The first is signal-to-noise ratio (SNR) [47] defined as:

$$SNR = 10\log\left(\frac{\sum_{a=1}^{M} S^2(a)}{\sum_{a=1}^{M}\left(S(a) - S'(a)\right)^2}\right) . \qquad (43)$$

The second is the Segmental Signal-to-Noise Ratio (SSNR) [69] which is an improvement with respect to conventional SNR measure and it was created to handle the dynamic nature of non-stationary signals such as speech. The definition of SSNR is:

$$SSNR = \frac{1}{N}\sum_{m=1}^{N} SNR_m \quad . \qquad (44)$$

N is the number of frames in the signal

The SNR does not take into account the specific characteristics of the human auditory system, but it can just give a general idea of imperceptibility [52]. Thus, we also employed one of the most popular methods called mean opinion score (MOS) [45,53,52 and 70] which conducts to provide a better test of inaudibility based on human perception. Ten listeners participated in the practical test and asked to classify the difference between the original and the watermarked speech/audio in terms of 5-points Mean Opinion Score (MOS) with impairment scale defined in Table 1 [52]. To measure the quality of the proposed speech/audio signal, we averaged values of all participants.

Table 1: MOS grading scale

MOS	Description
5	Imperceptible
4	perceptible but not annoying
3	Slightly annoying
2	Annoying
1	Very annoying

III.4.2. Robustness

Robustness is a measure of the resistance of the watermark against attempts to eliminate or corrupt it, intentionally or accidentally, by different kinds of digital signal processing attacks. For the evaluation of robustness, this simulation examines the bit error rates (BER) between the original watermarking image and the extracted watermarking image. BER is defined by the following expression [70]:

$$BER = \frac{B_{ERR}}{N} \times 100\% \quad . \qquad (45)$$

Where B_{ERR} is the number of erroneous bits and N is the total number of bits

Zero means that the attack doesn't have any effect on the watermark and the extraction is successful. Also we employed normalized correlation coefficient (NC) which expresses the similarity between extracted watermarking image and original watermarking image after being attacked and it is defined by the following expression [71]:

$$\mathrm{NC}(\mathrm{w}, \mathrm{w}') = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} w(i,j)w'(i,j)}{\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N} w^2(i,j)}\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N} w'^2(i,j)}} \quad (46)$$

Where N*N is the size of watermark. *W(i,j)* and *W'(i,j)* are the watermark and recovered watermark images, respectively. One is the best value for NC and it shows that the inserted watermark is extracted successfully.

In order to test the robustness of the proposed algorithm, separately we attack the watermarked version using typical signal processing manipulations

a) **AWGN:** Add white Gaussian noise to the vector watermarked speech/audio signal, measuring the power of the audio-speech before adding noise.
b) **Re-sampling:** The watermarked speech/audio was down-sampled to half the original sampling rate and then up-sampled back to the original sampling rate.
c) **Re-quantization:** 16 bits per sample watermarked speech/audio signals is quantized down to 8 bits per sample.
d) **Echo:** We add an echo signal with a different delay and decay of to the watermarked speech/audio signal.
e) **Amplification:** The amplitude of the watermarked speech/audio signal is rescaled by ±10%,±15%, ±20% and 30%.
f) **Cropping:** We set the number of samples of the watermarked speech/audio signal to zero randomly.

III.4.3. Capacity

Data payload is identified as the number of bits embedded in a one-second audio part [10], and is measured in bits per second (bps). Assume that S the length of the original speech signal in seconds and K is the amount of embedded watermark bits, the capacity of the proposed scheme C is expressed as [36]:

$$C = \frac{K}{S} bps \qquad (47)$$

III.5. Conclusion

Third chapter divided into three main parts, we gave new proposed schemein the first part, the method introduced was blind, based on DWT and DCT transformation and employed sub-sampling technique, the embedding of watermark in norm space. The scheme also used Arnold transform for encryption the watermark. The part two introduced other blind method based on hybrid DWT/DCT and used sub-sampling. The last part included all measurement and attacks used in the experiments to assess the performance of our two schemes.

Chapter IV

Blind secured scheme for audio/speech based on DWT- DCT- sub_sampling- Norm Space results

IV.1. Introduction

This chapter presents all results and discusses on it, whereas all simulations are implemented on Windows PC having Intel 2.2GHz processor and 2GB RAM. All the experiments are performed using MATLAB 7.10.0 on different speech/audio signals which are stored as 16 bit mono wave file, and frequency 44100 Hz.

In order to evaluate the performance of the proposed scheme in real conditions, simulations are performed on different lengths of speech/audio signals included and also different types of human speech signals (male and female) and different languages (English and French).

All of the digital speech/audio files are downloaded from reference [72], SQAM file (Sound Quality Assessment Material) recording for subjective tests. We edit the speech/audio file to change stereo to mono and we use two binary images as watermarks (UZAD image which it used in all experiments and star image which it used only in the experiments results in tables 4, 5 and 6), Fig.11, Fig.12 show them, respectively :

a) Original b) Scrambled image

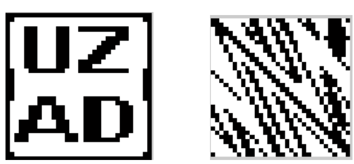


Figure 11: Watermark image (UZAD)

a) Original b) Scrambled image

Figure 12:Watermark image(STAR)

IV.2. Imperceptibility

Tables 2 and 3 show values of different measurements for different speech/audio signals results from our proposed method (DWT, DCT, Sub-Sampling, Norm Space, Arnold), so it is clear that the SNR satisfy the requirement of international federation of the phonographic industry (IFPI) with the SNR above 20 db, and it can be up to 30 db which means that our proposed scheme can get better perceptual quality than the previous methods. In addition, we can see that the SSNR is greater than the SNR which means that there is no camouflage.

However, the values of MOS resulting from our proposed method are high, which indicates that the watermarked speech and audio signals are perceptually indistinguishable from the original ones.

Table 2: SNR, SSNR and MOS of Speech type signal

Speech	SNR	SSNR	MOS
spme50_1	29,7432	35,2420	4,4
spmf52_1	30,2990	35,1564	4,6
spfe49_1	30,5078	35,4074	4,6
average	30,1833	35,2686	4,53

Table 3: SNR, SSNR and MOS of Audio type signal

Audio	SNR	SSNR	MOS
bass47_1	30.0425	35.5654	4,7
gspi35_2	32.1148	33.5571	4,8
average	31,0786	34,5612	4,75

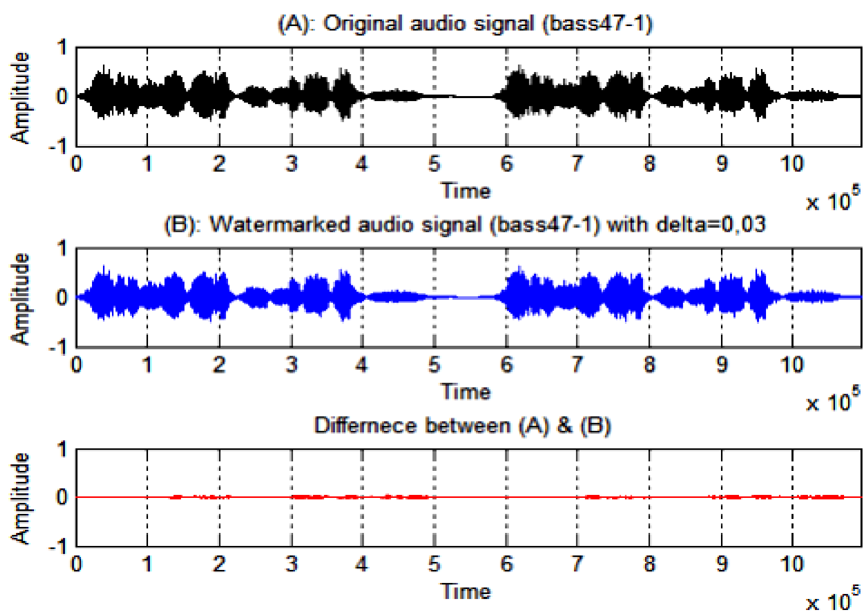


Figure 13: Waveforms of the original and watermarked audio (bass47_1) and difference between them

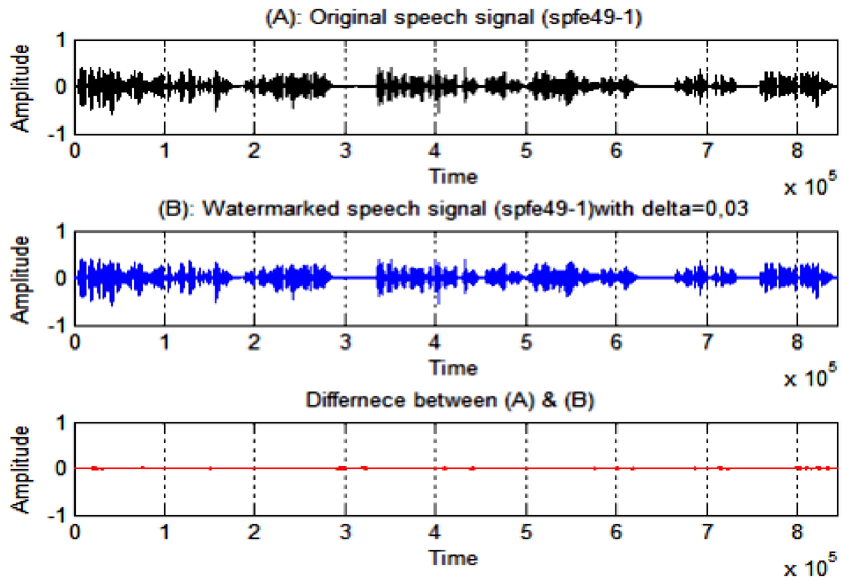


Figure 14: Waveforms of the original and watermarked speech (spfe49_1) and difference between them

Fig.13 illustrates the time waveforms of the original and watermarked audio signal and differences between them respectively, which present the inaudibility by our algorithm. It can be seen that there is only a little visual difference which indicates that our algorithm possesses good transparency.

By observing the waveforms in Fig.14 of the original speech signal (A) and the watermarked version (B) and the difference between them, we can conclude that there is almost no difference.

Fig.15 shows the SNR and SSNR versus the Δ (quantization step) for audio and speech signal (the left: spfe49_1 speech and on the right: gspi35_2 audio). As seen, whenever Δ increases, SNR and SSNR decrease. This is because the norm values are far from their original state (where the bits are embedded), and thus there are a distortions in the original speech/audio signals. Also we can observe that the values of SSNR didn't come down inferior the values of SNR and always stay on up which indicates that there is no camouflage using the process of embedding the watermark.

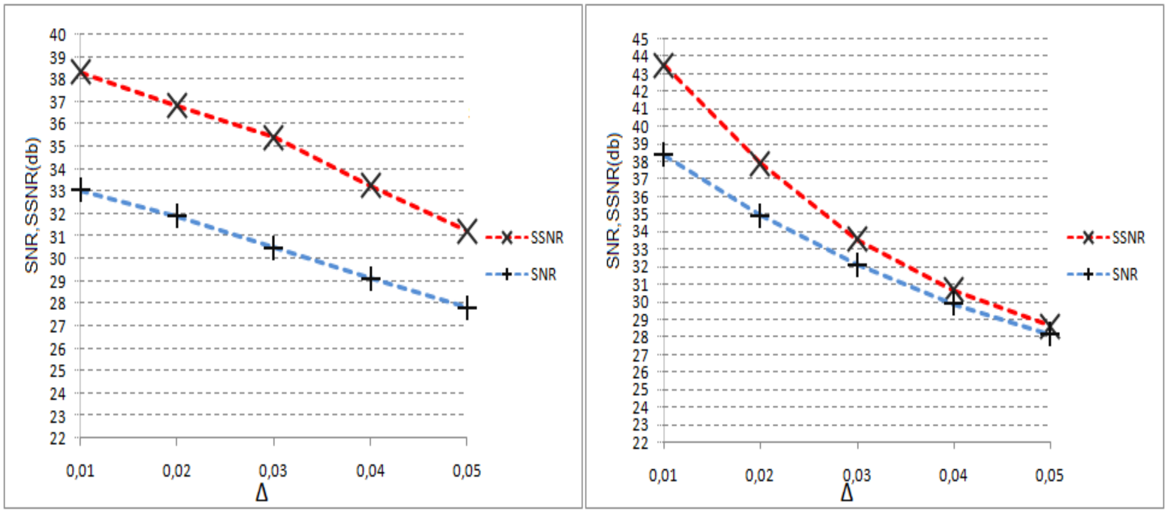


Figure 15: SNR and SSNR versus the Δ for audio and speech signal (on the left: spfe49_1 speech and on the right: gspi35_2 audio)

IV.3. Robustness

Table 4: Results of robustness against different type of signal processing attacks for audio signal (bass47_1)

The attacks		**Watermark images**					
		UZAD			STAR		
		SNR between WAS and AWAS	BER %	NC	SNR between WAS and AWAS	BER %	NC
Without attacks		Inf	00	1	Inf	00	1
AWGN		18.0719	00	1	18.0062	00	1
Echo (0.13,0.33)		17.5284	00	1	17.4828	00	1
Resampling		40.7000	5.0781	0.9595	41.6001	4.5898	0.9544
Re-quantizaton		31.5877	00	1	31.5842	00	1
Cropping (10000)		20.3075	00	1	20.2556	00	1
Amplification	+20%	19.8671	00	1	20.7071	00	1
	-20%	20.7070	00	1	19.8670	00	1

Table 5: Results of robustness against different type of signal processing attacks for speech signal (spme50_1)

The attacks		Watermark images					
		UZAD			STAR		
		SNR between WSS and AWSS	BER %	NC	SNR between WSS and AWSS	BER %	NC
Without attacks		Inf	00	1	Inf	00	1
AWGN		18.0519	00	1	18.0042	00	1
Echo (0.15, 0.32)		12.1942	00	1	12.2625	00	1
Resampling		34.7870	4.9805	0.9603	35.0483	4.4922	0.9554
Re-quantizaton		31.5584	00	1	31.5546	00	1
Cropping (10000)		19.0726	00	1	18.9774	00	1
Amplification	+20%	21.2669	00	1	21.9869	00	1
	-20%	21.9868	00	1	21.2668	00	1

Table 6: Results of robustness against different type of signal processing attacks for speech signal (spmf52_1)

The attacks		Watermark images					
		UZAD			STAR		
		SNR between WSS and AWSS	BER %	NC	SNR between WSS and AWSS	BER %	NC
Without attacks		Inf	00	1	Inf	00	1
AWGN		18.0614	00	1	18.0050	00	1
Echo (0.12, 0.3)		16.1849	00	1	16.6254	00	1
Resampling		30.2428	4.9805	0.9603	30.3407	4.4922	0.9554
Re-quantizaton		32.0982	00	1	32.0967	00	1
Cropping (3000)		24.9535	00	1	24.6027	00	1
Amplification	+20%	22.6162	00	1	23.2362	00	1
	-20%	23.2361	00	1	22.6161	00	1

Table 4, Table 5 and Table 6 show the robustness of our proposed method using different audio and speech signals (bass47_1, spme50_1 and spmf52_1) without attack and with various attacks. The low SNR between watermarked speech/audio signal (WSS/WAS) and attacked watermarked speech/audio signal (AWSS/AWAS) demonstrates that the majority of attacks used for evaluation of the robustness were very strong such as: AWGN, adding Echo, cropping and amplification attacks. However the majority of the BER values are zeros and the majority of NCs values are ones which means that the process of detection can detect the inserted watermark successfully. It indicates that the watermark system adopted has good robustness performances. So that all attacks can't degrade the watermark except in re-sampling attack, but that's not a problem because the BER is low in this situation and we can still identify our watermark.

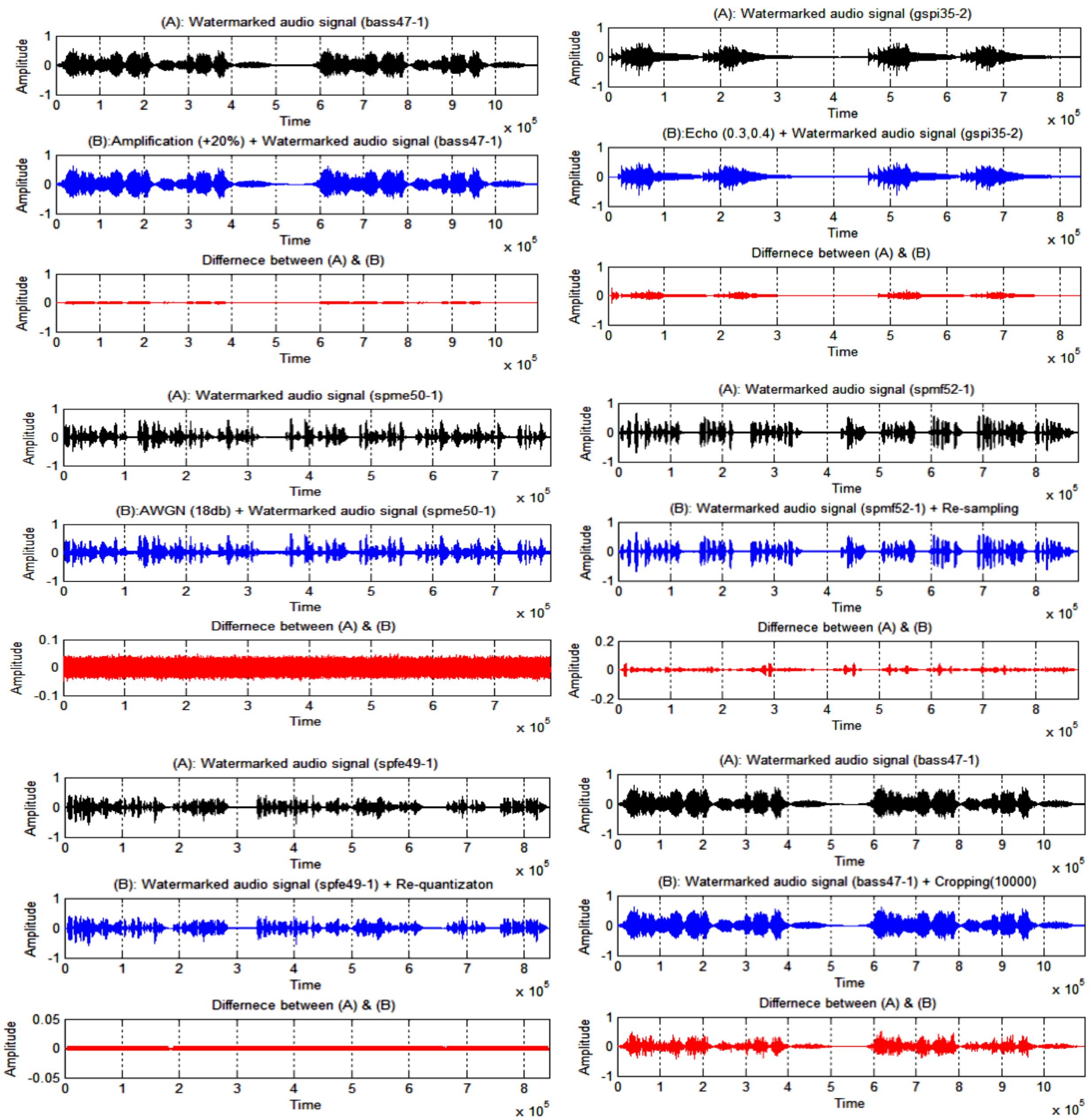


Figure 16: The used different attacks and their effects on original watermarked signals

In Fig.16, we can observe that the attacks used are very strong and effects on the signal. This figure explains more the strong attacks used so that there exists a little difference by the attacks: re-quantization and re-sampling. The difference is noticeable in the attack of amplification and AWGN. Big differences are observed in the echo and cropping attacks between watermarked between watermarked speech/audio signal and the attacked speech/audio signal.

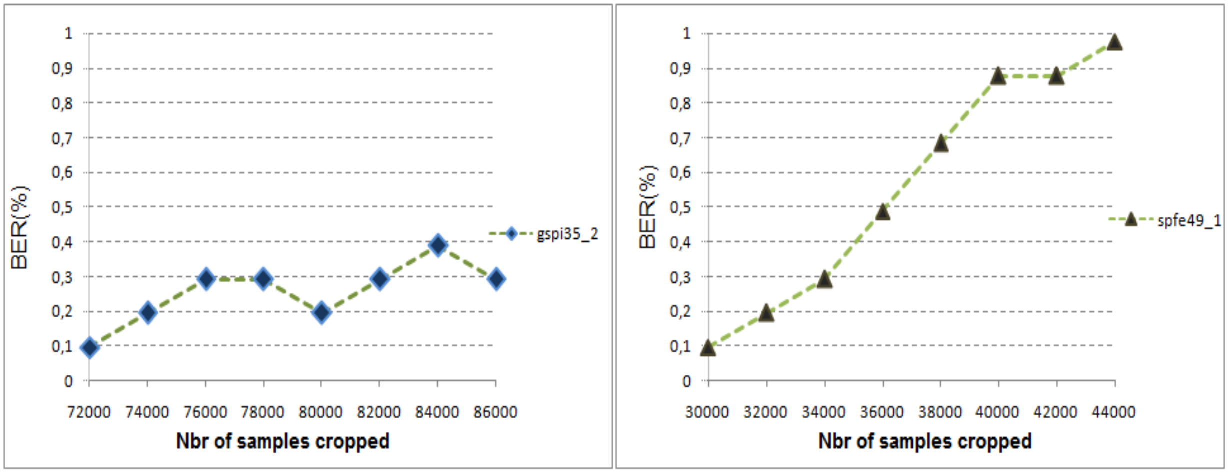


Figure 17: BER vs cropping for audio-speech signal (on the left gspi35_2 audio, on the right spfe49_1 speech)

Fig.17 illustrates the BER values versus increasing number of samples that are cropped in the audio and speech signals. BER remains small under 1% although thousands of samples were set as zero randomly. Although the cropping was changed by 14 thousands cropped samples, the BER remains small and did not exceed 1%.

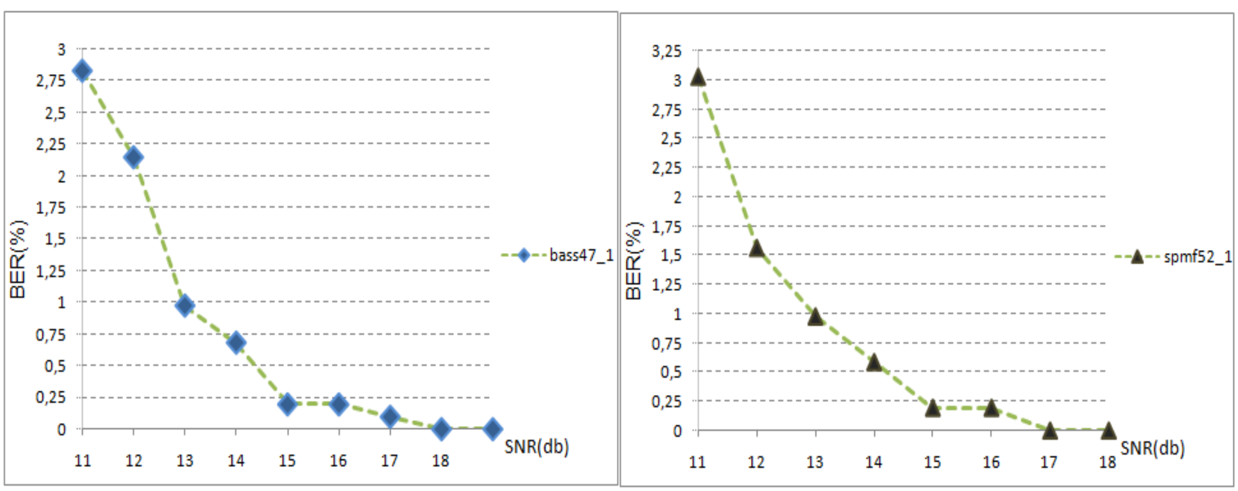


Figure 18: BERs vs AWGN attacks for audio-speech signal (on the left bass47_1 audio, on the right spmf52_1 speech)

Fig.18 shows the BER after different SNR of AWGN attacks. Although all of these attacks are strong and influential on the signal significantly, BER is small at SNR=11db (<3%) and null at SNR=18db. This confirms the robustness of the watermark inserted in speech/audio signal. The lower the strength of AWGN SNR, the more obvious is the watermark.

IV.4. Capacity

The capacity is not too high as shown in table 7, but it is sufficient as the conditions of IFPI are set to 20b/s a satisfied because the goal is reached, the watermarking is very robust and high imperceptibility is attained.

Table 7: Capacity measures for different audio and speech signals

Audio/Speech	**bass47_1**	**gspi35_2**	**spme50_1**	**spmf52_1**	**spfe49_1**
capacity	41.19	53.87	57.03	51.17	53.36

IV.5. Comparisons

From the comparison results in Table 8, we can see that our proposed (DWT, DCT, Sub-sampling, Norm-space, Arnold) scheme can obtain a relatively high imperceptibility and good payloads results, since SNR and MOS results are higher than almost all other published methods selected for comparison. It demonstrates the preference for our scheme. Besides, the payload in our scheme is lower than in [73] and [63] but, it is relatively high compared to the other selected methods.

Table 8: Summary of comparisons with seven methods cited in literature

Methods	**Average of SNR (db)**	**Capacity b/s**	**Type**	**Average of MOS**
DWT-SVD in [29]	20,7	27,56	Speech	4,4
	21,2		Audio	4,65
SVD-AQ in [73]	30,3	172,39	Audio	-
DWT-AMM in [63]	21,932	200	Speech	3,25
CCCD in [74]	25,777	49	Speech	-
DWPT-Multiplication in [59]	28,08	31,25 -125	Speech	4,11
Adaptive DWT SVD in [51]	24,37	45,9	Audio	4,46
Method in [10]	30,0675	17,2	Audio	-
Our proposed scheme (DWT- DCT- Sub sampling - Norm space – Arnold)	31,0786	41.19-53.87	Audio	4,53
	30,1833	51.17-57.03	Speech	4,75

Table 9: Comparison between our proposed scheme and scheme in reference [41] for Audio signal

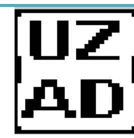
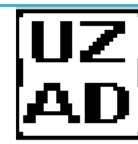
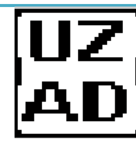
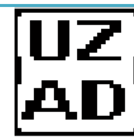
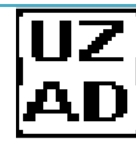
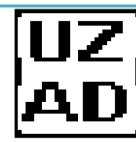
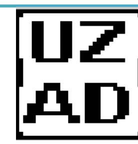
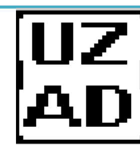
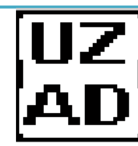
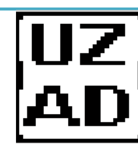
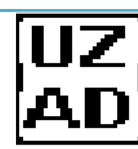
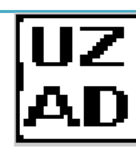
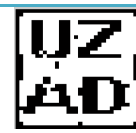
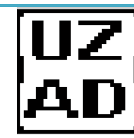
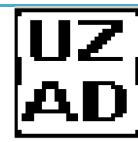
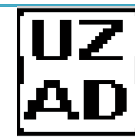
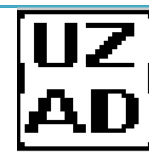
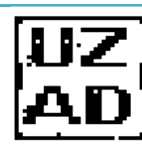
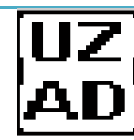
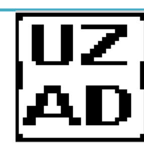
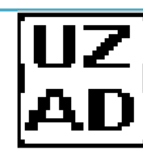
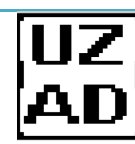
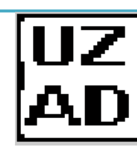
Audio	attacks	Factor (power)	BERs of		NCs of		Detected watermark	
			Scheme in [41]	Proposed scheme	Scheme in [41]	Proposed scheme	Scheme in [41]	Proposed scheme
gspi35_2	AWGN	18 db	00	00	1	1		
	Re-sampling	44100-22050-44100 Hz	00	5.5664	1	0.9558		
	Re-quantization	16-8-16 bits	00	00	1	1		
	Echo	(0.1,0.4)	00	00	1	1		
		(0.3,0.4)	8.6914	00	0.9274	1		
	Amplification	+15%	26.1719	00	0.7591	1		
		-15%	33.4961	00	0.6764	1		
	Cropping	30000	0.8789	00	0.9928	1		
		70000	45.8984	0.0977	0.7447	0.9992		

Table 10: Comparison between our proposed scheme and scheme in reference [39] for Speech signal

Speech	attacks	Factor (power)	BERs of		NCs of		Detected watermark	
			Scheme in [39]	Proposed scheme	Scheme in [39]	Proposed scheme	Scheme in [39]	Proposed scheme
spfe49_1	AWGN	18 db	1.9531	00	0.9841	1		
	Re-sampling	44100-22050-44100 Hz	34.3750	5.1758	0.7026	0.9586		
	Re-quantization	16-8-16 bits	00	00	1	1		
	Echo	(0.1,0.2)	16.6016	00	0.8608	1		
	Amplification	+10%	1.0742	00	0.9913	1		
		-10%	00	00	1	1		
	Cropping	10000	3.5156	00	0.9713	1		
		20000	7.1289	00	0.9414	1		

Authors in [41] and [39] proposed blind watermarking schemes for the audio and speech signals. We compared our proposed design with these published schemes.

Table 9 and Table 10 summarize the comparisons between our proposed watermark detection results and results of schemes in [41] and [39] against various attacks. We observe that the robustness of embedded watermark in our design is better than the embedded watermark in schemes of [41] and [39].

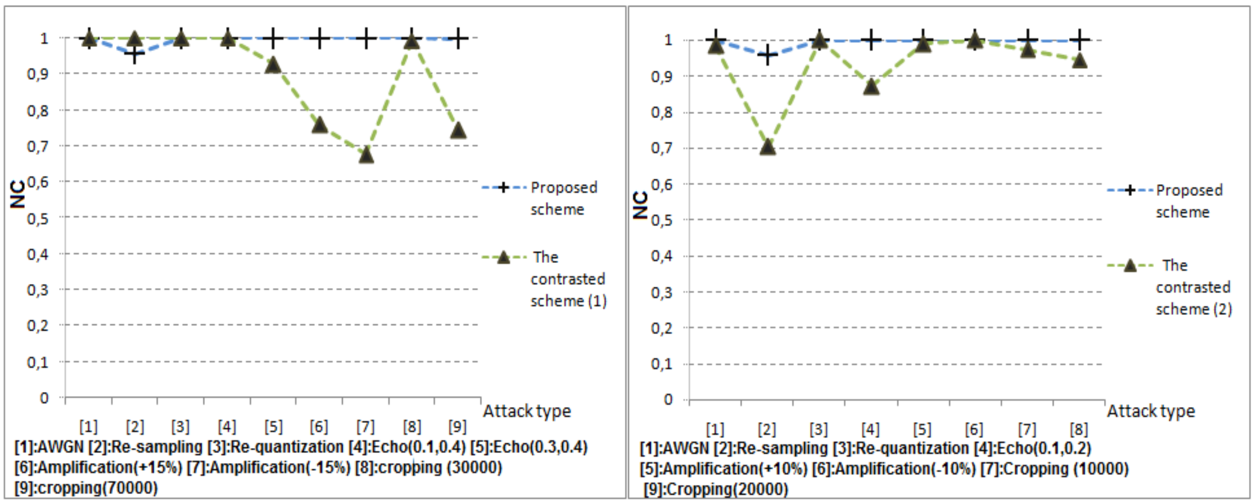


Figure 19: Efficiency comparison between the proposed scheme and other two schemes: the contrasted scheme (1) in [41], and the contrasted scheme (2) in [39]

In Fig.19, the two graphs illustrated well comparison results between our proposed scheme and the two published schemes in references [41] and [39]. Under nine (9) signal processing attacks types, we observe the steady robustness of our proposed design against all strong attacks. Advantages of our proposed design are resumed as:

- It is more robust than the schemes in [41] and [39].
- Our SNR is greater than the SNR determined from scheme of [41] which means better imperceptibility.
- Extraction is blind in our proposed design, without using original signal.
- Extracting without using parameter Δ (the Δ used in the embedding process).
- We can apply both on speech signals and audio signals.

IV.6. Conclusion

The new blind scheme for speech and audio signals watermarking based on DWT, sub-sampling, DCT transform and the embedding in the vector norm was evaluated in this chapter. We performed all necessary experiments to ensure the efficiency as well as the fully blind detection is accomplished without using the original speech/audio signal and the insertion parameter is not required. The proposed design, compared to other schemes presented in literatures, makes an excellent tradeoff between security, capacity, imperceptibility and robustness against signal processing attacks at random payload for different types of audio/speech signals. The decomposing with sub-sampling abates a little robustness against the re-sampling attack but gives our proposed design other advantages against other attacks and allows the imperceptibility to remain high.

Chapter V

Blind scheme for biometric speech watermarking using DWT-DCT-sub_sampling results

V.1. Introduction

This chapter provides the results of proposed scheme based on sub-sampling, DWT and DCT. The chapter presents efficiently study of the proposed scheme concerning imperceptibility, robustness, capacity also execution speed. On the other hand the comparisons were achieved. The same PC and Matlab version mentioned in precedent chapter used in these experiments, simulations are performed on different lengths of speech signals including different natures of signals (male and female) and different languages (English, French, German). All of the speeches are downloaded from [24] SQAM (Sound Quality Assessment Material Recordings for Subjective Tests) file. We edit the speech file to change lengths. Table 11 represents the speeches used in our experiments. Also we saved biometric images (fig.20) from [75] represent digital fingerprints then resized to fitted size and convert from greyscale to binary images. The binary images used as watermarks, the different watermarks embedded within different speech signals. Also we employed other binary image (fig .21) in the comparisons part.

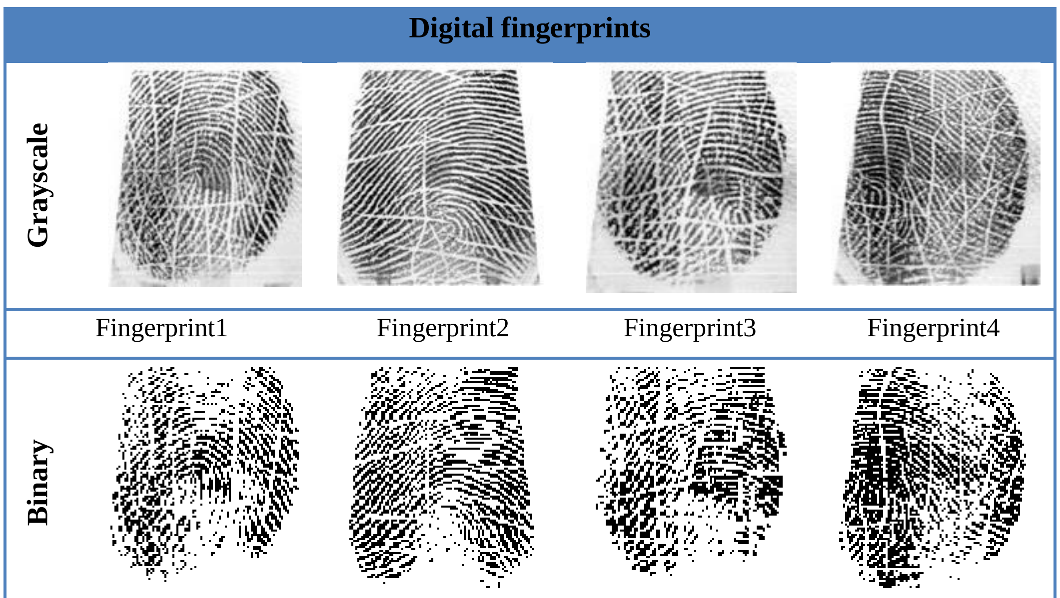


Figure 20: Different fingerprints (115×99 bits)

Table 11: Speech properties

name	type	Mono/stereo	Nbr bits	Frequency	Length (seconds)	Man/Woman	Language	Fingerprint (watermark)
spfe49_1	wav	Mono	16	44100 Hz	19.187	Woman	English	Fingerprint1
spme50_1	wav	Mono	16	44100 Hz	16.857	Man	English	Fingerprint2
spfg53_1	wav	Mono	16	44100 Hz	16.537	Woman	German	Fingerprint3
spmf52_1	wav	Mono	16	44100 Hz	20.01	Man	French	Fingerprint4
bass47_1	wav	Mono	16	44100 Hz	24.860	Man	Unknown	Fingerprint1

UZAD image

Figure 21: Watermark image (32x32 bits)

V.2. Imperceptibility

Fig.22 represents the evolution of the SNR values with different parameters Δ and demonstrates there is a counter proportionality. Table 12 gives the accurate values for quantitative evaluation for different speech signals with Δ variation. All of the SNR values superior to the minimum value imposed by IFPI (20db).

Fig.23 represents the evolution of SNR with variations of speech signals lengths. Table 13 gives the accurate values for quantitative evaluation. The SNR values of our proposed scheme are increasing with the increase of speech signals length because of the distortion become little.

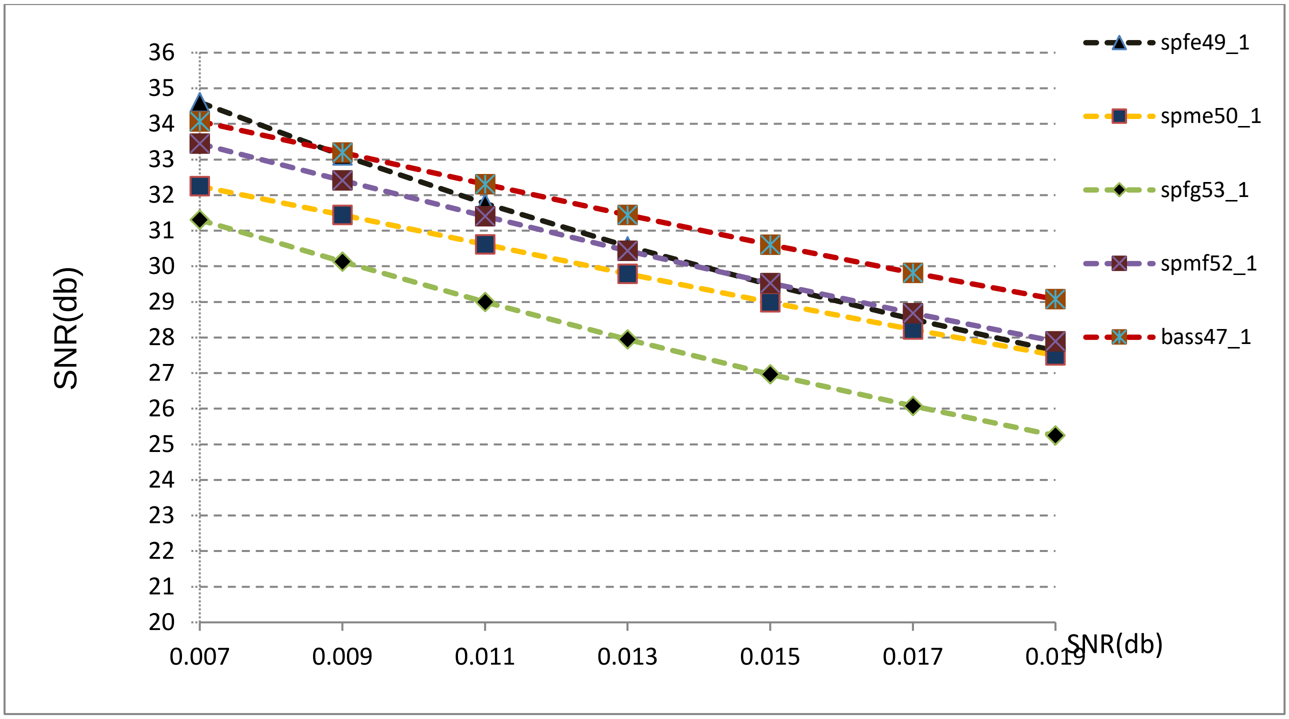


Figure 22:SNR in function with Δ

Table 12: SNR evolution with variation of Δ for different speech signals

Δ	spfe49_1	spme50_1	spfg53_1	spmf52_1	bass47_1
0,007	34,605	32,2467	31,3025	33,4421	34,0692
0,009	33,1133	31,443	30,1248	32,4146	33,1955
0,011	31,7675	30,6104	28,9958	31,4011	32,3068
0,013	30,5631	29,7846	27,9419	30,434	31,4376
0,015	29,4828	28,9855	26,9685	29,5258	30,6052
0,017	28,5083	28,2226	26,0719	28,6785	29,8168
0,019	27,6235	27,4993	25,2452	27,8895	29,0739

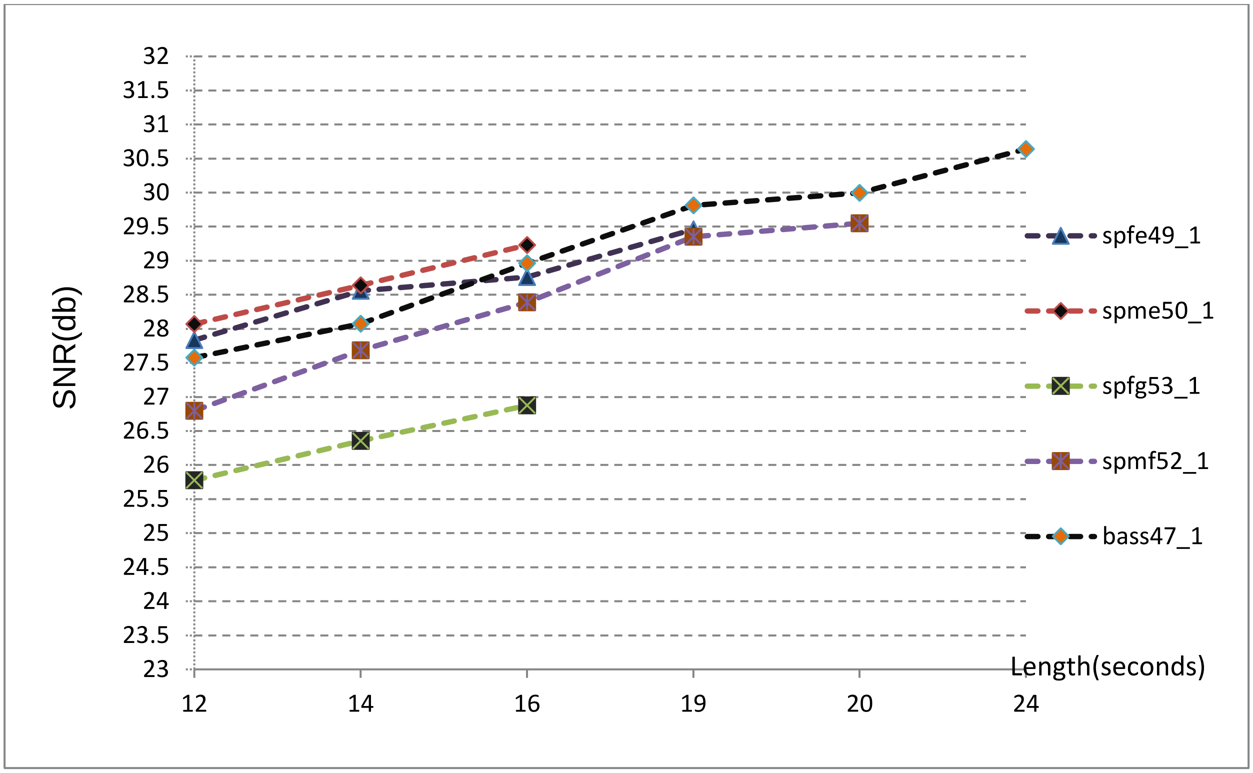


Figure 23: SNR in function with speech signals length

Table 13: SNR versus length signals

length(seconds)	spfe49_1	spme50_1	spfg53_1	spmf52_1	bass47_1
12	27,8334	28,0703	25,7751	26,7974	27,5772
14	28,5575	28,6397	26,3541	27,6819	28,0754
16	28,7585	29,23	26,8744	28,3867	28,9631
19	29,4596	-	-	29,3519	29,8111
20	-	-	-	29,5497	29,9978
24	-	-	-	-	30,6402

The table 14 show that the imperceptibility evaluated with two aspects subjective evaluation test (MOS) and objective evaluation test (SNR), the all listeners can’t find any difference between watermarked and original versions of speech signal which confirm and authenticate the values obtained by objective evaluation test.

Table 14: Imperceptibility with MOS with Δ=0.03

Values of	spfe49_1	spme50_1	spfg53_1	spmf52_1	bass47_1
SNR	27.2104	27.6872	25.3668	27.9902	28.7193
MOS	5.0	5.0	5.0	5.0	5.0

V.2. Robustness

In order to evaluate the robustness of the suggested scheme, many attacks were applied including: additive noise (AWGN), re-quantization, cropping, amplification and adding Echo. All experiment of robustness test based on Δ=0.03.

V.2.1. AWGN attack

Table 15 presents that the different speech signals attacked with different powers of AWGN attack and illustrates the SNR values between watermarked signals and attacked watermarked signals. Although the power of attack is large, almost of the BER values are zeros and the NC values are 1. For that we can state that our proposed scheme is robust for AWGN attacks. Fig.24 illustrates the watermarked signal (spfe49_1), attacked watermarked signal and the difference between them. It demonstrates that the AWGN attack used is big.

Table 15: Different speech segments attacked with different AWGN

signal	awgn snr (db)	SNR between WS & AWS	BERs	NCs
spfe49_1	18	17.9894	00	1
spme50_1	24	24.0048	3.5222	0.9742
spfg53_1	17	16.9875	00	1
spmf52_1	18	18.0070	00	1
bass47_1	16	18.0069	00	1

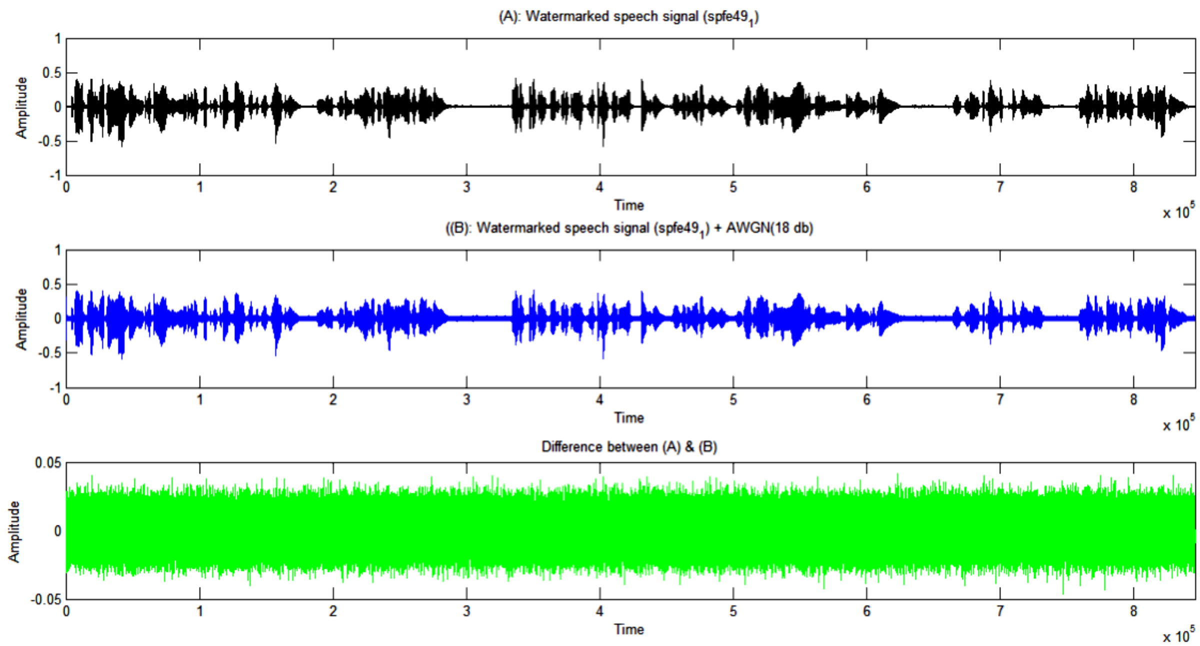


Figure 24: Original speech signal and watermarked speech signal attacked with AWGN and the difference between them

V.2.2. Re-quantization attack

Table 16 shows the SNR values between the watermarked speech signals and the quantized watermarked speech signals. Also it shows that almost of the values of BER are zero and value of NC are one after the attack. Fig.25 shows the watermarked speech signal (spme50_1), quantized watermarked signal and difference between them and illustrates that the difference is small.

Table 16: SNR BETWEEN THE WATERMARKED SIGNAL AND ITS QUANTIZED VERSION

signal	SNR between WS & QWS	BERs	NCs
spfe49_1	30.2550	00	1
spme50_1	31.7286	3.4431	0.9759
spfg53_1	29.3259	00	1
spmf52_1	31.3381	00	1
bass47_1	31.2944	00	1

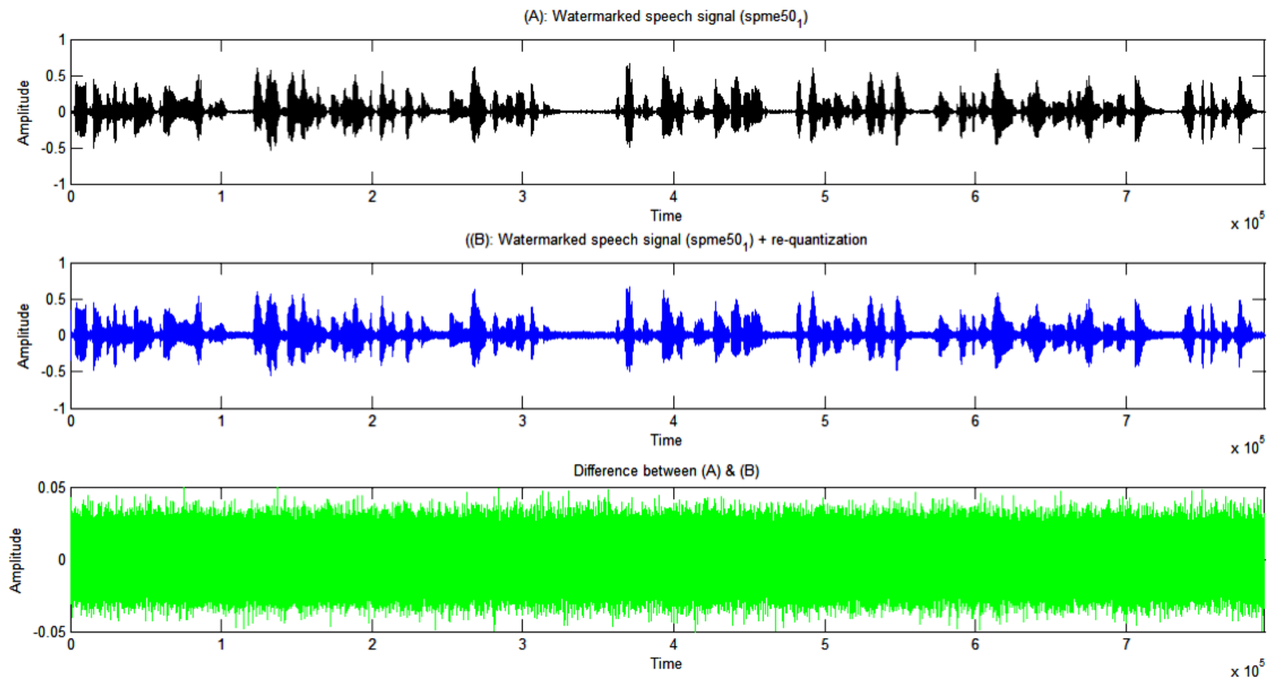


Figure 25: The difference between watermarked speech signal and its quantized version

V.2.3. Cropping attack

Table 17 illustrates the values of SNR between the watermarked speech signals and the cropped watermarked speech signals and the number of samples set randomly to zero. It shows the majority of the BER values are zero and NC values are 1. Even though the attack is very strong, we can identify our watermark without difficulty. Fig.26 illustrates the watermarked speech signal (spfg53_1), cropped watermarked signal and the difference between them. It also shows that the difference is very large.

Table 17: SNR between watermarked speech signals and its cropped version

signal	SNR between WS & CWS	Nbr of cropped samples	BERs	NCs
spfe49_1	16.1605	21000	00	1
spme50_1	17.3147	15000	4.7870	0.9663
spfg53_1	15.5859	20000	00	1
spmf52_1	16.9426	18000	00	1
bass47_1	17.4578	16000	00	1

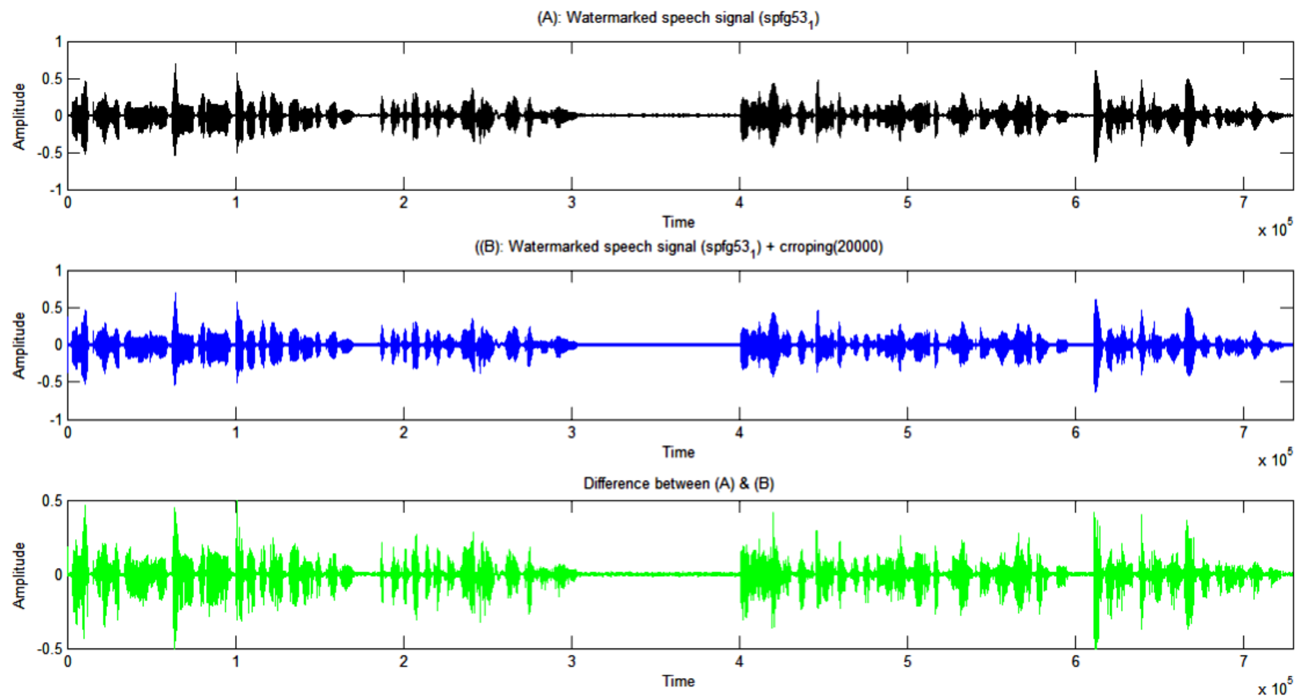


Figure 26: The difference between watermarked speech signal and its cropped version

V.2.4. Echo attack

We add an echo signal with a different delay and decay of to the watermarked speech signal. Table 18 is shows the SNR values between watermarked speech signals and watermarked speech signals with echo, and illustrates that almost of the values of BER are zero and value of NC are 1. Although the attack is very strong (the SNR between WS & EWS), we can detect our watermark easily. Fig.27 illustrates the watermarked speech signal (spmf52_1), watermarked signal with echo and the difference between them which is very big.

Table 18: SNR between watermarked speech signals and WS with added echo

signal	SNR between WS & EWS	Echo(delay, decay)	BERs	NCs
spfe49_1	6.3633	0.4,0.6	00	1
spme50_1	10.6253	0.2,0.4	4.5411	0.9680
spfg53_1	8.2686	0.3,0.5	00	1
spmf52_1	6.9247	0.4,0.6	00	1
bass47_1	12.9082	0.2,0.6	00	1

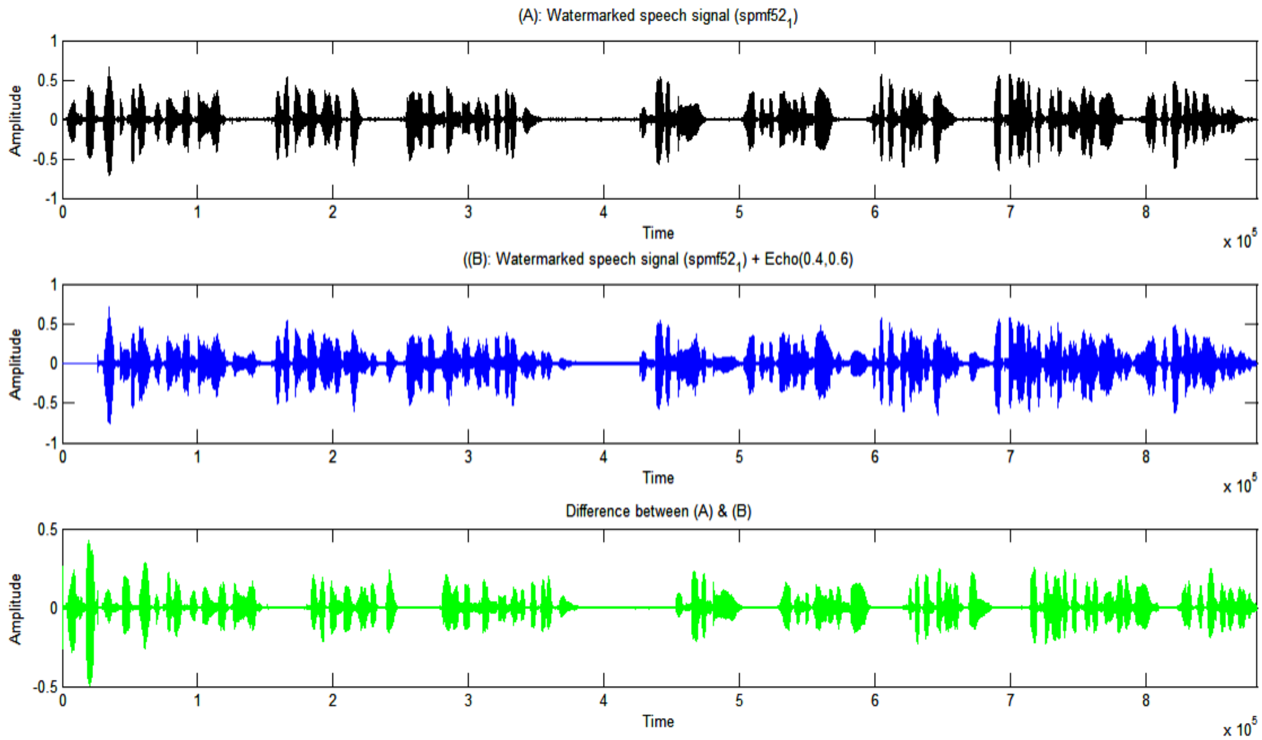


Figure 27: The difference between watermarked speech signal and WSS with added echo

V.2.5. Amplification attack

The amplitude of the watermarked speech signal is rescaled. A positive and negative rate of scaling indicates that the amplitude is amplified and attenuated, respectively. Table 19 shows the SNR values between watermarked speech signals and amplified watermarked speech signals, and shows that the majority of the values of BER are zero and values NC are 1. Though the attack is strong, we can identify our watermark easily. Fig.29 illustrates the watermarked speech signal (bass47_1), the amplified watermarked signal and the difference between them. It is observed that the difference is very big.

Table 19: SNR between watermarked signal and its amplified version

signal	SNR between WS & AMWS	Factor	BERs	NCs
spfe49_1	19.0362	+20%	00	1
	19.9561	-20%	00	1
spme50_1	21.5217	+20%	3.3377	0.9766
	22.2216	-20%	3.3377	0.9766
spfg53_1	22.3301	+20%	00	1
	22.9700	-20%	00	1
spmf52_1	22.9111	+20%	00	1
	23.5110	-20%	00	1
bass47_1	16.3558	+30	00	1
	17.5858	-30	00	1

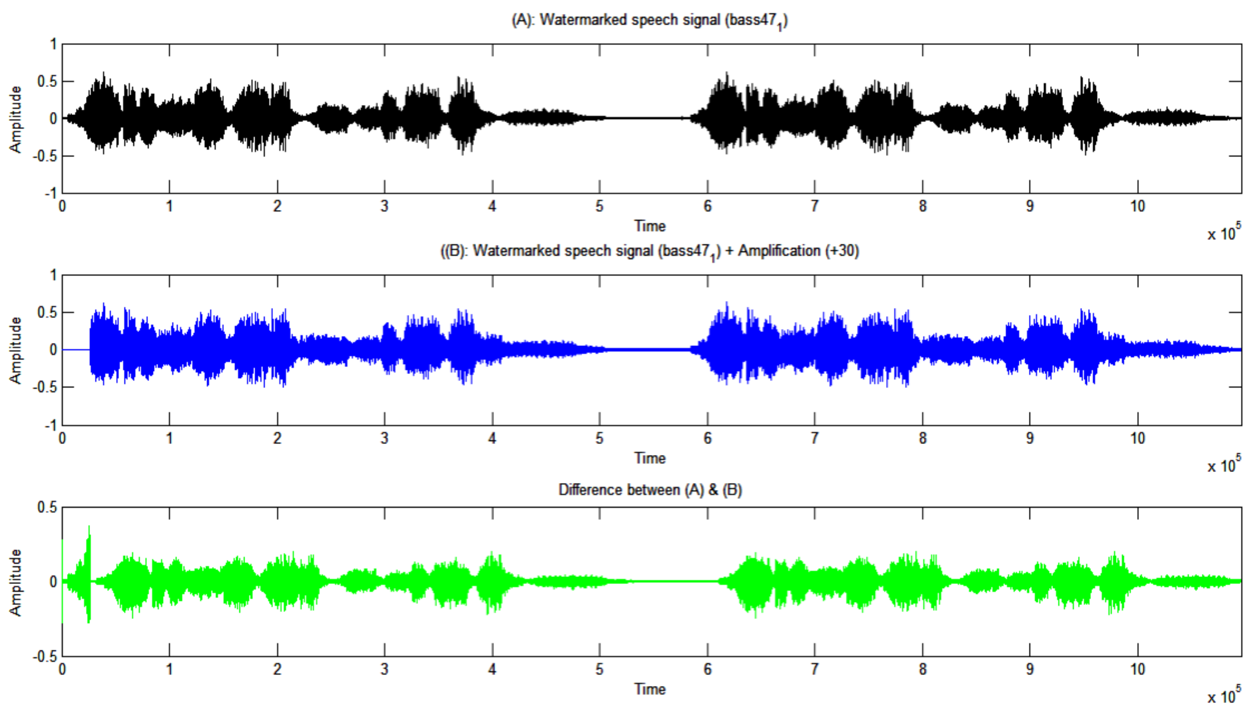


Figure 28: The difference between watermarked speech signal and its amplified version

Experiments show the strength and robustness of our proposed method and the SNR between watermarked speech signal (WS) and attacked watermarked speech signal (AWS) is small in all types of great attacks on the signal indicating the strength of the attack having a significant impact on the signal, to the level of losing its importance and quality and thus be unusable. Which means that the watermark resists until the signal becomes deficient. We can influence the watermark by greater attacks but it will not be useful if we lose completely the signal.

V.3. Capacity

Table 20 show the data payload of different speech signals. All of the capacities are widely superior to the minimum value imposed by IFPI (20 bits per seconds).

Table 20: Capacity of the watermarked speech signal

Watermark image 115x99 bits					
Speech	spfe49_1	spme50_1	spfg53_1	spmf52_1	bass47_1
Capacity (b/s)	593	675	688	568	457

V.4. Comparisons

To perform the fair comparison we change the watermark from binary fingerprint image to the watermark image shown in fig 21, so that, our proposed scheme has an equal or greater capacity than other three intended schemes for comparisons. Also choose the point where our proposed has more imperceptibility, on the other hand in the first comparison we used other speech signal “SP1” produced from speech “spfe49_1”, in last comparison we used also other speech signal “sp6” produced from “spmf52_1”. Tables 21 ,25 and 27 show the information On which the comparisons are based, Table 22, 26 and 28 show the comparisons with the other proposed schemes based on robustness aspect, the comparisons were performed with apply many popular attacks .

V.4.1. Comparison with results in [39]:

Fig.29 illustrates the original speech signal (SP1), the watermarked speech signal and the difference between them. It is obvious that the difference is extremely small and the watermark is spread on the entire signal with uniformity. Fig 30 illustrates the original speech signal (SP1), the watermarked signal and the difference between them. It is clear that the difference is very on small some parts of signal and very big on some parts and the watermark is distributed on the entire signal without uniformity.

Table 21: Comparison with scheme proposed in [39] based on SNR and capacity

Method / Parameters	Scheme proposed in [39]	Proposed scheme
Δ	0.036	0.015
SNR (db)	33.2417	34.8096
Capacity (b/s)	204.8	204.8

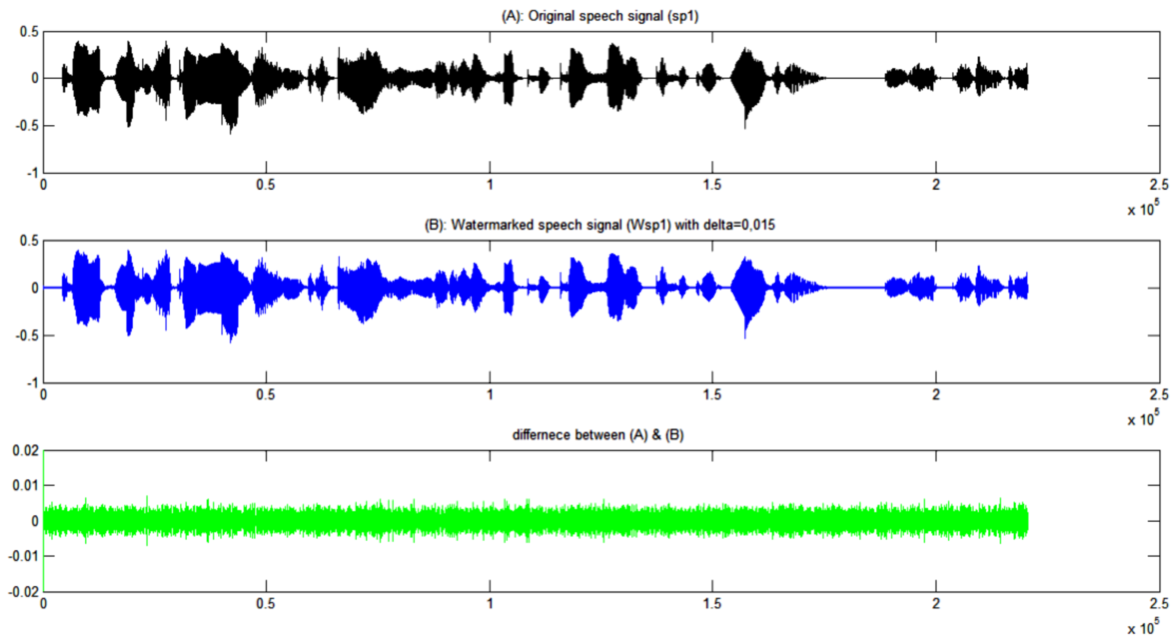


Figure 29: Results of our proposed scheme

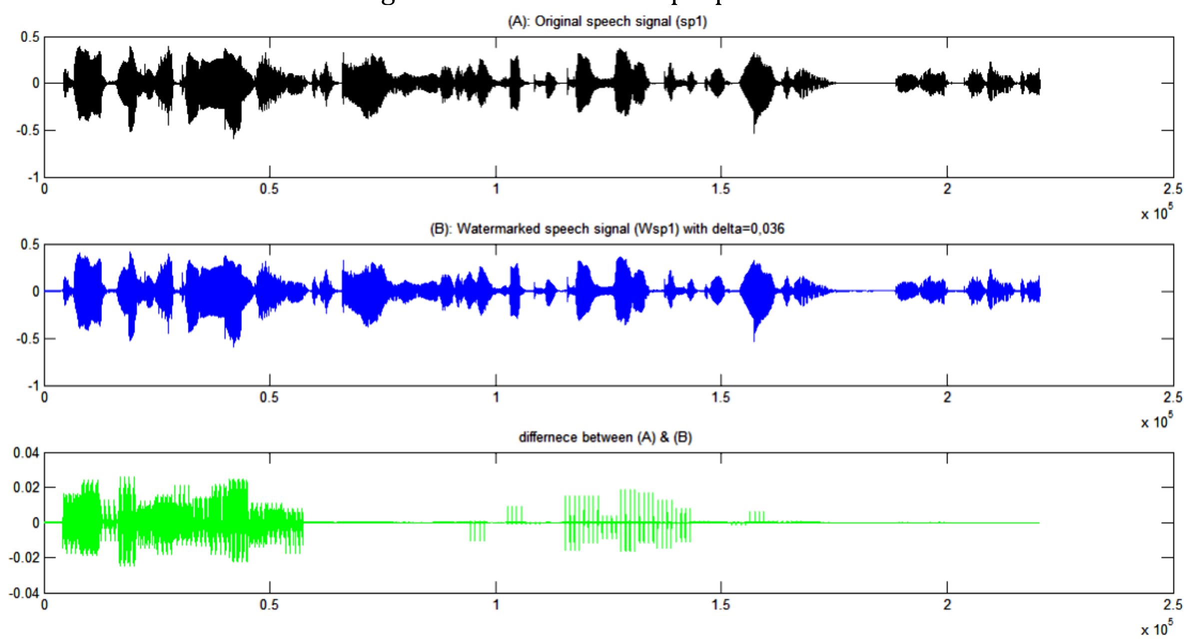
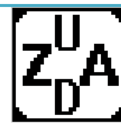
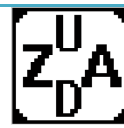
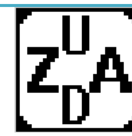
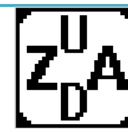
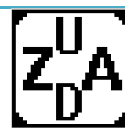
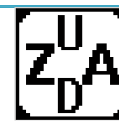
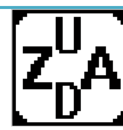
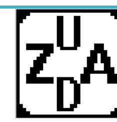
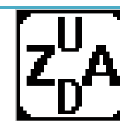


Figure 30: Results of the proposed scheme in [39]

Table 22: Comparison with scheme proposed in [39] based on different attacks using speech signal sp1

		Scheme proposed in [39]			**Proposed scheme**		
		BERs %	NCs	Images detected	BERs %	NCs	Images detected
Without attack		00	1		00	1	
AWGN	35 db	00	1		00	1	
	30 db	00	1		00	1	
	24db	0.8789	0.9933		00	1	
Re-quantization	Down (8bits)	00	1		00	1	
Cropping (1300 samples)	Nbr Beginning	00	1		00	1	
	Random	3.4180	0.9738		00	1	
Echo	(0.3,0.2)	47.0703	0.6015		00	1	
Amplification	+20%	90.8203	0.1156		00	1	
	-20%	96.5820	0.0446		00	1	

NBR: as defined in 'III.4.2.f.Cropping attack' (the samples cropped are attacked with AWGN).

Table 23: Comparison between elapsed times in our proposed and proposed in [39] (embedding)

	Embedding time (seconds)		
	Speech signals		
Methods	SP1	spmf52_1	bass47_1
Scheme proposed in [39]	2.955488	4.938818	5.875246
proposed	0.405195	0.738192	0.860793

Table 24: Comparison between elapsed times in our proposed and proposed in [39] (extracting)

	Extracting time (seconds)		
	Speech signals		
Methods	SP1	spmf52_1	bass47_1
Scheme proposed in [39]	1.006378	1.088234	1.145674
proposed	0.179673	0.318821	0.361185

Tables 23 and 24 show the elapsed time by the embedding and extraction process of our proposed and scheme in [39], although two schemes can embedded and extract the watermark in real time, everyone can easily observe that our scheme has high speed in execution and speeder than the other one.

V.4.2. Comparison with results in [10]:

Fig.31 illustrates the original speech signal (bass47_1), the watermarked signal and the difference between them. It is obvious that the difference is extremely small and the watermark is spread on the entire signal with uniformity. Fig 32 illustrates the original speech signal (bass47_1), the watermarked signal and the difference between them. It is clear that the difference is very small on some parts of signal and very big on some parts and the watermark is distributed on the entire signal without uniformity.

Table 25: Comparison with scheme proposed in [10] based on SNR and capacity

Method / Parameters	Scheme proposed in [10]	Proposed scheme
Δ	/	0.035
SNR (db)	33.39	34.6131
Capacity (b/s)	17,2	41.19067

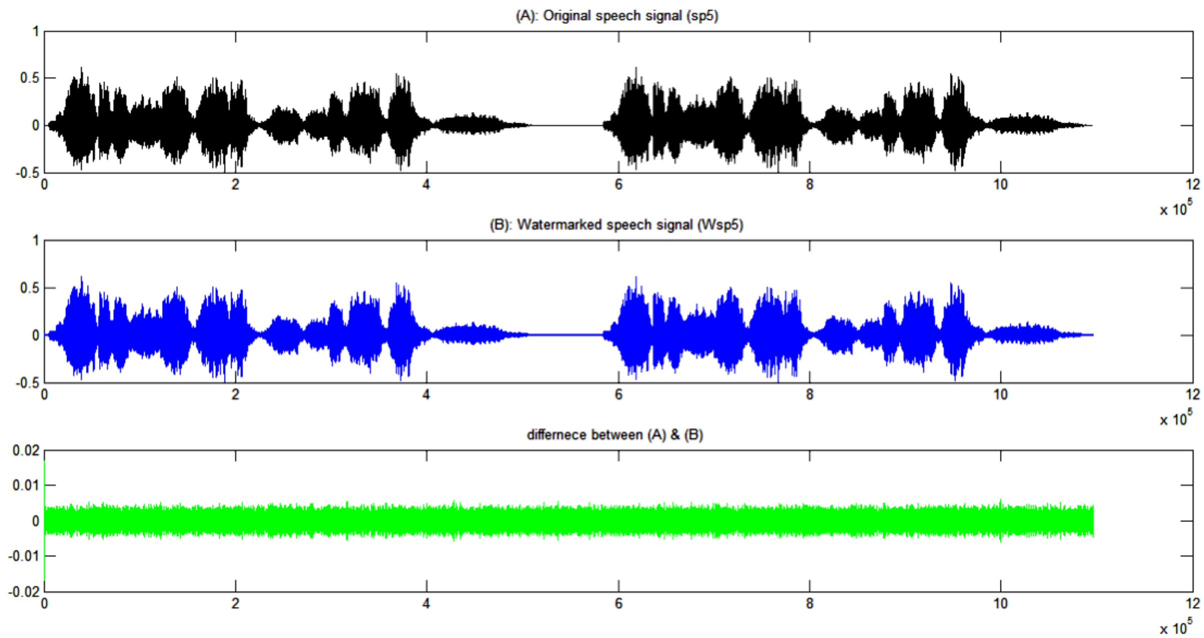


Figure 31: Results of our proposed scheme

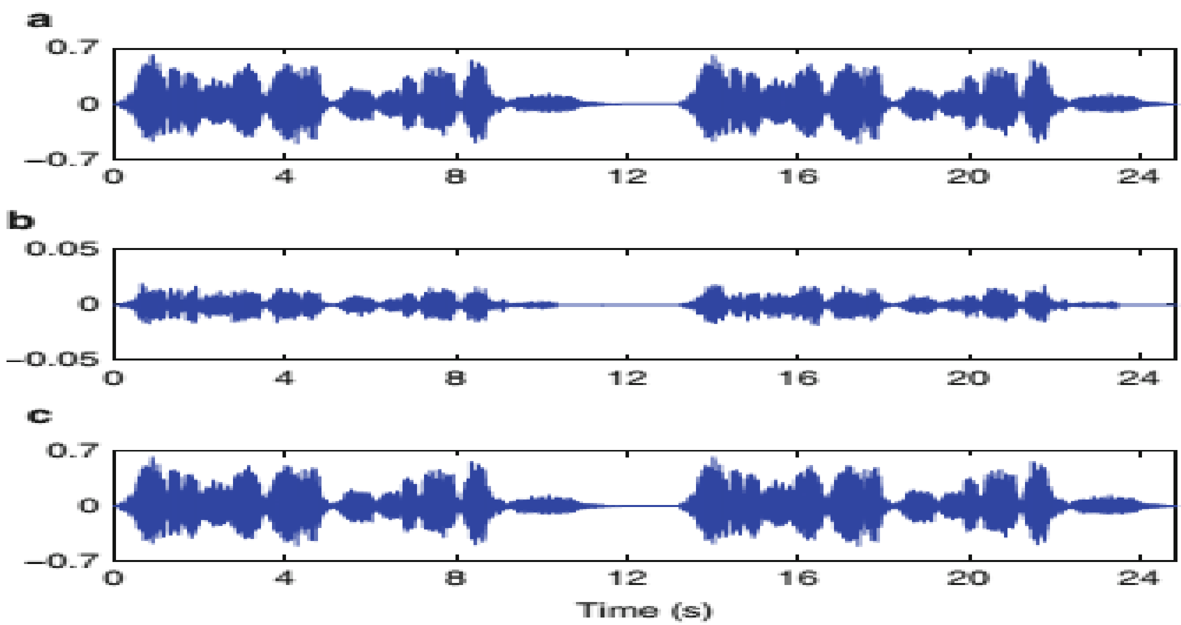
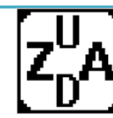


Figure 32: Results of Scheme proposed in [10] (a: original; b: difference (a, c) ; c: watermarked)

Table 26: Comparison with scheme proposed in [10] based on different attacks using speech signal sp5

		Scheme proposed in [10]				Proposed scheme		
		BERs %		NCs	Images detected	BERs %	NCs	Images detected
		B*	AS+					
Without attack		00	00	-		00	1	
AWGN	40db	6.86	5.71	-		00	1	
	36 db	11.71	9.14	-		00	1	
	30db	x	18.57	-		00	1	
Re-quantization	Down (8bits)	17.71	16.00	-		00	1	
Cropping (samples)	8x25ms	x	00	-		00	1	
Echo	(0.3,0.2)	0.57	0.57	-		00	1	
Amplification	+20	00	00	-		00	1	
	-20	00	00	-		00	1	

*: Basic Detection +: Adaptive synchronization

V.4.3. Comparison with results in [48]:

Fig.33 illustrates the original speech signal (SP6), the watermarked speech signal and the difference between them. It is obvious that the difference is extremely small and the watermark is spread on the entire signal with uniformity. Fig 34 illustrates the original speech signal (SP6), the watermarked speech signal and the difference between them. It is clear that the difference is very small on some parts of signal and very big on some parts and the watermark is distributed on the entire signal with a poor form.

Table 27: Comparison with scheme proposed in [48] based on SNR and capacity

Method / Parameters	Scheme proposed in [48]	Proposed scheme
Δ	0.96	0.02
SNR (db)	27.1210	33.7528
Capacity (b/s)	172.39	172.39

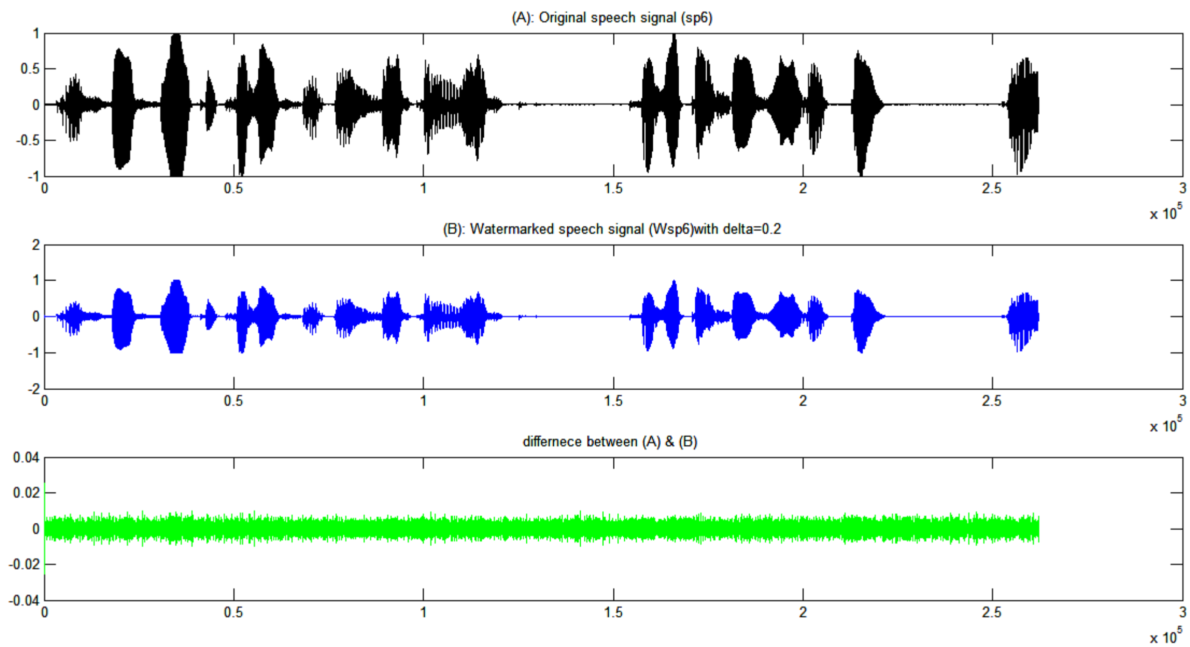


Figure 33: Results of our proposed scheme

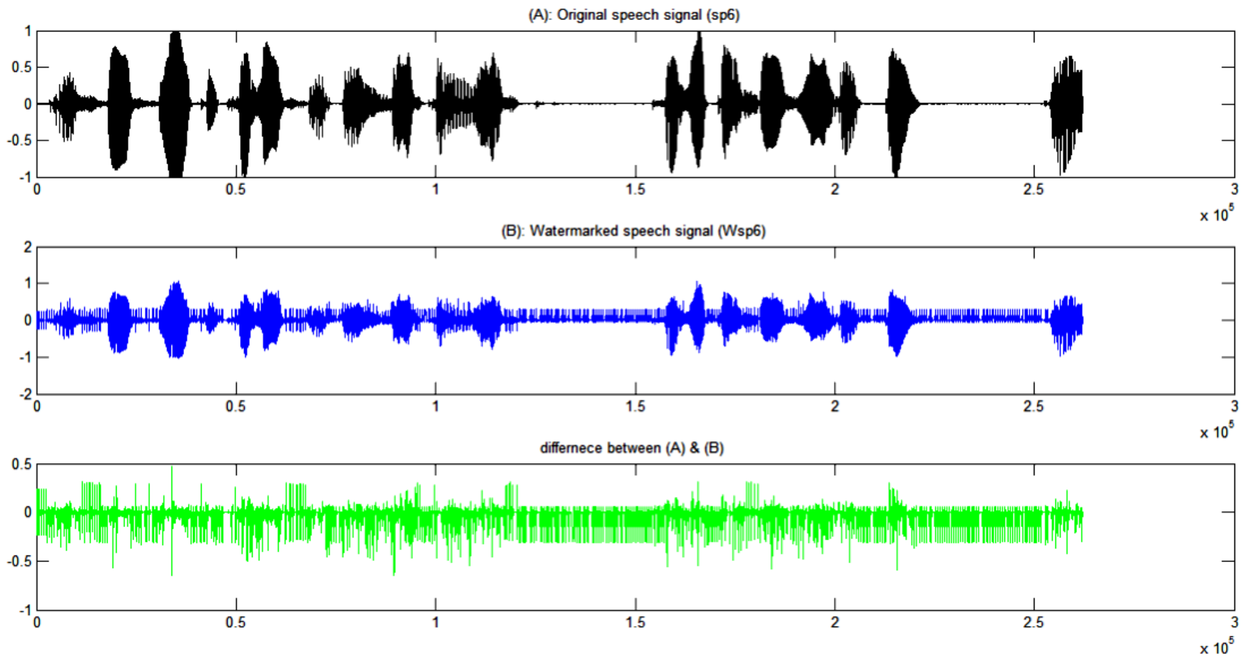
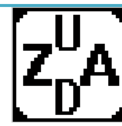
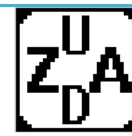
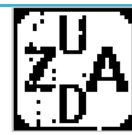
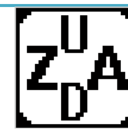
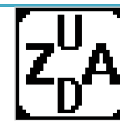
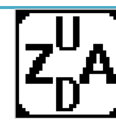
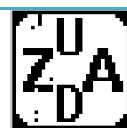
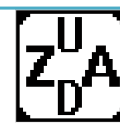
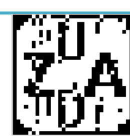
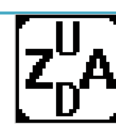


Figure 34: Results of Scheme proposed in [48]

The proposed method works well. It is known that in most of watermarking methods there is an inverse proportionality between robustness and imperceptibility. We tried to find a trade-off by keeping imperceptibility with increasing strength and robustness. To preserve imperceptibility, we exploited the correlation between each two successive samples by sub-sampling the signal. To enhance robustness, we space between these sub-samples values. This is done using Δ; but since each adjacent two samples are extremely close to each other, the small variations in Δ certainly keep better signal imperceptibility and will separate them clearly which will give superior robustness.

Table 28: Comparison with scheme proposed in [48] based on different attacks using speech signal sp6

		Scheme proposed in [48]			**Proposed scheme**		
		BERs %	NCs	Images detected	BERs %	NCs	Images detected
Without attack		00	1		00	1	
AWGN	30 db	1.1719	0.9910		00	1	
	24db	2.0508	0.9843		00	1	
Re-quantization	Down (8bits)	0	1		00	1	
Cropping (1300 samples)	Nbr Beginning	1.1719	0.9910		00	1	
	Random	0.7813	0.9940		00	1	
Echo	(0.3,0.2)	8.8867	0.9314		00	1	
Amplification	+20%	8.7891	0.9314		00	1	
	-20%	7.4219	0.9423		00	1	

Nbr: as defined in 'III.4.2.f .Cropping attack' (the samples cropped are attacked with AWGN).

V.5. Conclusion

This chapter includes the results of implementation of a new scheme. The new scheme is a hybrid approach of DWT and DCT for watermarking speech signals by biometric data. Sub-sampling is made before transforms operations and the biometric data (fingerprint) is embedded then. Inverse process is realized on the watermarked and hardly attacked and noised speech signal in real conditions. Experimental results performed on different lengths of speeches signals and also different types of signals (male and female) and different languages (English, French, German), indicate that the proposed scheme is robust against different attacks and noises compared to some previous recently published works with good imperceptibility and better performance in embedding capacity, in addition it has a high speed in the execution. According to the designed scheme can has high robustness, high imperceptibility and also can carry a lot of bits (high capacity), we can guarantee that this scheme suitable for many applications such in biometric systems, fingerprinting, copyright protection.

General conclusion

The completely understanding of the conflict between the requirements of watermarking system confirms that to design and implement a scheme for speech and audio watermarking is complicated and not easy, although that, we can found perfect solutions, then create two new schemes for this purpose and satisfy the requirements. The proposed schemes jointly exploits benefits of DWT, DCT, sub-sampling and norm space to obtain an effective blind watermarking systems with good auditory quality, practical resistance against mainly attacks, high data payload and low computation in execution.

The Reaching to high energy regions where human auditory system is less sensitive is by applying DWT and DCT, whereas the approximation coefficients after DWT include most of the signal energies, the DCT offers to storage the high energies in a small number of samples. The correlation between two vectors after sub-sampling operation in speech and audio signal is very high. The norm space let us embedded the watermark bits in significant position, which the norm related to all samples of the vector.

The first proposed algorithm based on DWT, DCT, sub-sampling and norm space, also we employed Arnold transform to encrypt the watermark bits. Through the results in chapter 4 the proposed scheme produced acceptable results, whereas, the extraction of the watermark is without using original speech and audio signals, besides the extraction without using quantization step (Δ),the SNR confirmed the imperceptibility and SSNR validated trusty values of SNR and proved that there is no camouflage, the capacity is satisfactory, the embedded watermark can resist hard attacks. Comparisons with results of other schemes concerning the inaudibility, robustness and capacity authenticate the preference of our proposed scheme.

The second proposed algorithm, only based on DWT, DCT and sub-sampling, the insertion of biometric watermark bits were randomly. The watermark was a digital fingerprint. The chapter 5 presented perfect results of the proposed scheme; while, the extraction of biometric watermark was blindly and without employed the parameter of insertion (Δ), the SNR and capacity were widely superior to the minimum values imposed by IFPI which indicated the proposed scheme has the strength to embed a great number of bits imperceptibility, the extraction after strong attacks was easily which demonstrated the robustness. The fair comparisons achieved through four metrics; imperceptibility, robustness, payload and execution speed with some recently schemes established the strength of proposed scheme.

Precedent points giving us three ideas for the future works. The first idea is making the insertion of the parameter (Δ) adaptive, which offers an increase in the imperceptibility requirement but we will exploit it to enhance the number of embedded bits. The second idea is applying the schemes on Arabic language speeches and apply all necessary improvements, due to the reason that Arabic language is practiced and difficult and also classified in one of the first ranks in the world. The third suggested idea using a greyscale fingerprint image as a watermark instead of binary fingerprint image with implementation of all the required alterations, because the greyscale image can offer more security in biometric systems.

List of scientific productions

- **International Publication**

1. Ahmed Merrad & Slami Saadi **« Blind speech watermarking using hybrid scheme based on DWT/DCT and sub-sampling »,** Multimedia Tools and Applications journal 2018, DOI 10.1007/s11042-018-5939-z.

2. Slami Saadi, Ahmed Merrad & Ali Benziane **«Novel secured scheme for blind audio/speech norm-space watermarking by Arnold algorithm»,** Signal Processing journal 154(2019) 74–86, https://doi.org/10.1016/j.sigpro.2018.08.011.

- **International Communications**

1. Ahmed Merrad, Slami Saadi & Ali Benziane & Ahmed HAFAIFA **« Robust Blind Approach for Digital Speech Watermarking»,** the second international conference on natural language and speech processing (ICNLSP 2018), Alger .

Bibliography

[1] Suraj Kumar Dubey, Vivek Chandra "Steganography, Cryptography and Watermarking: A Review" International Journal of Innovative Research in Science, Engineering and Technology, February 2017, DOI:10.15680/IJIRSET.2017.0602076 pp2595-2599.

[2]Hardikkumar V. Desai "Steganography, Cryptography, Watermarking: A Comparative Study" Journal of Global Research in Computer Science, December 2012, 33-35.

[3] Michael E. Osadebey and Apostolos A. Georgakis "Spread spectrum wavelet watermarking system" DML Technical Report: DML-TR-2005:02 ISSN Number: 1652-8441, DEPARTMENT OF APPLIEDPHYSICS ANDELECTRONICS UME° AUNIVERISTY,SWEDEN.

[4] Antony, Sobin c. Sherly. A.P " Audio Steganography in Wavelet Domain – A Survey".International Journal of Computer Applications, Vol 52, No.13, pp 975 ... Journal of Applied Sciences, Engineering and Technology, Jul 2012.

[5] Michael Arnold, Martin Schmucker, Stephen D. Wolthusen "Techniques and Applications of Digital Watermarking and Content Protection", 2003 ARTECH HOUSE, INC.

[6] Channapragada Rama Seshagiri Rao, Munaga V.N.K. Prasad "Digital Watermarking Techniques in Curvelet and Ridgelet Domain", SpringerBriefs in Computer Science (2016), DOI 10.1007/978-3-319-32951-2.

[7] Sascha Zmudzinski "Digital Watermarking for Verification of Perception-based Integrity of Audio Data" a thesis PhD Berlin west, 2017.

[8] Patrick loo "a watermarking using complex wavelets" a dissertation submitted to the university of Cambridge for the degree of doctor of philosophy, thesis 2002.

[9] Mohammad Ali Nematollahi, Chalee Vorakulpipat and Hamurabi Gamboa Rosales "Digital Watermarking Techniques and Trends" Springer Science+Business Media Singapore 2017.

[10] Yiqing Lin, Waleed H. Abdulla, "Audio Watermark: A Comprehensive Foundation Using MATLAB", ISBN 978-3-319-07974-5 (eBook). DOI 10.1007/978-3-319-07974-5. Springer Cham Heidelberg New York Dordrecht London, 2015.

[11] Pranab Kumar Dhar and Tetsuya Shimamura, "Advances in Audio Watermarking Based on Singular Value Decomposition" SPRINGER BRIEFS IN ELECTRICAL AND COMPUTER ENGINEERING (2015), DOI 10.1007/978-3-319-14800-7.

[12] Mahmoud El-Gayyar "Watermarking Techniques Spatial Domain Digital Rights Seminar" Media Informatics University of Bonn Germany May 06th.

[13] Yiqing Lin, Waleed H. Abdull "audio watermarking for copyrights protection", technical report school of engineering report no.650 the university of Auckland, private bag 92019,Auckland, New Zealand 2007 .

[14] Borko Furht, Edin Muharemagic and Daniel Socek “MULTIMEDIA ENCRYPTION AND WATERMARKING” 2005 Springer Science+Business Media.

[15] David Megias, Jordi Herrera-Joancomarti, and Julia Minguillon “Robust Frequency Domain Audio Watermarking: A Tuning Analysis” Third International Workshop, IWDW 2004 Seoul, South Korea, October 30 - November 1, 2004.

[16] Juergen Seitz “Digital watermarking for digital media” Information Science Publishing (an imprint of Idea Group Inc.) 2005.

[17] Nedeljko Cvejic, Tapio Seppänen “Digital Audio Watermarking Techniques and Technologies: Applications and Benchmarks”, 2008 by IGI Global Information Science Reference.

[18] Raul Martinez-Noriega “LDPC coded Watermarks for audio: Algorithms tolerant to compression and desynchronization” a dissertation submitted in partial satisfaction of the requirements for the degree doctor of engineering , March 2011, the university of Electgro-communications, Cair .

[19] Yong Xiang, Guang Hua and Bin Yan “Digital Audio Watermarking Fundamentals, Techniques and Challenges” SpringerBriefs in Electrical and Computer Engineering Signal Processing 2017.

[20] Frank Y. Shih “Digital Watermarking and Steganography Fundamentals and Techniques “2008 by Taylor & Francis Group, LLC

[21] Feng-Hsing Wang, Jeng-Shyang Pan and Lakhmi C. Jain “Innovations in Digital Watermarking Techniques” Studies in Computational Intelligence,Volume 232, 2009 Springer-Verlag Berlin Heidelberg.

[22] Xing He “Signal Processing, Perceptual Coding and Watermarking of Digital Audio: Advanced Technologies and Models” 2012 by IGI Global Information Science Reference.

[23] Rohit Thanki • Komal Borisagar • Surekha Borra “Advance Compression and Watermarking Technique for Speech Signals” SpringerBriefs in Electrical and Computer Engineering 2018, https://doi.org/10.1007/978-3-319-69069-8.

[24] Amit Kumar Singh • Basant Kumar, Ghanshyam Singh • Anand Mohan ”Medical Image Watermarking Techniques and Applications” Springer International Publishing AG 2017, DOI 10.1007/978-3-319-57699-2 .

[25] Christian Rathgeb • Andreas Uhl • Peter Wild “Iris Biometrics” Springer Science+Business Media, LLC 2013.

[26] S. M. Deokar and B. Dhaigude, “Blind audio watermarking based on Discrete wavelet and Cosine transform”, 2015, International Conference on Industrial Instrumentation and Control (ICIC), College of Engineering Pune, India. May 28-30,2015.

[27] V. Mehta and N. Sharma, “Secure Audio Watermarking based on Haar Wavelet and Discrete Cosine Transform”, International Journal of Computer Applications (0975 – 8887), Volume 123 – No.11, August 2015.

[28] A. Tiwari and M. Sharma, "Comparative Evaluation of Semi Fragile Watermarking Algorithms for Image Authentication", Journal of Information Security, Vol. 3 No. 3, 2012, pp. 189-195, DOI: 10.4236/jis.2012.33023.

[29] C. Yong-mei, G. Wen-qiang and D. Hai-yang, ”An Audio Blind Watermarking Scheme Based on DWT-SVD”, JOURNAL OF SOFTWARE, VOL. 8, NO. 7, JULY 2013, 1801-1808.

[30] E. Brannock, M. Weeks and R. Harrison, ”Watermarking with Wavelets simplicity leads to Robustness”, (2008) IEEE.

[31] L. Debnath,”Wavelet Transform and Their Application”, Birkhäuser (2002).

[32] M. A. Osman, N. H. Ali, "Audio Watermarking Based on Wavelet Transform", Applied Mechanics and Materials Vols 229-231 (2012) pp 2784-2788, Trans Tech Publications, Switzerland, doi:10.4028/www.scientific.net/AMM.229-231.2784.

[33] A.E. Villanueva-Luna, A. Jaramillo-Nuñ ez, D. Sanchez-Lucero, C. M. Ortiz-Lima, J. Gabriel Aguilar-Soto, A. Flores-Gil, M. May-Alarcon, ”De-Noising Audio Signals Using MATLAB Wavelets Toolbox”, Engineering Education and Research Using MATLAB, Dr. Ali Assi (Ed.), ISBN: 978-953-307-656-0, InTech, (2011). http://www.intechopen.com/books/engineering-education-and-researchusing-matlab/de-noising-audio-signals-using-matlab-wavelets-toolbox.

[34] Mohammad Ali Nematollah , S.A.R Al-Haddad , Shyamala Doraisamy, M. Iqbal Bin Saripan “Digital Audio & speech Watermarking Basedon the Multiple Discrete Wavelets Transform and Singular Value Decomposition” 2012 Sixth Asia Modelling Symposium, DOI 10.1109/AMS.2012.54.

[35] Achintya Singhal, Anurag Narayan Chaubey, Chandra Prakash “Audio Watermarking Using Combination of Multilevel Wavelet Decomposition, DCT and SVD” IEEE 2011.

[36] M. Hemis, B. Boudraa and T. M. Meksen,” New secure and robust audio watermarking algorithm based on QR factorization in wavelet domain”, Int. J. Wavelets Multiresolut Inf. Process. 13, 1550020 (2015) , https://doi.org/10.1142/S0219691315500204.

[37]Pejman Rasti, Gholamreza Anbarjafari and Hasan Demirel “colour Image Watermarking based on Wavelet and QR Decomposition” 2017 IEEE.

[38]Qingtang Su,Yugang Niu ,Gang Wang ,ShaoliJia ,JunYue “Color image blind watermarking scheme based on QR decomposition” Volume 94, January 2014, Pages 219-235 Signal Processing.

[39] P. K. Dhar, “A Blind Audio Watermarking Method Based on Lifting Wavelet Transform and QR Decomposition”, 8th International Conference on Electrical and Computer Engineering 20-22 December, 2014 , Dhaka, Bangladesh 136-139.

[40] X.Wang, P.Wang, P.Zhang, H.Yang, “A blind audio watermarking algorithm by logarithmic quantization index modulation”, Multimed Tools Appl DOI 10.1007/s11042-012-1259-x. Springer Science+Business Media New York 2012.

[41] X. Wang, P.Wang, P. Zhang, S. Xu, H. Yang, “A norm-space, adaptive, and blind audio watermarking algorithm by discrete wavelet transform”, Signal Processing 93 (2013) 913–922, http://dx.doi.org/10.1016/j.sigpro.2012.11.003 .

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

[43] Subir andAmit M. Joshi, “DWT-DCT based Blind Audio Watermarking using Arnold Scrambling and Cyclic Codes”, 2016 3rd International Conference on Signal Processing and Integrated Networks (SPIN).

[44]Xing He “Signal Processing, Perceptual Coding and Watermarking of Digital Audio:Advanced Technologies and Models”2012 by IGI Global.

[45] B.Y. Lei, I. Y. Soon and Z. Li, “Blind and robust audio watermarking scheme based on SVD–DCT”, Signal Processing 91 (2011) 1973–1984 , doi:10.1016/j.sigpro.2011.03.001.

[46] Dhar PK, Shimamura T, Blind SVD-based audio watermarking using entropy and log-polar transformation, Journal of Information Security and Applications (2014), http://dx.doi.org/10.1016/j.jisa.2014.10.007.

[47] H.-T. Hu and L.-Y. Hsu, “ Robust, transparent and high-capacity audio watermarking in DCT domain”, Signal Processing 109 (2015) 226–235, http://dx.doi.org/10.1016/j.sigpro.2014.11.011 .

[48] Pranab Kumar DHAR, Tetsuya SHIMAMURA, “Blind Audion Watermarking in Transform Domain Based on Singular Value Decomposition and Exponential-Log Operations”, RADIOENGINEERING, Vol.26, NO.2, June 2017. DOI:10.13164/re.2017.0552.

[49] Shijun Xiang, Jiwu Huang,”Robust Audio Watermarking Against the D/A and A/D Conversions”. CoRR abs/0707.0397 (2007).

[50] H.Yang, D. Bao, X. Wang, P. Niu,” A robust content based audio watermarking using UDWT and invariant histogram”, Multimed Tools Appl (2012) 57:453–476. DOI 10.1007/s11042-010-0644-6. Springer Science+Business Media, LLC 2010.

[51] V. Bhat, K. I. Sengupta and A. Das, “An adaptive audio watermarking based on the singular value decomposition in the wavelet domain”, Digital Signal Processing 20 (2010) 1547–1558, doi:10.1016/j.dsp.2010.02.006.

[52] A. Al-Haj, A. Mohammad and L. Bata, “DWT–Based Audio Watermarking”, The International Arab Journal of Information Technology, Vol. 8, No. 3, July 2011,326-333.

[53] B.Y. Lei, I.Y. Soon, F. Zhou, Z. Li and H. Lei, “A robust audio watermarking scheme based on lifting wavelet transform and singular value decomposition”, Signal Processing 92 (2012) 1985–2001 doi:10.1016/j.sigpro.2011.12.021 .

[54] Jyotirmayee Mishra, Pune M.V.Patil, “An Effective Audio Watermarking using DWT-SVD”, Department of Electronics, BVDUCOE, BVDUCOE, Pune J.S.Chitode, 2013 International Journal of Computer Applications.

[55] Ali Al-Haj,”A dual transform audio watermarking algorithm”, Multimed Tools Appl, DOI 10.1007/s11042-013-1645-z. Springer Science+Business Media New York 2013.

[56] H. T. Hu , H.H. Chou, C.Yu and L.Y. Hsu, "Incorporation of perceptually adaptive QIM with singular value decomposition for blind audio watermarking", EURASIP Journal on Advances in Signal Processing, 2014:12, http://asp.eurasipjournals.com/content/2014/1/12 .

[57] B. Lei, F. Zhou, E. L.Tan, D. Ni, H. Lei, S. Chen and T. Wang, "Optimal and secure audio watermarking scheme based on self-adaptive particle swarm optimization and quaternion wavelet transform", Signal Processing (2014), http://dx.doi.org/10.1016/j.sigpro.2014.11.007 .

[58] H.T. Hu, L.Y. Hsu and H.H. Chou, "Variable-dimensional vector modulation for perceptual-based DWT blind audio watermarking with adjustable payload capacity", Digital Signal Processing (2014), http://dx.doi.org/10.1016/j.dsp.2014.04.014 .

[59] M.A. Nematollahi , H.G. Rosales, M.A. Akhaee and S.A.A. Al-Haddad, "Robust digital speech watermarking for online speaker recocognition", Hindawi publishing corporation mathematical problems in engineering 2015, Volume 2015, Article ID 372398, http://dx.doi.org/10.1155/2015/372398 .

[60] Nair U.R, Birajdar G.K, "Audio watermarking in wavelet domain using Fibonacci numbers", Signal and Information Processing (IconSIP), International Conference on. IEEE, 2016:1-5.

[61] Elshazly A.R, Nasr M.E, Fuad M.M, et al., "Synchronized double watermark audio watermarking scheme based on a transform domain for stereo signals", Electronics, Communications and Computers (JEC-ECC), 2016 Fourth International Japan-Egypt Conference on. IEEE, 2016: 52-57.

[62] A.Kaur, M.K.Dutta,K.M. Soni,N.Taneja, « Localized & self adaptive audio watermarking algorithm in the wavelet domain », Journal of Information Security and Applications 33 (2017) 1–15, http://dx.doi.org/10.1016/j.jisa.2016.12.003 2214-2126/©2017 Elsevier Ltd.

[63] H. T. Hu, S. J. Lin and L. Y. Hsu, "Effective blind speech watermarking via adaptive mean modulation and package synchronization in DWT domain", EURASIP Journal on Audio, Speech, and Music Processing 2017:10, doi:10.1186/s13636-017-0106-4.

[64] H. T. Hu, L.Y. Hsu and H.H. Chou, "Perceptual-based DWPT-DCT framework for selective blind audio watermarking", Signal Processing (2014),[12pages], http://dx.doi.org/10.1016/j.sigpro.2014.05.003 .

[65] Sujit M.Deokar Bhaveek Dhaigude, "Blind audio watermarking based on Discrete wavelet and Cosine transform", 2015 international conference on industrial instrumentation and control (ICIC).

[66] Pranab Kumar Dhar "Studies on Digital Audio Watermarking Using Matrix Decomposition" A Dissertation Submitted to the Graduate School of Science and Engineering in Partial Fulfillment of the Requirements for the Degree of DOCTOR OF PHILOSOPHY in Mathematics, Electronics and Informatics, Saitama University, JapanSeptember 2014.

[67] A. Benoraira, K. Benmahammed and N. Boucenna, "Blind image watermarking technique based on differential embedding in DWT and DCT domains", EURASIP Journal on Advances in Signal Processing (2015) 2015:55, DOI 10.1186/s13634-015-0239-5.

[68] Evelyn Brannock, Michael Weeks, Robert Harrison, "Watermarking with Wavelets simplicity leads to Robustness", (2008) IEEE.

[69] Chu Wai.C, “speech coding algorithms”, foundation and evolution of standardized coders (2003).

[70] S. Shokri, M. Ismail, N. Zainal and M. Moghaddasi, “Audio-Speech Watermarking Using a Channel Equalizer“, Wireless Pers Commun (2017), DOI 10.1007/s11277-017-4095-5.

[71] V. Bhat , K. I.Sengupta and A. Das, “An audio watermarking scheme using singular value decomposition and dither-modulation quantization”, Multimedia Tools Appl (2010), DOI: 10.1007/s11042-010-0515-1.

[72] http://sound.media.mit.edu/resources/mpeg4/audio/sqam/, last check at 24/09/2018 11:31 .

[73] M. Ogura, Y. Sugiura and T. Shimamura, “SVD Based Audio Watermarking Using Angle-Quantization”, International Conference on Electrical, Computer and Communication Engineering (ECCE), February 16-18, 2017, Cox’s Bazar, Bangladesh.

[74] Z. Liu , J. Huang , X. Sun and C. Qi, ”A security watermark scheme used for digital speech forensics”, Multimedia Tools Appl (2016), DOI 10.1007/s11042-016-3533-9 2016.

[75]http://english.ia.cas.cn/db/201611/t20161101_169922.html last check at 08/09/2018 10:35