



DOCTORAL THESIS

3rd Cycle Doctoral (D-LMD)

Presented by

AMRAOUI ABDELKADER

With a view to obtaining the doctoral diploma in 3rd Cycle Doctoral (D-LMD)

Branch: Electronics

Specialty: Electronics

Topic

Speaker identification and verification security improvement
using speech watermarking

Supported, on 26 / 09 / 2023, before the jury composed of:

Last and first name	Grade	Institution of affiliation	Designation
Mr : Elhadj Ahmed Ghazel	Professeur	Université de Djelfa	President
Mr : Slami SAADI	Professeur	Université de Djelfa	Supervisor
Mr : Abbassi Ahmed	MCA	National school bousada	Co-Supervisor
Mr : Bessissa Lakhdar	Professeur	Université Djelfa	Examiner
Mr : Sone Abdelkader	MCA	Université de Laghouat	Examiner
Mr : Mitiche Lahcene	Professeur	Université de Djelfa	Examiner

Djelfa University , FST, 2023



THESE DE DOCTORAT

Doctorat 3^{ème} Cycle (D-LMD)

Présentée par

AMRAOUI ABDELKADER

En vue de l'obtention du diplôme de Docteur en 3^{ème} Cycle D-LMD

Filière : Electronique

Spécialité : Electronique

Thème

Amélioration de la sécurité d'identification et de vérification du locuteur en utilisant le tatouage de la parole

Soutenue publiquement, le ... /... /2023, devant le jury composé de :

Nom et Prénom	Grade	Etablissement de rattachement	Désignation
Mr :Elhadj Ahmed Ghazel	Professeur	Université de Djelfa	Président
Mr : Slami Saadi	Professeur	Université de Djelfa	Directeur de thèse
Mr : Abbassi Ahmed	MCA	Ecole Nationale bousada	Co- Dir de thèse
Mr : Bessissa Lakhdar	Professeur	Université de Djelfa	Examineur
Mr : Sone Abdelkader	MCA	Université de Laghouat	Examineur
Mr : Mitiche Lahcene	Professeur	Université de Djelfa	Examineur

Université de Djelfa, FST, 2023

بسم الله الرحمن الرحيم

إِنَّ الَّذِينَ ارْتَدُّوا عَلَيَّ أَدْبِرْهُمْ مِّنْ بَعْدِ مَا تَبَيَّنَ لَهُمُ الْهُدَىٰ الشَّيْطَانُ
سَوَّلَ لَهُمْ وَأَمْلَىٰ لَهُمْ (26) ذَلِكَ بِأَنَّهُمْ قَالُوا لِلَّذِينَ كَرِهُوا مَا نَزَّلَ
اللَّهُ سَنُطِيعُكُمْ فِي بَعْضِ الْأُمْرِ وَاللَّهُ يَعْلَمُ أَسْرَارَهُمْ (27) فَكَيْفَ
إِذَا تَوَفَّتْهُمُ الْمَلَائِكَةُ يَضْرِبُونَ وُجُوهَهُمْ وَأَدْبِرْهُمْ (28) ذَلِكَ بِأَنَّهُمْ
إِتَّبَعُوا مَا آسَخَطَ اللَّهُ وَكَرِهُوا رِضْوَانَهُ فَأَحْبَطَ أَعْمَلَهُمْ (29) أَمْ
حَسِبَ الَّذِينَ فِي قُلُوبِهِم مَّرَضٌ أَنْ لَّنْ يُخْرِجَ اللَّهُ أَضْغَانَهُمْ
(30) وَلَوْ نَشَاءُ لَأَرَيْنَاكَهُمْ فَلاَعْرِفْتَهُمْ بِسِيمَاهُمْ وَلَتَعْرِفَنَّهُمْ فِي
لَحْنِ الْقَوْلِ وَاللَّهُ يَعْلَمُ أَعْمَالَكُمْ (31) وَلَنَبْلُوَنَّكُمْ حَتَّىٰ نَعْلَمَ الْمُجْهِدِينَ
مِنْكُمْ وَالصَّابِرِينَ وَنَبْلُوَنَّكُمْ أَخْبَارَكُمْ (32) إِنَّ الَّذِينَ كَفَرُوا وَصَدُّوا
عَنْ سَبِيلِ اللَّهِ وَشَاقُّوا الرَّسُولَ مِنْ بَعْدِ مَا تَبَيَّنَ لَهُمُ الْهُدَىٰ لَن
يُضْرَبُوا اللَّهُ شَيْئًا وَسَيُحِبُّ أَعْمَلَهُمْ (33) .

صدق الله العظيم

Remerciement

Monsieur le président du juré Pr: Elhadj Ahmed Ghezal, messieurs les membres du juré de discussion de thèse ,Pr: Sone Abdelkader ,Pr:Bessissa Lakhdar, Pr:Mitiche Lahcene , Pr: Saadi slami et le Dr: Abbassi Ahmed .

Avec une immense gratitude je me tiens devant vous aujourd'hui, à l'occasion de cette réunion de jurés du soutenance de mon thèse de doctorat . Je tiens tout d'abord à exprimer mes sincères remerciements à chacun d'entre vous, président du juré et membres du jurés , pour votre présence .

je remercie mon directeurs de thèse , le Pr: Saadi slami et le co-directeur le Dr: Abbassi Ahmed , pour leur soutien constant, leurs conseils . Je souhaite également remercier tous les enseignants qui ont partagé leurs connaissances et leur expertise avec moi , et toutes les personnes qui mon aidé, et tous ceux qui m'ont soutenu , de près ou de loin, à la préparation de cette thèse . Sans oublier ma grande et petite famille, et mes amis .

Table des matières

Remerciements

Résumé

Abstract

Introduction générale13

CHAPITRE (I)

Le son et la parole

I -1 -1 Introduction17

I -1-2 Historique du son17

I -1 -3 Définition de son18

I -1 -4 Propagation des sons18

I -1 -5 Catégories des sons19

I -1 -6 Types d'ondes sonores.....20

I -1 -7 Caractéristiques du son20

1- Fréquence

2- Amplitude

3- Longueur d'onde

4-Vitesse21

5- L'intensité du son22

I -1 -8 Les types du son23

1 - Signaux de temps continus :23

2 - Signaux analogiques :23

3 - Signaux périodiques :23

La parole

I - 2 -1	Domaine d'études de prononciation	24
1	- Niveau phonétique	
2	- Niveau acoustique	
I - 2 -2	Structure physiologie de l'appareil auditif et phonétique humain	25
	l'appareil auditif humain	25
3- a	L'oreille externe	
3 -b	L'oreille moyenne	
3- c	L'oreille interne	
	l'appareil phonétique humain	26
I - 2 -3	Mécanisme production de parole	26
1-	La glotte :	27
2-	Le larynx :	27
3-	les cordes vocales :	28
I - 2 -4	La couche de voix	29
I - 2 -5	Les couches vocales pour les hommes et Les femmes	29
I - 2 -6	Conclusion de chapitre	30

CHAPITRE II : Le Signal de parole et leur caractéristiques , LA biométrie et la tatouage .

II - 1	Le signal	32
II - 1	théorie du signal.....	32
II - 3	Traitement du signal.....	33
1-	Historique	
2-	la définition de La transformation de Fourier	
3-	Le champ d'application :	
II - 4	Le bruit	34
II - 5 - 1	Classification des signaux.....	34

2 - Classification phénoménologique	35
II - 6 Le signal parole	
II - 7 Définition de signal parole	36
1- Les sons voisés	
2- Les sons non voisés	
II - 8 Les caractéristiques de signal parole	37
II - 11 Physiologie du système de production de la parole.....	44
II - 12 Production de Parole.....	45
II - 13 Modèle de production de la parole	45
1-Les approches articulatoires	
2-Les approches électriques	
II - 14 Modèle auto régressif.....	45
II - 15 Modèle Auto-Régressif à Moyenne Ajustée (ARMA).....	48
II - 16 la biométrie.....	48
1-Histoire de la biométrie	
2-Définition biométrie	
II - 17 les Modalités biométriques les plus importantes.....	50
II - 18 Fonctionnement d'un système biométrique.....	52
II - 19 Evaluation des Systèmes biométrique.....	52
II - 20 Evaluation des performances des Systèmes biométriques	52
II - 21 La précision d'un système biométriques	54
II - 22 Modalité empreinte digitale	55
II - 23 Modalité visage.....	56
II - 23-1 La fonctionnement de la reconnaissance faciale	56
2- L'histoire du système de reconnaissance faciale.....	57
II - 24 Modalité de reconnaissance par iris	58
1 Historique iris	58

	2	définition58
	3	Architecture générale de système reconnaissance d'iris	
		1- L'acquisition	
		2- La segmentation	
		3- La normalisation	
II - 25		Modalité ADN60
		1- Définition	
		2- Historique	
		3- Analyse de l'ADN	
II - 26		Modalité par voix62
		1- L'Historique62
		2- La Biométrie comportementale62
		3- les Avantages et inconvénients de la modalité voix62
		1- Les Avantages	
		2- Les Inconvénients	
II -26		Le Système biometrique62
II - 27		Les autres modalités biométriques63
II - 28		Les techniques de la transmission sécurisée de l'information	
II - 29		Cryptographie.....	64
II - 30		La stéganographie65
II - 31		Le tatouage (watermarking).....	65
		1- Historique de tatouage (watermarking)	
		2- Définition de tatouage numérique (watermarking)	
		3- Système de tatouage (watermarking)	
		4- Les Propriétés d'un système de tatouage (watermarking)	
		5- Compromis entre robustesse, imperceptibilité et capacité de tatouage.	
		6- Classification du tatouage numérique	

7- Les applications du tatouage numérique	
8- Protection des droits d'auteur	
II - 31-9 Les attaques menaçant le tatouage.....	70
II - 31-10 l'attaque par tatouage	70
1- Dynamique	
2- Filtre	
3 - Bruit	
4- Conversion	
5- Compression avec perte	
6 - Sample Permutation	
7 - Time stretch (également connu sous le nom de pitch shift)	
Conclusion de chapitre.....	71
Chapitre III Etude ,Experiences et discussion	
III -1 Introduction de chapitre	73
III -2 historique de Transform d'ondelette :.....	73
III -3 Transformée discrète en ondelettes (DWT).....	73
III -4 Transformée continue en ondelettes (CWT) :.....	75
III -5 Etude et discussion de la premier partie.....	75
1- Caractéristiques de la transformée en ondelettes	
2- organigramme de première partie	
3- discussion de première partie	
4- résultat de première partie	
III -6 Etude et discussion de la deuxième partie.....	79
1- Loi de Shannon	
2- Filtrage	
3- Caractéristiques du signal vocal	
III -7 Etude et discussion de la troisième partie	84

1-	discussion de troisième partie	
2-	organigramme de troisième partie (WER)	
3-	passage par zéro	
4-	organigramme de la troisième partie	
5-	Taux d'énergie d'ondelettes proposé par (WER)	
III -8	Etude et discussion de la quatrième partie	88
1-	definition	
2-	Traitement du signal primaire	
3-	L'étape de filtrage	
4-	segmentation de signal	
5-	Fénetre (windowing)	
6-	La fonction d'autocorrélation	
7-	Extraction et extraction de caractéristiques	
8-	La méthode d'analyse par prédiction linéaire LPC	94
9-	Les coefficients cepstraux prédictifs linéaires LPCC.....	95
10-	processus de comparaison et de performance.....	97
III -9	Etude et discussion de la cinquième partie	98
III -10	Watermark	101
III -11	Conclusion de chapitre troisième	102
Chapitre IV Resultat et Duscusions		
IV -1	Introduction de chapitre	104
IV -2	première partie	
1-	Résultats de 1 partie	104
2-	Conclusion.....	110
IV -3	Resultat partie 2	111
1-	Introduction	
2-	Conclusion	

IV -4	Resultat partie 3.....	113
IV -5	Resultat de partie 4.....	118
	1- pré-traitement	
	2- filtrage du signal	
	3- Framming (Couper le signal en trames)	
	4- fenêtrage (Windwing)	
	5- Energie	
	6- Energie acourt terme	
	7- Autocorrelation court -terme	
	8- Analyse le LPC	
IV -6	cinquième partie	
	1- Résultats de 1 partie	
	2- Conclusion	
IV -7	Conclusion et discussion de chapitre 4	130
	Conclusion générale.....	132
	Bibliographie	134

ملخص

تعد تقنية التعرف على الكلام من أهم تقنيات التعرف في النماذج الحديثة التي تم تطوير العديد من الأنظمة المختلفة من حيث الأساليب المستخدمة لاستخراج الميزات ، والتعرف على الكلام ، وهي مقسمة إلى جزئين: التعرف على الكلام ، والتحقق من المتحدث . لقد تناولت أطروحتنا في الجزء الثاني من الدراسة تجارب على تحديد جنس المتحدث من خلال نهج جديد قائم على خوارزمية ومعياري جديد لاستخراج مجموعة من القيم التي تحدد جنس المتحدث وحقق معدل التعرف على نوع المتحدثين يصل إلى 98٪ لتحديد نوع المتحدث ، هذا المعيار قريب في مجاله لقيم التردد الأساسي لإشارة الكلام واعتمد على نسبة الطاقة ، وعلى عدد تقاطعات مع الصفر للإشارة بعد تطبيق تحويل المستوى الثالث المويجي ، بعد اختيار أنسب موجات الأم ، والفرع الأنسب ، الذي كان موضوع دراسات وتجارب في الأجزاء الأولى. اما الجزء الاول من دراستنا للفصل 3 من الأطروحة ، هو تجارب على 38 فرعاً من 6 أمهات مختلفة يتم تطبيقها على 6 إشارات مختلفة من جنسين مختلفين ، و تمت مقارنة النتائج بين الأموجات الأم ، وفروعها لكلا الجنسين وبالنسبة للجزء الرابع فهو لتحديد المتكلم باستخدام المتنبي الخطي وبخصوص الجزء الخامس من الفصل الثالث فقد تم اضافة الوشم الرقمي للإشارة الصوتية لحمايتها من اي هجوم الكتروني باستخدام المصفوفات المكونة من خصائص الإشارة .

Abstract

Speech recognition technology is one of the most important recognition technologies in modern models. Many different systems have been developed in terms of the methods used to extract features, and to recognize speech, and it is divided into two parts: speech recognition, and speaker verification. Our thesis in the second part of the study dealt with experiments on determining the gender of the speaker through a new approach based on an algorithm and a new criterion to extract a set of values that determine the gender of the speaker and achieved a rate of recognition of the gender of the speaker up to 98% to determine the type of speaker, this criterion is close in its field to the values of The basic frequency of the speech signal depended on the energy ratio, and on the number of intersections with zero of the signal after applying the third level wavelet transform, after selecting the most appropriate mother waves, and the most appropriate branch, which was the subject of studies and experiments in the first parts. As for the first part of our study for chapter 3 of the thesis, it is experiments on 38 branches of 6 different mothers that are applied to 6 different signals of different sexes, and the results were compared between the mother wavelets, and their branches for both sexes, and for the fourth part it is to determine the speaker using linear prediction and Concerning the fifth part of the third chapter, a digital tattoo has been added to the audio signal to protect it from any electronic attack by using arrays composed of signal properties.

Keywords: Wavelets, DWT, Energy, Zero crossing, Watermark , speaker sex

INTRODUCTION GENERALE

Ces dernières années, le domaine de l'intelligence artificielle a connu un grand développement. Ce qui a permis d'économiser le temps et les efforts. Il est devenu facile pour la machine ou l'ordinateur de mettre en œuvre des commandes et des instructions vocales sans se référer aux applications traditionnelles qui nécessitent que l'être humain enseigne la machine par une communication directe. Les premiers logiciels de reconnaissance vocale sont basés sur un vocabulaire limité des mots, et des phrases, et ne peuvent pas les reconnaître que s'ils sont prononcés très clairement. Dans les recherches, la reconnaissance vocale ne cherche qu'à identifier la voix d'un locuteur. La reconnaissance automatique de la parole est devenue un vaste domaine. Elle est entrée dans la plupart des domaines de la vie à partir d'utilisations dans la technologie, les communications, le domaine de la sécurité, et diverses utilisations, telles que la conversion de la parole en textes écrits, et autres via une large gamme d'applications, par contre des programmes plus avancés ont la capacité d'accepter la parole naturelle, les dialectes et les différentes langues.

La reconnaissance de la parole fait partie des domaines de recherche importants, qui a eu beaucoup d'attention ces dernières années, et est devenu l'un des principaux défis de la de la vérification de la parole qui est d'assurer la sécurité du signal de parole. Étant donné que la parole est une forme de communication très personnelle et sensible, il est nécessaire d'empêcher l'accès non autorisé au signal vocal ou à son contenu. Dans les années 1940, des chercheurs ont mis au point un appareil spectrographe qui prend un enregistrement et trace l'intensité du son dans l'enregistrement à différentes fréquences. Dans les années 80, De nombreuses entreprises ont développé des systèmes de reconnaissance de la voix. Et dans les années 90, les entreprises qui travaillaient sur la voix comme IBM, ont mis au point un système pour le marché commercial. Selon la société américaine - International Biométrie Groupes (IBG), cette technologie biométrique détenait 4,3 % du marché en 2000 [1]. Selon leurs objectifs premiers, nous avons trois technologies de protection et de sécurité de l'information qui sont : La cryptographie, La stéganographie et le tatouage numérique, qui sont des technologies qui traitent la protection et la sécurité de l'information. Cette dernière a connu un grand développement en raison de la plus grande importance qu'il requiert dans les domaines de la sécurité et des communications cryptées dans les domaines sécuritaires et militaires. C'est à partir de l'année 1995 que le tatouage numérique a suscité beaucoup d'attention et s'est développé très rapidement. Il existe de nombreux sujets de recherches, des méthodes de travail et des systèmes pratiques qui ont été développés dans ce domaine. C'est à travers la création d'ateliers et des conférences dédiés à la sécurité de l'information, la sécurité, la dissimulation des données et la sécurité multimédia. Le magazine EURASIP a traité en 2006 un sujet sur la sécurité de l'information, qui a mis en lumière le dynamisme du domaine .

Le but de la recherche

L'objectif de cette recherche est de trouver des nouvelles voies, des méthodes pour identifier, et vérifier le locuteur, et aussi de diversifier la protection du signal de parole contre les différentes attaques de signal de parole, au moyen des méthodes de mesure du signal, et des nouvelles critères. La vérification, et la comparaison avec des recherches antérieures dans toutes les étapes de cette la recherche.

Les outils de recherche

Dans nos expériences, nous avons utilisé le programme 4K Video Downloader pour Windows pour télécharger, des clips vidéo, et des phrases anglais (clips audio) de format (mp4) à partir d'Internet, puis nous avons converti ces clips au format (wav) en utilisant Coverio en ligne, et Cloudconvert en ligne. Puis, on a découpé ces clips chargés en phrases de longueur définie à l'aide du programme (Audacity 2.4.2), avec une fréquence d'échantillonnage de $f_e=11025$ hz. Ensuite, nous avons analysé, et étudié ces signaux wav, en utilisant la boîte à outils wavlets d'un programme matlab a13 pour un ordinateur des spécifications suivantes :

PC Acer basé sur x67, processeur : Intel® Core™ i3-2348M CPU : 2.30GHz, Version de SMBIOS 2.7, Système d'exploitation win 10

Recherches et études antérieures

Il n'y a qu'un petit nombre des recherches antérieures sur la comparaison des ondelettes les plus appropriées pour la reconnaissance vocale. Les recherches antérieures sur le critère « Énergie retenue » est également similaire au critère utilisé dans notre recherche, avec la différence d'application au type de signal, et type d'ondelette mère et leurs branches. On peut aussi dire que la recherche liée au critère PRD (Percentage Root Mean Square Différence) est le plus utilisé, et c'est un critère d'évaluation de la différence entre les signaux, en plus il est facile pour la comparaison, et l'utilisation [2].

Parmi les études les plus proches de notre recherche figurent la recherche [3], [4] [5] chacune d'entre elles portait sur une comparaison des ondelettes mères les plus appropriées dans les analyses du signal ECG. Cette recherche a fait une étude comparative sur 4 (quatre) types d'ondelettes mères pour choisir l'ondelette la plus appropriée pour l'extraction des caractéristiques.

De nombreux algorithmes sont développés pour l'extraction des caractéristiques acoustiques du signal parole qui forment l'élément clé de la détection, et l'identification du locuteur. La fréquence fondamentale f_0 est l'une des plus importantes de ces caractéristiques. Grâce à notre nouvelle approche que nous avons abordée dans l'article qui nous a été publiée en 2022, et qui montre que la vitesse de calcul direct du taux d'énergie des ondelettes ou << Wavelet Energy Rate >> (WER), est plus rapide, et plus facile pour déterminer le sexe de l'orateur, par rapport aux méthodes qui dépendent de la plage à laquelle appartiennent les valeurs calculées en extrayant les valeurs f_0 des courbes, et dépend de la valeur ajoutée à la valeur extraite de la courbe. Comme la méthode de 1/3 maximum central qui est basée sur l'autocorrélation de l'analyse LPC (Linear Predictive Coding). La spécificité de cette méthode est de rejeter les composantes des fréquences supérieure à 600 Hz qui peuvent être corrompues par le bruit. Après l'élimination des hautes fréquences [6] En plus notre approche est plus facile que la méthode d'extraction de fréquence fondamentale basée sur les fonctions d'autocorrélation sans fenêtre et normalisée [7], qui présente un algorithme fondamental d'estimation de fréquence d'un signal de parole bruyant.

Contenu de cette recherche et les étapes de recherche

La thèse contient quatre chapitres, dans le premier chapitre nous traitons le concept de son et ses caractéristiques, et aussi la parole, ainsi que des caractéristiques du système auditif et vocal humain. Nous avons également abordé les étapes de la production de la parole, les organes humaines responsable de la production de la parole. Nous avons présenté une comparaison entre le son dans la nature, et le son des autres êtres vivants, ainsi que le signal parole des deux sexes.

Pour le deuxième chapitre, nous avons présenté une étude détaillée sur le concept et les caractéristiques du signal de la parole. Nous avons également présenté une étude sur les signaux, sur la biométrie, leur histoire, leurs types, et leurs caractéristiques avec une comparaison entre les différentes biométries et leurs utilisations. Au cours de ce chapitre, nous avons abordé les caractéristiques des tatouages numériques, et leur utilisation dans le signal parole. Le troisième chapitre, qui est un chapitre d'application, étude, et expériences. Il est divisé en quatre phases, qui sont des articles de recherche qui ont été publiés chacun séparément. La première phase, du troisième chapitre, traite la façon de choisir l'ondelette mère, et la branche appropriée dans le troisième niveau de la transformation en ondelettes discrète lorsqu'elle est appliquée au signal de parole, et comparée à d'autres signaux tels que l'ECG. Cette étude a été publiée dans la revue des sciences et technologies, Publié sous le titre « L'ondelette mère est la meilleure dans le traitement du signal de parole ». Quant à la deuxième phase du troisième chapitre, est consacrée à une comparaison de calcul de l'énergie retenue la plus élevée dans le troisième niveau de transformée en ondelettes discrète chez les deux sexes parlants, et la comparaison des taux d'énergie d'un signal du parole de niveau trois à l'aide de (DWT) de l'ondelette mère Haar, Meyer, biorthogonal, coiflets, symlets, daubechies. Pour la troisième phase de chapitre 3, nous avons présenté une nouvelle approche, et un nouveau critère afin de déterminer le locuteur des deux sexes selon la plage des valeurs par rapport à la plage des valeurs de la fréquence fondamentale, que nous avons nommé « Taux d'Énergie d'Ondelettes » ou « Wavelet Energy Rate » (WER). C'est une nouvelle approche pour l'identification et la vérification du sexe du locuteur à l'aide de l'énergie de premier niveau DWT, et du passage à zéro. Cette recherche a été aussi publiée dans la revue « Engineering, Technology & Applied Science Research Vol. 12, No. 6, 2022, 9570-9578 ». La quatrième phase du troisième chapitre a été consacrée à une nouvelle méthode de sécurisation du signal de parole au moyen d'un watermark. C'est une nouvelle méthode destinée à diversifier la protection du signal de parole contre les attaques extérieures et autres.

Le quatrième chapitre explique les résultats des tableaux, les données, et les chiffres obtenus à partir du troisième chapitre, on y a un résumé et une discussion complète de ces résultats.

CHAPITRE (I)
LE SON ET LA PAROLE

Le son

I-1-1 Introduction

Dans ce premier chapitre, nous avons dressé un bref historique du son depuis le XVIIe siècle, donné une définition, étudié les caractéristiques les plus élémentaires du son, comparé les différents types des sons produits par les organismes vivants, la nature et les instruments de musique. Nous avons également présenté les systèmes auditifs et vocal de l'être humain, défini leurs composants et discuté des organes importants responsables de l'audition dans l'oreille humaine, ainsi que des organes responsables de la production de la parole dans le corps humain. En outre, nous avons décrit les étapes de la production du son et de la parole, comparé des modèles équivalents et similaires d'aides auditives et d'aides à la parole, et présenté une comparaison de la portée auditive entre les humains et d'autres organismes.

I-1-2 Historique du son

L'histoire du son remonte à l'Antiquité, lorsque les humains ont découvert pour la première fois le pouvoir de la vocalisation pour communiquer et faire de la musique. Au fil du temps, les gens ont développé des instruments pour amplifier et améliorer le son, comme les anciens amphithéâtres grecs et les colisées romains.

Au début du XVIIe siècle, Marin Mersenne (1588-1648), un religieux et mathématicien français, a présenté en 1627 ses recherches sur l'acoustique et a réalisé la première mesure de la vitesse du son. Il a également publié son premier traité de l'harmonie universelle sur la musique, où il a rassemblé les théories des anciens et des modernes. Mersenne a déterminé la relation entre la hauteur musicale et la fréquence, quantifié l'expérience de Pythagore et montré que la hauteur du son est inversement proportionnelle à la longueur de la corde et proportionnelle à la racine carrée de la tension.

L'un des premiers pionniers de l'étude du son fut Robert Boyle (1627-1691), est un physicien et chimiste irlandais, est le père de la philosophie naturelle moderne, dans les années 1660 mena des expériences sur la relation entre la pression atmosphérique et le son.

À la fin du XVIIe siècle, le scientifique Isaac Newton (1642-1727), mathématicien et physicien anglais, a expliqué la propagation du son et présenté en 1687 dans son livre "Philosophiæ naturalis principia mathematica" les principes du mouvement des corps. Au XVIIIe siècle, en 1710, Isaac Newton a également proposé la théorie de la vitesse du son.

Au début du XIXe siècle, le mathématicien et physicien français Jean Baptiste Joseph Fourier (1768-1830) a publié son livre "Théorie analytique de la chaleur" en 1822. Grâce à ses recherches, Fourier a introduit l'analyse de Fourier, en particulier la série de Fourier, qui propose la théorie de l'analyse des ondes complexes en termes des ondes simples qui les composants. Plus tard dans le même siècle, Hermann von Helmholtz, un scientifique et physicien allemand (1821-1894), a mené des recherches révolutionnaires sur la nature des ondes sonores, développant le concept de résonance et découvrant les principes de l'harmonie musicale ainsi que la théorie de l'audition et le timbre des sons. En 1877, John William Strutt

Rayleigh (1842-1919), un physicien anglais et lauréat du prix Nobel de physique en 1904, a écrit "The Theory of Sound", où il présente les principes du son. [8] ,[9]

Au XXe siècle, l'étude du son continue d'évoluer avec le développement des nouvelles technologies telles que le microphone, le phonographe et le haut-parleur. Cela a conduit à des avancées majeures dans des domaines tels que l'enregistrement musical, les télécommunications et la reconnaissance vocale.

La science de l'acoustique a continué de progresser au cours du XXe siècle. Le physicien hongrois Georg von Békésy (1899-1972) a utilisé des modèles pour comprendre les mécanismes de l'audition et analysé la perception des fréquences pour déterminer la tonalité du son. Les ultrasons ont été étudiés dans les années soixante et développés pour des applications telles que la production et l'imitation des sons, ainsi que les domaines des champs de fréquence ultrasonores. Aujourd'hui, le son continue de jouer un rôle crucial dans notre vie quotidienne, allant de la musique que nous écoutons aux sons de la nature, en passant par la technologie que nous utilisons pour communiquer et nous connecter les uns aux autres [8] ,[9] .

I-1-3 Définition de son

Le son est une onde acoustique qui se propage à travers un milieu élastique, tel que l'air. C'est la sortie audible des vibrations d'un corps dans un milieu spécifique. D'un point de vue physique, le son est une énergie qui se propage sous forme de vibrations dans un milieu compressible (liquides, gaz, air, solides). Cette vibration crée une perturbation dans le milieu environnant, qui se propage sous forme d'ondes de pression. Les ondes sonores sont perçues par l'oreille humaine comme des sensations auditives, et elles sont utilisées dans de nombreux domaines tels que la musique, la communication, la médecine et la surveillance. Un son est pur lorsque sa fréquence de vibration est bien définie. Les ondes sonores ont des caractéristiques telles que la fréquence, l'amplitude et la longueur d'onde qui déterminent la façon dont elles sont perçues par les humains et peuvent être utilisées pour la communication, le divertissement, la recherche scientifique et diverses autres fins.

De plus, certains insectes, tels que les abeilles, produisent des sons en faisant vibrer leurs ailes dans l'air, tandis que d'autres insectes produisent des sons en déplaçant ou en frottant des parties de leur corps. En outre, différents instruments de musique émettent des sons lorsqu'ils sont utilisés. [10]

Parmi les facteurs qui affectent la transmission du son et sa vitesse, on trouve la nature du corps dans lequel le son est transmis, la température et la densité. Les molécules du corps servent de support ou de porteur pour le son et l'énergie. Elles transmettent l'énergie et le son.

I-1-4 Propagation des sons

La vibration provoque le déplacement des molécules du milieu. Cependant, ces molécules ne se déplacent pas de manière globale, elles vibrent autour de leur position d'équilibre sans qu'il y ait de déplacement net de matière. La propagation du son est un transfert d'énergie, et lorsque la perturbation prend fin, les molécules reviennent à leur position d'origine. C'est ce que nous appelons une onde sonore. Au fur et à mesure que l'onde sonore se propage, l'énergie qu'elle transporte se dissipe et diminue. On peut représenter la propagation des ondes sonores par des

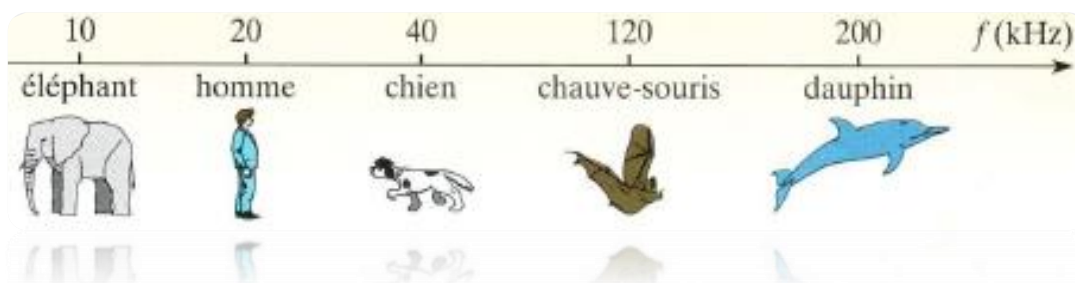
cercles concentriques qui entourent le point central où tombe un objet, la figure (1-1) présente l'exemple, comme un caillou lorsqu'on le jette dans un milieu calme, comme de l'eau.



Figure (1-1) : la propagation des ondes à la surface [11]

La plage d'audition de l'être humain est limitée et dépend de la fréquence du son, qui va de 20 Hz à 20 000 Hz. Ainsi, les sons de 20 Hz à 200 Hz sont des sons à fréquence grave (lente), tandis que ceux de 2001 Hz à 20 000 Hz sont des sons à fréquence aiguë (haute). Cette plage d'audition varie d'un individu à l'autre. L'être humain ne peut pas entendre les sons les plus bas en dessous de la fréquence de 20 Hz, correspondant à 0 dB de décibels, qui sont appelés infrasons (les fréquences situées en dessous de cette plage). Ce sont des fréquences que les humains ne peuvent ni entendre ni sentir, contrairement aux animaux tels que les chats, les chiens et les rongeurs tels que les rats. Par exemple, les glissements de la croûte terrestre qui conduisent aux tremblements de terre et aux volcans. Les scientifiques ont prouvé que les sons d'une intensité inférieure à 30 dB sont masqués par d'autres sons pour l'être humain et ne retiennent pas leur attention [10].

De même, les fréquences supérieures à 20 000 Hz et au-delà de 130 dB sont appelées des ultrasons (ondes ultrasonores), qui peuvent causer des dommages au système auditif humain (l'oreille) lorsque leur valeur augmente. Cela est indiqué dans le tableau (1-1), et le figure (1-2).



Figure(1-2) : Limite des fréquences audibles selon les espèces [12]

I-1-5 les catégories des sons

Les sons peuvent être divisés en trois catégories :

1 - (Au-dessus) Ultrason

Ce sont des ondes dont les fréquences causent de la douleur au système auditif humain supérieur à 20000Hz et se caractérisent par leurs courtes longueurs.

2 - Les ondes audio

Ces fréquences se situent dans la gamme audible pour l'ouïe humaine, soit entre 20 Hz et 20000 Hz.

3 - (Au-dessous de son) infrasons

Sa fréquence est inférieure à 20 Hz, et les humains ne peuvent pas l'entendre et les autres créatures peuvent l'entendre.

I-1-6 Les Types d'ondes sonores

Il existe différents types d'ondes sonores, qui sont déterminés par la manière dont les particules du milieu vibrent lorsqu'elles sont traversées par l'onde. deux principaux types d'ondes sonores les plus courantes sont :

Les ondes longitudinales sont des ondes dans lesquelles les particules du milieu se déplacent parallèlement à la direction de propagation des ondes. telque Les ondes sonores dans l'air.

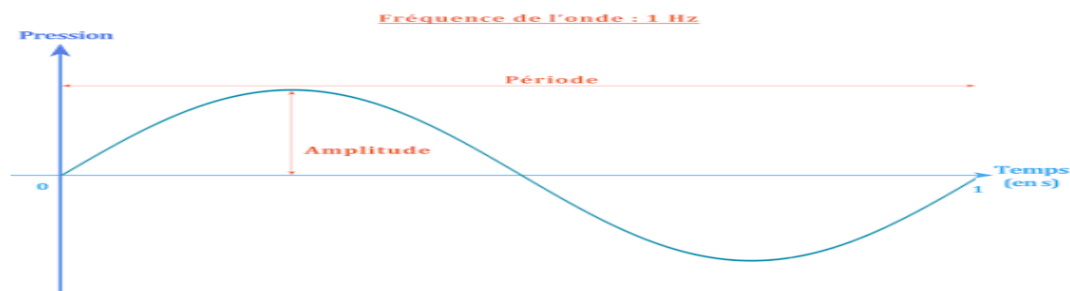
Les ondes transversales sont des ondes dans lesquelles les particules du milieu se déplacent perpendiculairement à la direction de propagation des ondes.telque Les ondes lumineuses , les ondes sur une corde de guitare . En général, les ondes sonores sont des ondes longitudinales car elles traversent un milieu en comprimant et en dilatant les molécules du milieu dans la même direction que l'onde se déplace. Cependant, il existe certains types d'ondes sonores, telles que les ondes de surface, qui ont à la fois des composantes transversales et longitudinales [13].

I-1-7 Les Caractéristiques du son

Les caractéristiques du son ils sont nombreuses, ce qui permet aux chercheurs de l'étudier et de le distinguer. Le son peut être analysé comme un signal qui varie dans le temps. Parmi les caractéristiques du son , la fréquence, l'amplitude, la longueur d'onde et la vitesse .

1- La Fréquence :

La fréquence est le nombre de cycles ou périodes d'une onde qui se produisent par seconde, exprimée en unité de temps. La fréquence est utilisée pour décrire la hauteur d'un son, c'est-à-dire sa hauteur tonale, qui dépend de la fréquence de l'onde sonore. Si la fréquence est élevée, le son est aigu, tandis que si la fréquence est basse, le son est grave. L'unité de mesure de la fréquence est le hertz (Hz). La fréquence est égale à la vitesse de l'onde sonore divisée par la longueur d'onde. La fréquence ne change pas lorsqu'on passe d'un milieu à un autre, le figure(1-3) présente fréquence d'onde = 1Hz [10]



Figure(1-3) : fréquence d'onde = 1Hz

2- L'Amplitude :

L'amplitude ou le volume sonore est mesurée en décibels (dB). Exprime la force du signal et représente la hauteur de la longueur d'onde ou le point la plus haut à partir sa position d'équilibre donc la hauteur du son.

3- Longueur d'onde :

La longueur d'onde est la distance entre deux points opposés de deux ondes successives ou la distance parcourue pendant une période complète, détermine sa fréquence et sa hauteur, mesurée en mètres (m). les longueurs d'onde plus longues produisant des sons plus graves. La longueur d'onde sonore peut être calculée à l'aide de la formule (1-1)[10].

$$1-1 \quad \lambda = c/f$$

λ : La longueur d'onde sonore (m)

c : La vitesse de propagation du son (m/s)

f : la fréquence du son (Hz)

4- La Vitesse :

Le changement de milieu de propagation du son entraîne une modification de la vitesse du son. En effet, plus les particules de matière sont proches les unes des autres dans le milieu de propagation, plus la transmission des vibrations sonores est rapide. Ainsi, la vitesse du son est plus faible dans les gaz que dans les liquides, et plus faible dans les liquides que dans les solides. Le tableau (1-1) présente les vitesses du son dans différents milieux. La célérité (vitesse de propagation) du son est la même dans tous les milieux. Dans l'air, la vitesse du son est de 343 m/s à une température de 20°C, et de 331 m/s à 0°C. Dans l'eau, la vitesse du son est de 1407 m/s à 0°C, et de 1450 m/s à 15°C. La pression et la température jouent un rôle important dans la propagation des ondes sonores dans l'air. La vitesse du son est considérablement plus faible que celle de la lumière, elle est estimée à 299 792 km/s (mesure effectuée par le Bureau international des poids et mesures). La formule (1-2) donne une approximation de la vitesse du son dans l'air [14] .

La vitesse de son dans les corps	m/s
Aluminium	5000
Le bois	4110
Le verre	4540
Briques	3650
Acier	5200

le tableau (1-1) la vitesse de son dans les corps.

$$1-2 \quad c = 20.05\sqrt{T}$$

C : la vitesse du son (m/s)

T : la température (°C)

5 - L'intensité du son :

L'unité d'intensité du son est le watt par mètre carré (W/m²). Cela signifie que l'intensité sonore est une mesure d'énergie, correspondant à la quantité d'énergie sonore traversant chaque unité une surface unitaire perpendiculaire à la direction de propagation du son .Cependant, comme la gamme d'intensités sonores pouvant être perçues par l'oreille humaine est très large, une échelle logarithmique appelée échelle des décibels (dB) est utilisée pour exprimer les niveaux d'intensité sonore [15].L'échelle des décibels est basée sur le seuil d'audition, qui est le niveau d'intensité sonore le plus bas pouvant être perçu par l'oreille humaine [15]. Le seuil d'audition est défini à 0 dB, et chaque augmentation de 10 dB représente une multiplication par dix de l'intensité sonore [14]. Par exemple, un son avec un niveau d'intensité de 40 dB est dix fois plus intense qu'un son avec un niveau d'intensité de 30 dB.

Le tableau (1-2) présente les niveaux d'intensité sonores audibles par l'oreille humaine. L'intensité sonore est définie par la formule (1-3).

$$1-3 \quad I = w/(4\pi r^2)$$

I : Intensité accoustiques (W/m^2)

w : puissance accoustiques (w)

r : distance de la source et un point de mesure (m^2)

Son	l'intensité du son watts/m ²	l'intensité du son dB
Son aigu .seuil de la douleur	10^{-2}	140db
Insupportable	10^{-4}	120db
Son Très fort	10^{-6}	110db
Son fort (Bruits d'usine)	10^{-7}	90db
	10^{-9}	70db
Son de parole	10^{-10}	60db
	10^{-11}	50db
Son faible à 100 Khz	10^{-12}	40db
Son Très faible à 10 Khz	10^{-14}	20db
Le son le plus faible que l'oreille humaine puisse entendre à 3 Khz	10^{-16}	0db
	10^{-17}	-10db

tableau (1-2) : l'intensité du son au niveau de l'appareil auditif humain [16] .

I-1-8 Les types du son

Les signaux peuvent être classés en fonction de diverses caractéristiques tels que la nature du signal analogique ou numérique , ou le domaine temporel ou fréquentiel, soit de leur continuité , leur déterministe ou aléatoire , etc.

1 - Signaux de temps continus : Ces signaux sont définis dans le temps . (exemple signal électrique ou sonore) . D'autre part les Signaux à temps discret , ces signaux ne sont définis qu'à des instants discrets et les valeurs des signaux sont généralement représentées à l'aide d'une séquence de nombres.

2 - Signaux analogiques : Ces signaux sont continus en temps et en amplitude (exemple: signal électrique analogique ou sonore analogique) .D'autre part les Signaux numériques , ces signaux sont discrets en temps et en amplitude , utilisés pour représenter des signaux analogique échantillonnés (exemple : les signaux audio et vidéos numériques) . Les signaux numériques sont généralement représentés à l'aide d'une séquence de nombres binaires.

3 - Signaux périodiques : Les signaux périodiques se répètent à intervalles de temps fixe . Ils peuvent être représentés par une fonction mathématique périodique (exemple : signal sinusoïdal, signal carré) , D'autre part les Signaux aléatoires , ne peuvent pas être décrits par une équation mathématique ou un algorithme (exemples : bruit , . . .)[17].

Le parole

I-2-1 Définition

La parole est un moyen de communication fondamental entre les êtres humains . La parole est un élément essentiel de la communication verbale et non verbale, car elle permet de transmettre des informations, des émotions et des idées. Elle est produite par les vibrations des cordes vocales, qui génèrent des propriétés acoustiques spécifiques telles que la fréquence, l'intensité et le timbre. La voix humaine est considérée comme complexe et distinctive, car elle reflète les caractéristiques individuelles de chaque personne, telles que le sexe, l'âge, la santé et l'origine géographique.

"" الآية 30 من سورة محمد ""

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

[وَلَوْ نَشَاءُ لَأَرَيْنَاكُمْ فَلَعَرَفْتَهُمْ بِسِيمَاهُمْ ۚ وَلَتَعْرَفَنَّهُمْ فِي لَحْنِ الْقَوْلِ ۗ وَاللَّهُ يَعْلَمُ أَعْمَالَكُمْ]

(وَلَتَعْرَفَنَّهُمْ فِي لَحْنِ الْقَوْلِ ۗ)

أي :

فيما يبدو من كلامهم الدال على مقاصدهم ، يفهم المتكلم من أي الحزبين هو بمعاني كلامه وفحواه ، وهو المراد من لحن القول ، كما قال أمير المؤمنين عثمان بن عفان ، رضي الله عنه : ما أسر أحد سريرة إلا أبداها الله على صفحات وجهه ، وفلتات لسانه . وفي الحديث " : ما أسر أحد سريرة إلا كساه الله جلبابها ، إن خيرا فخير ، وإن شرا فشر . " وقد ذكرنا ما يستدل به على نفاق الرجل ، وتكلمنا على نفاق العمل والاعتقاد في أول " شرح البخاري " ، بما أغنى عن إعادته هاهنا . وقد ورد في الحديث تعيين جماعة من المنافقين.

----- تفسير ابن كثير -----

Les scientifiques ont prouvé que chaque personne possède une voix spécifique qui ne peut être imitée. La voix est considérée comme l'une des empreintes digitales qui peut être utilisée dans des nombreux domaines, y compris le domaine de la sécurité, pour identifier le locuteur. Le nombre des caractéristiques vocales est estimé à 100, comprenant des caractéristiques physiologiques liées aux différences entre les systèmes vocaux humains, tels que la longueur des cordes vocales, des sinus, des lèvres et autres, ainsi que des caractéristiques liées à la personnalité de l'orateur, notamment le ton, la vitesse et le timbre de la parole.

I-2-1 Le Domaine d'études de prononciation

Le domaine d'étude de la prononciation est un sous-domaine de la phonétique, qui est la branche de la linguistique qui s'intéresse aux sons de la parole ,qui concerne la production, la transmission et la réception des sons de la parole dans une langue particulière. La prononciation, en particulier, se concentre sur l'articulation des sons et la manière dont ils sont perçus et interprétés par les auditeurs. Les études de prononciation impliquent souvent l'analyse acoustique et la transcription phonétique des sons de la parole, ainsi que l'observation de la position de la langue, des lèvres et de la mâchoire lors de la production de ces sons.[18],[19]

1 -Le Niveau phonétique

Le niveau phonétique se réfère à l'étude des sons du langage parlé, à savoir les phonèmes. Au niveau phonétique, le langage est composé de sons, également appelés phonèmes. Les phonèmes sont les unités de base du langage parlé qui peuvent être combinées pour former des mots. Il existe trois principaux aspects de la phonétique : la phonétique articulatoire, la phonétique acoustique et la phonétique auditive [20].

2- Le Niveau acoustique

Le niveau acoustique est une mesure de l'intensité sonore d'un signal sonore. Il est généralement utilisé pour décrire la puissance sonore ou la pression sonore d'un son. Le niveau acoustique peut être utilisé pour mesurer la force ou la faiblesse d'un signal sonore . Il est souvent utilisé pour évaluer les niveaux de bruit et les niveaux sonores dans divers environnements .Cette mesure implique l'étude de la fréquence, de l'amplitude et de la durée des ondes sonores, ainsi que de leurs propriétés spectrales.[21].

I-2-2 Structure physiologie de l'appareil auditif et phonétique humain

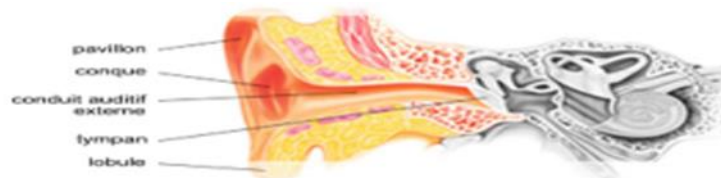
L'être humain possède un système complexe pour communiquer avec les autres êtres humains. Ce système est composé de deux parties complexes et profondes dans leur fonctionnement : le système auditif et le système de phonation et de parole.

1- l'appareil auditif humain

Le système auditif humain est un mécanisme complexe du corps humain qui nous permet de percevoir et de traiter le son . Le système auditif humain est un récepteur complexe , Il implique l'oreille, le nerf auditif et le cerveau, qui travaillent ensemble pour convertir les ondes sonores dans l'air en signaux électriques que le cerveau interprète comme un son.ce système auditif a la capacité d'identifier les sons dans une gamme limitée, allant de 20 Hz à 20 000 Hz selon le tableau (1-2). L'oreille est composée de trois parties : l'oreille externe, l'oreille moyenne et l'oreille interne.

1 - a L'oreille externe :

L'oreille externe , est constituée de la membrane tympanique et de l'armature auriculaire externe, son rôle est d'amplifier et de capter le son , puis de l'envoyer sous forme de vibrations au conduit auditif . L'oreille externe a également un rôle de protection du tympan le long avec d'autres rôles.[22]



Figure(1-4) : L'oreille externe

1- b L'oreille moyenne :

L'oreille moyenne , se compose du marteau, de l'enclume et de l'étrier, sa tâche est de transmettre des signaux ou des vibrations sonores à l'oreille interne , et travaille à réduire les sons forts grâce à l'action des muscles de l'étrier et du tympan. [22]



Figure(1-5) : L'oreille moyenne

1- c L'oreille interne :

L'oreille interne est responsable de la conversion du son en signaux nerveux , qui est interprété par le cerveau et transmis au cerveau via le nerf auditif et se compose des cellules de la cochlée, du vestibule [22] .



Figure(1-6) : L'oreille interne

2- l'appareil phonétique humain

Le système articulatoire humain nous permet d'exprimer différents sons de la parole dans notre langue et constitue un dispositif essentiel pour la communication humaine et la communication d'idées. Il se compose des organes impliqués dans la production des sons de la parole, ce sont les poumons, le larynx, le pharynx, la cavité buccale et la cavité nasale, ainsi que la langue, les lèvres et les dents.

I-2-3 Mécanisme production de parole

La production de la parole et du son chez l'être humain est un processus complexe qui implique plusieurs organes constituant le système vocal humain. La parole est produite à travers différentes cavités spécifiques par lesquelles l'air provenant des poumons passe, notamment la bouche et le nez, en passant par la trachée située à son extrémité supérieure, puis le larynx. Le système nerveux intervient dans la régulation de la respiration, ce qui a un effet sur la voix d'une personne en termes de tonalité et de volume. Ainsi, la voix d'une personne en colère sera différente de celle d'une personne triste, d'une personne souffrant d'un rhume ou de tout autre état psychologique influant sur la voix.

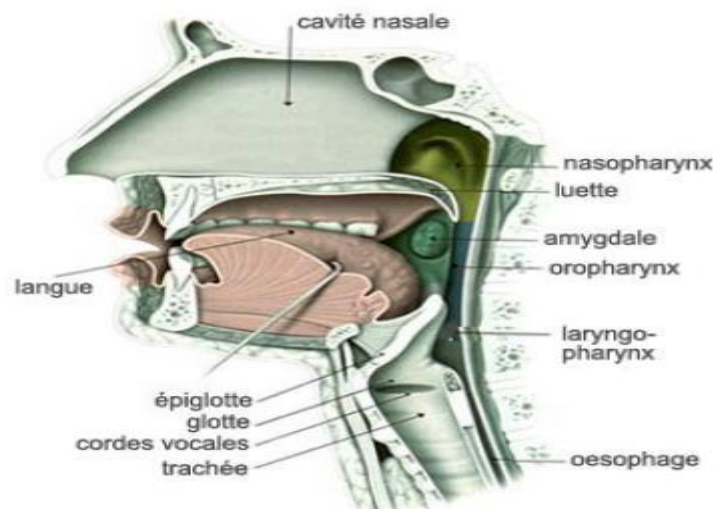
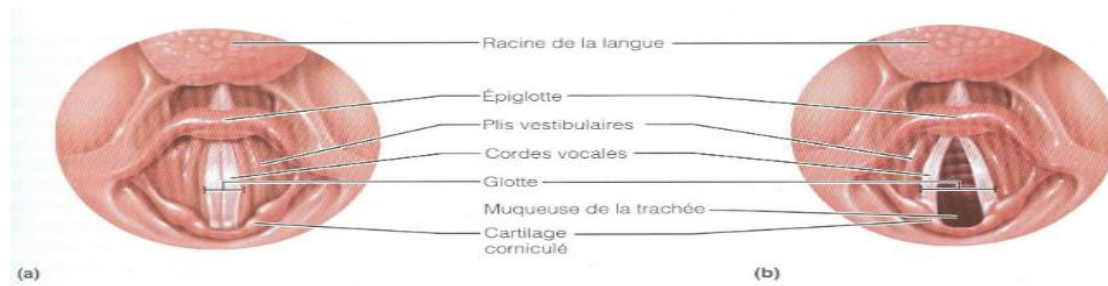


Figure (1-7) : : Vue de face de l'anatomie l'appareil vocal humain[23]

1- La glotte :

La glotte est constituée de deux cordes vocales, situées dans le larynx et situées au-dessus de la cavité de l'épiglotte et La trachée joue un rôle clé dans le transport de l'air des poumons et relie le larynx aux poumons.[24]



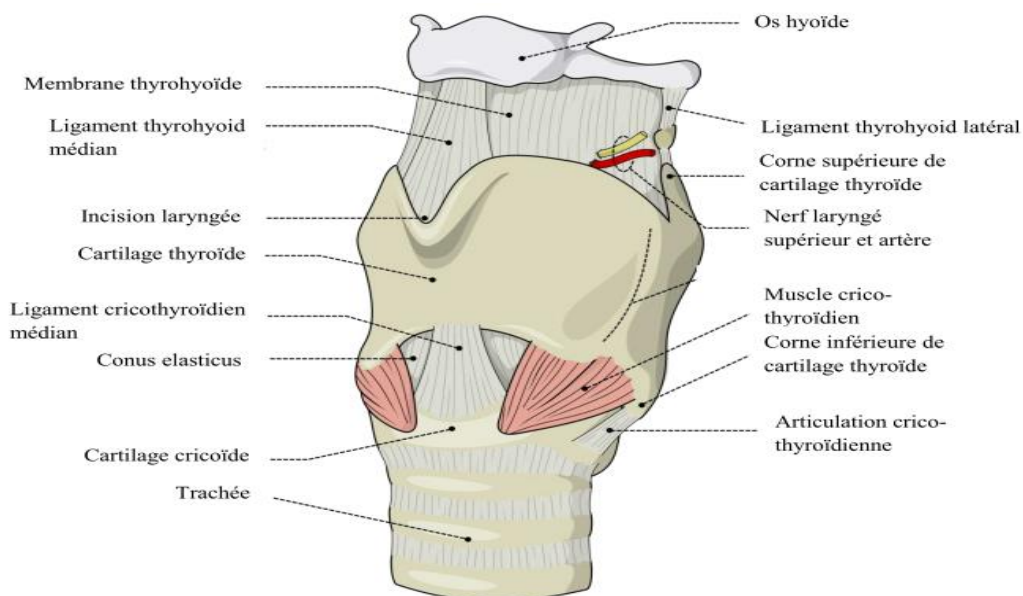
Figure(1-8) : Vue laryngoscopique de la glotte [25]

2- Le larynx

Le larynx est situé dans la région du cou, au-dessus de la trachée. Il joue un rôle important dans la production de la parole en se composant de muscles et de cartilages mobiles qui entourent une cavité. Selon le tableau, la taille du larynx est en moyenne 20% plus grande chez les hommes que chez les femmes.

Le larynx contient les cordes vocales qui, lorsqu'elles se rapprochent et vibrent, produisent un signal périodique sous la forme d'impulsions de pression périodiques sur la conduit vocal pendant la prononciation. Le mouvement des cordes vocales vers l'avant et vers le haut aide également à la déglutition et joue un rôle dans la respiration, en permettant l'entrée d'air dans les poumons lorsque la glotte s'ouvre et que les cordes vocales s'éloignent l'une de l'autre.

En plus de son rôle dans la parole, le larynx est également important dans la régulation de la toux et dans la protection des voies respiratoires contre les particules étrangères. La figure (1-9) représente les organes du larynx.



Figure(1-9) : Vue de face de l'anatomie larynx[26]

3- Les cordes vocales :

Les cordes vocales sont situées dans le larynx et sont les deux principaux composants responsables de la production de la voix. Ce sont deux muscles symétriques qui agissent comme une valve. Chez les femmes, leur taille est plus petite que chez les hommes, et leur longueur est d'environ 18 mm chez les femmes et 22 mm chez les hommes. Lorsque les cordes vocales se rapprochent, elles ferment le larynx et empêchent les aliments de pénétrer dans la trachée. Lorsqu'elles vibrent rapidement en s'ouvrant et en se fermant, elles produisent des impulsions de pression périodiques sur le tractus vocal, qui sont converties en énergie sonore. La fréquence fondamentale du son est le nombre des vibrations par seconde résultant des mouvements rapides d'ouverture et de fermeture des cordes vocales, exprimé en Hertz. .[27]

Les longueurs des cordes vocales d'un homme adulte diffèrent des longueurs des cordes vocales d'une femme adulte et diffèrent des cordes vocales des enfants . Le tableau (1-3) présente ces longueurs .

sexe	homme adulte	femme adulte
Longeur cord vocal	22 mm	18-20mm

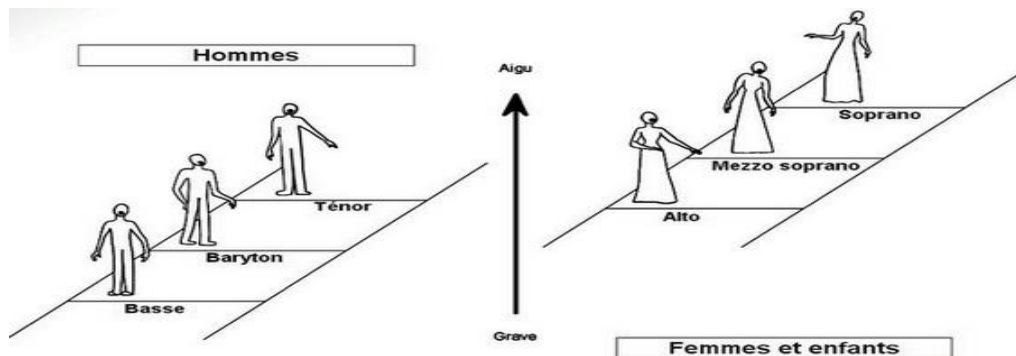
Le tableau (1-3) : Les longueurs des cordes vocales

I-2-4 La couche de voix

La taille du larynx et la longueur des cordes vocales ont une influence sur la fréquence fondamentale de la voix, qui est inversement proportionnelle à ces dimensions. Chez les femmes, la fréquence fondamentale est en moyenne plus élevée que chez les hommes, en raison de la plus petite taille du larynx et de la plus courte longueur des cordes vocales chez les femmes, comme indiqué dans le tableau suivant (1-4). [28] .

I-2-5 Les couches vocales pour les hommes et Les femmes

Les couches vocales des femmes et des hommes sont disposées du plus haut au plus bas, et le ton des femmes reste plus élevé que celui des hommes, selon le tableau (1-4) et la figure (1-10)



Figure(1-10) : ton de deux sexe du plus haut au plus bas

	Larynx	Couche acoustique	Longueur cord vocal
Femme	plus petite	Haut	plus court
Homme	20 % Plus grande	plus bas	plus long
Enfants	Fille meme avec Enfant	/	/

Le tableau (1-4) :Comparaison par sexe

Conclusion de chapitre

À travers ce qui a été discuté dans ce chapitre, nous pouvons dire qu'il y a un écart entre la parole et le son. on peut dire que la parole est une forme de communication humaine qui utilise le son comme moyen principal. Il utilise les cordes vocales, la langue, le larynx et d'autres organes pour produire des sons et les combiner en mots , et Un signal de parole est un signal audio généré par le système vocal humain et constitué d'un ensemble de composants vocaux, notamment des phonèmes, des syllabes , et les modèles d'intonation. D'autre part, le son est un phénomène physique qui peut être décrit comme des vibrations qui se propagent à travers un milieu tel que l'air, l'eau ou des matériaux solides. Ces vibrations peuvent être entendues par l'oreille humaine comme un son audible. Un signal audio est une forme d'onde complexe caractérisée par sa fréquence, son amplitude et sa durée .

CHAPITRE (II) LE SIGNAL DE PAROLE

Le signal de la parole

Dans le deuxième chapitre, nous allons classifier les signaux et définir des notions de puissances. Nous aborderons les phénomènes aléatoires et des signaux dits déterministe.

Dans ce chapitre, nous avons étudié le traitement du signal de la parole, et nous avons abordé clairement ses caractéristiques, Nous avons abordé la définition et l'explication des organes et des caractéristiques de l'appareil audio, et nous avons présenté le modèle équivalent pour celui-ci, ainsi que l'appareil audio et le modèle équivalent pour celui-ci. Dans ce deuxième chapitre, Nous avons également abordé un bref historique de la biométrie, et sa définition, et ses caractéristiques détaillées, et nous avons présenté un aperçu détaillé des principales biométries utilisées actuellement, et par le passé. et nous avons présenté Une ventilation détaillée Les étapes les plus importantes du développement de la sécurité des signaux vocaux les plus importantes, du Cryptographie en passant par stéganographie à le tatouage numérique.

II-1 - Le signal

Un signal est une grandeur physique mesurable dotée d'une unité et dépendant du temps, ou d'une autre variable, qui sert de support à l'information et qui représente physiquement l'information qu'il transporte de sa source à son destinataire en fonction de sa nature physique. Cette nature physique peut être très variable : optique, électronique, acoustique, courant, tension, force, température, pression, fréquence d'une onde électromagnétique (lumière visible, infrarouge, radio...), etc. De nos jours, la plupart des grandeurs physiques sont converties en signaux électriques, qui sont ensuite codés en signaux numériques binaires. Il existe très peu de mesures totalement analogiques. Ainsi, on peut dire que le rôle majeur du signal est la communication dans des disciplines telles que l'automatique, l'informatique et l'électricité. [29][30] Pour qu'il y ait communication, trois éléments sont indispensables .

- un émetteur, qui délivre un signal porteur de l'information
- un récepteur qui reçoit le signal et décode l'information que ce signal contient
- et une transmission du signal.

II-2 - La Théorie du signal

Une théorie scientifique doit atteindre la congruence entre les principes théoriques et les phénomènes observés, et permettre la possibilité de faire des prédictions sur certaines questions et rester valable face à de nouvelles expériences, ou elles doivent être corrigées.

La théorie du signal est un ensemble interconnecté d'idées, d'interprétations, d'observations et d'expériences qui s'appuient sur deux principes physiques et des hypothèses exprimées mathématiquement pour fournir une représentation du signal. Cette représentation pratique du signal permet de mettre en évidence ses principales caractéristiques, telles que la distribution fréquentielle et l'énergie.

La théorie du signal est la description mathématique des signaux et représente l'objectif principal de cette théorie. Elle repose sur un ensemble d'outils conceptuels qui permettent de construire des modèles théoriques basés sur les lois physiques exprimées en langage mathématique pour fournir une représentation du signal. Par conséquent, l'objectif fondamental de la théorie du signal est de décrire de façon mathématique ce qu'est un signal et ses propriétés. Il convient de distinguer deux branches principales du traitement du signal, qui sont la théorie du signal et la théorie de l'information. [31][32].

Traitement du signal

Théorie du signal	Théorie de l'information
Echantillonnage	Codage
Modulation	Cryptographie
Analyse spectrale	
Estimation	

Tableau (2-1) : présente les principales fonctions de Traitement du signal

II-3 - Traitement du signal

1- Historique

Les origines du traitement du signal reviennent au mathématicien et physicien français (1768 - 1830) par le développement d'une nouvelle méthode mathématique qui a remplacé la description d'un signal dans le temps ou dans l'espace par une description de ses fréquences composantes, appelée la "transformée de Fourier" représentée par la formule (2-1).

Le Traitement du Signal est une discipline indispensable. C'est un domaine dans le but d'extraire le maximum d'informations utiles d'un signal spécifique à l'aide d'une formulation mathématique adéquate et se concentre sur la physique appliquée perturbée par des signaux parasites (bruit), qui les masquent parfois complètement [32], [34].

Le Traitement du Signal est une discipline qui trouve son origine dans les sciences de l'ingénieur (particulièrement l'électronique, l'automatique), elle fait aujourd'hui largement, à de nombreux domaines, dans le but d'analyser le signal en fonction de leur nature (traitement de parole et traitement d'image et vidéo). Nous pouvons diviser le traitement du signal en deux catégories suivant les types d'opérations, et des principales fonctions, tel que suivant le tableau (2-1). Il y a deux types :

- type d'extraction des informations (Analyse, Détection, Estimation, Filtrage, Identification, Mesure, etc....).
- les types d'incorporations des informations (Codage, contrôle, Compression et la transmission, Modulation, Réduction du bruit, Synthèse, etc.....) [32], [34].

2 - La Transformation de Fourier

La transformation de Fourier $F(f)$ est une opération qui transforme une fonction intégrable sur \mathbb{R} en une autre fonction, décrivant le spectre fréquentiel de cette dernière. Si f est une fonction intégrable sur \mathbb{R} , sa transformée de Fourier est la fonction $F(f) = \hat{f}$ donnée par la formule (2-1) [33] :

$$(2-1) \quad F(f) = \int_{-\infty}^{+\infty} f(t) e^{-i\omega t} dx$$

Avec $\omega = 2\pi f$

3- Le champ d'application :

Pratiquement on peut dire que les techniques de traitement du signal et son champ d'application se situent donc dans tous les domaines concernés par la perception, la transmission ou l'exploitation des informations véhiculées par ces signaux et parmi ces domaines (Automatique , Astronomie , Automatique , Radar , Sonar , Télécommunications , Traitement d'image , ect) [32],[34].

4- Le bruit

Le bruit est un signal indésirable qui limite l'intelligibilité d'un signal utile et perturbe l'interprétation d'un signal. Le bruit est une perturbation superposée à un signal. La différenciation entre le signal et le bruit dépend de l'intérêt de l'utilisateur. Le rapport signal sur bruit, exprimé en décibels (dB), mesure la quantité de bruit contenue dans le signal et est défini par l'équation suivante [32] :

$$(2-2) \quad 10 \text{Log} \frac{P_S}{P_N} = \left(\frac{S}{N} \right)_{dB}$$

P_S : puissances du signal

P_N : bruit

II-5- Classifications des signaux

Un signal est une fonction d'une ou plusieurs variables. Ce signal physique est représenté par des fonctions à valeurs réelles $s(t)$ pour une variable réelle (t). Le signal a donc des caractéristiques telles que l'énergie bornée, une amplitude bornée, une continuité temporelle, une causalité ($s(t) = 0$ pour $t < 0$), un spectre du signal borné (tend vers 0 lorsque f tend vers ∞). Il existe d'autres critères de classifications des signaux en fonction de leur application tels que des signaux énergétique, morphologique, spectrale, ... etc :

1- Classification des signaux phénoménologiques

Ces signaux sont représentés par des fonctions $s(t)$ à une variable t (temps), Il présente deux types de principaux de Représentation temporelle (les signaux déterministes et les signaux aléatoires) :

1- a les signaux déterministes :

Un signal déterministe au contraire des signaux aléatoires/stochastiques, est un signal dont l'équation mathématique peut être calculée et représentée à tout moment et peut être prédite. Tout signal déterministe n'est pas obligatoirement périodique, On peut citer quelques fonctions mathématiques ainsi que leurs propriétés comme fonction échelon (unité), Fonction rectangle, Fonction signe, Fonction sinus cardinal, rampe, impulsion ou dirac, ..., donc Un signal déterministe peut être périodique ou non périodique.

Dans cette classe déterministe, il Ya les signaux périodiques, les signaux transitoires, etc... , on peut citer les signaux périodiques de période (T_p) si $f(t) = f(t + T_p)$ et (T_p) avec une fréquence fondamentale $F_p = 1/T_p$.

1- b Les signaux aléatoires :

Pour décrire le signal aléatoire (ou processus stochastique) il faut voir les propriétés statistiques, puisque leur comportement temporel est imprévisible, si ses valeurs ou réalisations dépendent du hasard et s'il ne possède pas de représentation analytique. on dit que le signal est stationnaire si leur propriétés statistiques invariantes dans le temps, donc leur valeur moyenne est indépendante du temps, et quel que soit le moment où l'on en observe une partie déterminée les résultats de leur analyse statistique restent les mêmes, l'observation de ce signal peut être caractérisée par des grandeurs statistiques ou fréquentielles. Un signal est dit ergodique si les moyennes sur plusieurs réalisations sont équivalentes à des moyennes temporelles correspondant à une seule épreuve, On peut dire que le signal est stationnaire et ergodique et peut être analysé en une seule expérimentation sur une longue période de temps [32].

2- Classification des signaux énergétiques :

La classification des signaux en tant que signaux énergétiques ou signaux de puissance est basée sur la quantité d'énergie contenue dans le signal (énergie est finie, borné dans le temps et dans l'amplitude, une durée finie, ..). Les signaux énergétiques et signaux de puissance est importante dans le traitement du signal.

3-Classification des signaux morphologiques :

La classification des signaux en tant que signaux morphologiques est basée sur les caractéristiques géométriques et temporelles du signal, sont généralement divisés en deux catégories : les signaux déterministes et les signaux aléatoires.

4- Classification des signaux spectrales :

La classification des signaux selon leur contenu spectral se fait en fonction de la répartition des fréquences dans le signal. Les signaux peuvent être classés en quatre catégories principales (Signaux à bande étroite , Signaux à bande large , Signaux à spectre continu , Signaux à spectre discret) .

II-6- Le signal parole

II-7- Définition de signal parole

La parole est la capacité du cerveau humain à donner des ordres et à contrôler les muscles des différents organes impliqués dans le processus d'articulation ou le système de la production de parole (l'appareil vocal humain) , Parmi ces organes , les cordes vocales, lèvres, la langue et le larynx , etc... .

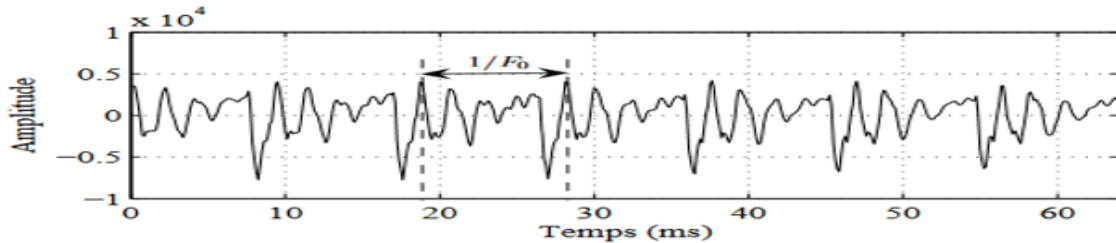
La parole est un moyen pour communiquer la pensée par l'intermédiaire de sons articulés , et c'est un aspect complexe et fascinant du comportement humain ,généralement il Ya plusieurs description de la parole telle que phonétique (phénomène de production de la parole) , syntaxique (étude descriptive des relations qui existent entre les mots) , l'acoustique [35],[36].

La perception de la parole implique le décodage et l'interprétation des signaux acoustiques pour en extraire le sens. Le système auditif humain est sensible aux propriétés spectrales et temporelles des sons de la parole, ce qui nous permet de distinguer différents phonèmes et mots.

Le signal de parole correspond aux changements de la pression de l'air induits par l'appareil vocal humain, considérés en fonction du temps. Ainsi, le signal de parole est un vecteur acoustique porteur d'informations d'une grande complexité. le signal de parole est un signal continu, non-stationnaire, d'énergie finie, dont les propriétés statistiques (moyenne, écart-type...) varient en fonction du temps. On peut distinguer deux grandes catégories de sons[35],[36]. :

1- a Les sons voisés

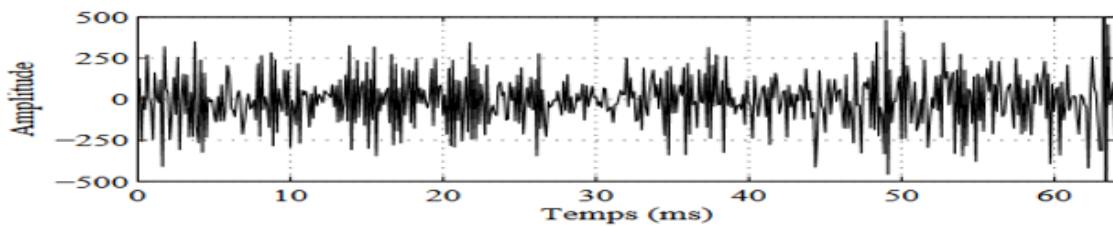
Par définition, sont des signaux quasi-périodiques qui possèdent un spectre fréquentiel très caractéristique. les sons voisés d'un parole, sont produits lorsque l'air traverse les cordes vocales et les fait vibrer. Il s'agit d'une suite d'impulsions périodiques d'air fournies par le larynx, caractérisée par une énergie élevée et une fréquence fondamentale (pitch). Généralement, les sons voisés ont une durée comprise entre 2ms et 20ms. [37]



Figure(2-1) : un signal d'un son voisé [37].

1- b- Les sons non voisés :

Sont des signaux apériodiques qui se produisent lorsqu'une perturbation du flux d'air provoque une contraction du conduit vocal. Ils sont équivalents à un bruit blanc filtré par le conduit vocal, et leur spectre n'a pas de structure particulière. Souvent, le spectre est le plus actif aux hautes fréquences.



figure(2-2) : un signal d'un son non voisés [37]

II-8- Les caractéristiques de signal parole

Les caractéristiques de ce signal sont appelées traits acoustiques. Chaque trait acoustique a une signification sur le plan perceptuel [38] La parole peut être caractérisée par plusieurs caractéristiques acoustiques, notamment :

2-a - Hauteur :

La hauteur ou la gravité perçue de la voix d'une personne, déterminée par la fréquence des vibrations de ses cordes vocales.

2-b - Intensité :

Le volume perçu de la voix d'une personne, déterminé par l'amplitude des ondes sonores qu'elle produit.

2- c- Durée :

La durée pendant laquelle un son ou un mot est produit, mesurée en secondes.

2-d - Formants :

Les fréquences de résonance du tractus vocal qui façonnent le spectre des sons de la parole et contribuent à leur caractère distinctif perceptif.

2-e - La fréquence fondamentale :

La périodicité d'un son voisé est déterminée par La fréquence fondamentale d'un signal de parole ou pitch , cette fréquence fait référence à la fréquence approximative de la structure quasi- périodique des signaux de parole voisés , est liée à la fréquence de vibration des cordes vocales (la fréquence du cycle d'ouverture/fermeture des cordes vocales), et par définition elle est l'inverse de la période de vibration des cordes vocales , Il est désigné par F_0 . Généralement la longueur des cordes vocales et des facteurs externes ,est liés à l'individu ont un impact significatif sur F_0 , cette fréquence fondamentale (F_0) , est évolue lentement dans le temps à l'intérieur des zones voisées .[39]

La fréquence fondamentale (F_0) , suivie par des raies (harmoniques) , multiples de F_0 de 3 à 5 formants ($F_1 - F_5$) à des fréquences plus élevées appelés formants , donc F_0 est visible comme une répétition après chaque échantillon . Dans le spectre de magnitude, de puissance ou de log-magnitude, le F_0 est visible sous la forme d'un pic à la fréquence $F_0=1/T_0$, où F_0 est la fréquence d'échantillonnage [40] , finalement en acoustique la fréquence fondamentale est définie au sens mathématique de Fourier[41] :

$$\forall x \in \mathbb{R} , T \in \mathbb{N}^+ \text{ si } f(x + T) = f(x)$$

T : periode et $f=1/T$ est le fondamental

A titre d'exemple si on à N pics p_k représente les fréquences de F_1 jusqu'à F_N , qui correspondent aux pics harmoniques kF_0 , alors vous pouvez approximativement calculer F_0 avec équation (2-3) suivant :

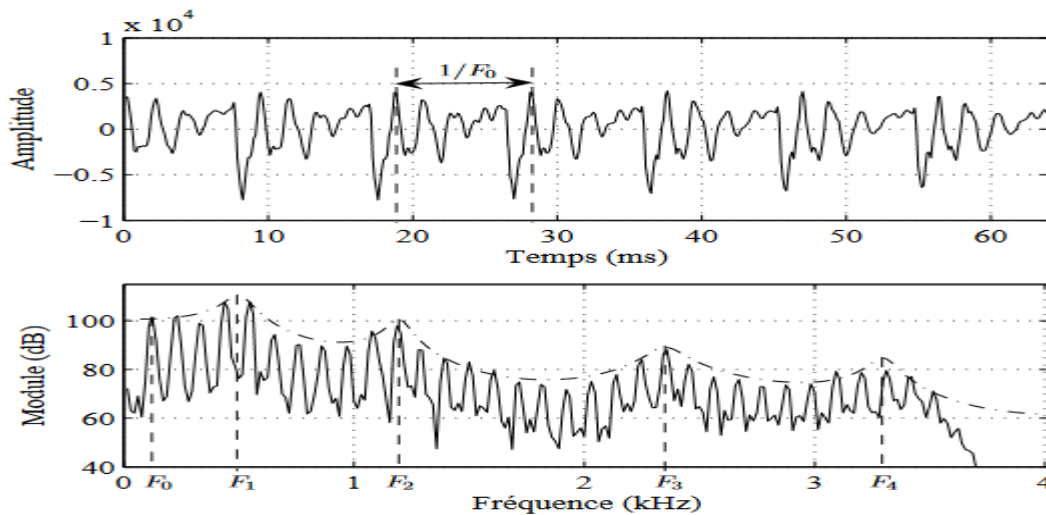
$$2-3 \quad F_0 \approx \frac{1}{N} \sum_{K=1}^N \frac{p_k}{K}$$

OU

$$F_0 \approx \frac{1}{N-1} \sum_{K=1}^{N-1} (p_{k+1} - p_k)$$

La fréquence fondamentale (F_0) , est se situent approximativement entre 80 et 600 Hz, où les hommes ont des voix plus basses que les femmes et les enfants . cette fréquence (F_0) approximativement peut varier [41]:

- de 80 à 200 Hz pour une voix masculine,
- de 150 à 450 Hz pour une voix féminine,
- de 200 à 600 Hz pour une voix d'enfant.



figure(2-3) : La fréquence fondamentale (tame d'un signal parole) , suivie par des raies (harmoniques f1-f5)[37]

2- f La spectre :

Les sons voisés du signal apparaissent sous la forme de successions de pics spectraux marqués, appelée enveloppe spectrale de type passe-bas, est elle-même présente les formants par kHz de bande passante, au contraire les sons non voisés qui se concentrent sur les hautes fréquences.

Le signal de parole est varié dans le temps, avec une excitation variant dans le temps, est de nature non stationnaire. La plupart des outils de traitement du signal étudiés sont conçus pour les signaux invariants dans le temps et une excitation invariante dans le temps, c'est-à-dire les signaux stationnaires. donc ces outils ne sont pas directement applicables au traitement de la parole, pour cela, nous devons découper tout le signal en trames, généralement la dure d'une trame est de 20 à 30 ms pour obtenir un temps stationnaire de courte durée d'une trame.

L'enveloppe spectrale ou spectre représente l'intensité de la voix selon la fréquence, elle est généralement obtenue par une analyse de Fourier à court terme, donc l'analyse et de synthèse de La transformée de Fourier à court terme (STFT) « Short Time Fourier Transform », sont fondamentaux pour décrire tout signal quasi-stationnaire (variant lentement dans le temps) tel que la parole. est appliquée dans l'analyse spectrale de la parole (signal échantillonné), en pondérant ces échantillons du signal (20 à 30 ms) par une fenêtre de pondération Hamming, et en effectuant un transformée de Fourier sur ces échantillons [42]:

L'avantage du cepstre, est son insensibilité au problème des formants (leur confusion avec les harmoniques dans le contexte de détection de fréquence glottale), et de la distorsion de la phase. Il permet par un simple filtrage passe bas (simple fenêtre rectangulaire) de récupérer la contribution vocale dans le domaine queférentiel. Dans le cas des voisés, le fondamental est correspond au pic le plus élevé. Par contre l'inconvénient de la méthode du centre est sa sensibilité au bruit, et à la taille de la fenêtre d'analyse. En plus la mesure du fondamental n'est pas toujours précise [41].

2- g - L'Energie :

La parole est produite à partir d'un système de voies vocales variant dans le temps avec une excitation . Par conséquent, le signal de parole est de nature non stationnaire. l'énergie est représentée par l'intensité du son , qui est liée à la pression de l'air en amont du larynx . Elle est plus forte pour les sons voisés , que pour les sons non voisés .

Le calcul de l'énergie totale du signal de parole n'est généralement pas inutile, mais son utilité dépend du contexte et des objectifs spécifiques de l'analyse du signal vocal. C'est parce que nous savons par la nature de sa production que la parole a une capacité ou une énergie qui varient dans le temps. Donc, ce qui est important dans le cas de la production de la parole, c'est un outil qui fournit des informations sur l'évolution de l'énergie au fil du temps. Ainsi, une méthode différente de traitement de la parole est nécessaire , et Pour être plus précis, les outils peuvent toujours supposer que le signal en cours de traitement est stationnaire.

Le signal vocal peut être stationnaire lorsqu'il est visualisé par une trame de 20 à 30 ms ,ce traitement est appelé traitement à court terme de la parole et peut être effectué dans le domaine temporel ou le domaine fréquentiel , et nous peuvent calculés à partir du traitement dans le domaine temporel de la parole L'énergie à court terme, le taux de passage par zéro à court terme , et l'autocorrélation à court terme [43] .

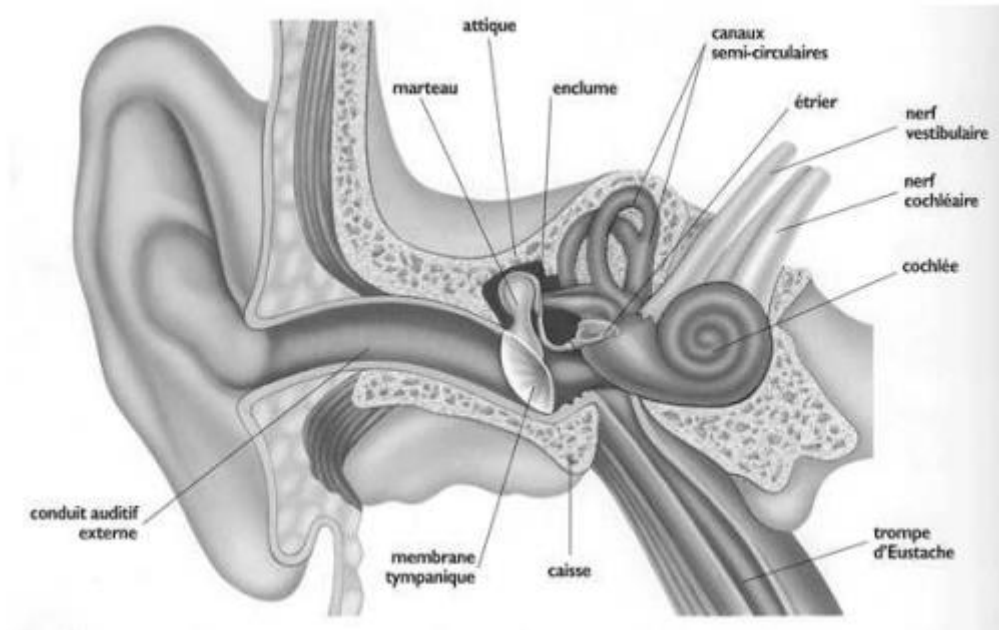
L'intensité du son se mesure en général en [dB] est une mesure logarithmique , et distingue un son fort d'un son faible. Par exemple une conversation normale a une intensité entre 20 et 50 dB, tandis qu'une discussion animée se mesure de 50 à 85 dB [43].

II-9- Perception auditive

L'onde acoustique est transmise de l'oreille externe à l'oreille interne où le tympan et les structures osseuses convertissent l'onde sonore en vibrations mécaniques qui sont transférées à la membrane basilaire à l'intérieur de la cochlée. La membrane basilaire vibre de manière sélective en fréquence le long de son étendue et effectue ainsi une analyse spectrale grossière (non uniforme) du son. Répartis le long de la membrane basilaire se trouvent un ensemble de cellules ciliées internes qui servent à convertir le mouvement le long de la membrane basilaire en activité neurale. Cela produit une représentation du nerf auditif en temps et en fréquence.

Donc on peut dire que L'organe majeur de l'audition est la membrane centrale de la cochlée appelée membrane basilaire, Elle renferme l'organe de Corti qui est constitué d'environ 25000 cellules ciliées raccordées au nerf auditif et baignant dans le liquide cochléaire , Chacune de ces cellules possède une réponse en fréquence qui dépend de sa position sur la membrane . On peut donc dire qu'une transformation fréquence-espace s'effectue le long de cette membrane. [44]

La figure (2-5) montre une vue de les trois sections de l'oreille humaine : l'oreille externe constituée du pavillon, et l'oreille moyenne commençant au niveau de la membrane tympanique, ou tympan, et comprenant trois petits osseuses (le marteau , l'enclume , l'étrier) , et l'oreille interne, constituée de la cochlée et de l'ensemble des connexions neuronales au nerf auditif .[45]

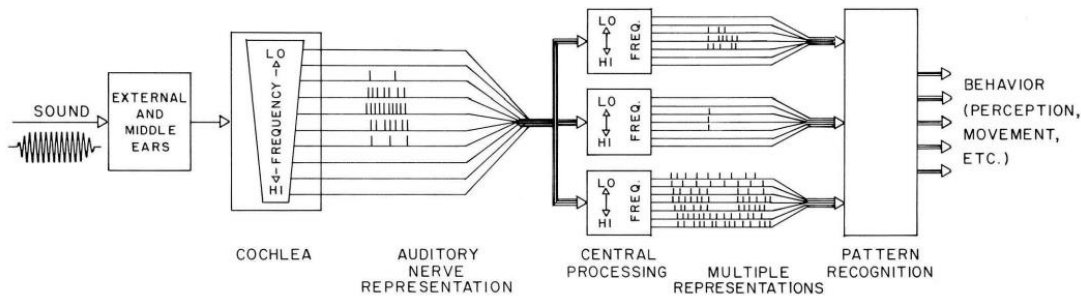


La figure (2-5) : présente l'appareil auditif humain (Structure de l'oreille).[45]

II-10- Mécanisme de l'audition

Le système auditif humain reste mal compris à ce jour, et seuls les mécanismes que le cerveau humain utilise pour percevoir le son ou la parole peuvent faire l'objet d'hypothèses. Cependant, plusieurs expériences des connaissances sur la façon dont les sons a été découverte qui utilisent des tonalités et des signaux sonores pour stimuler le système auditif humain de manière très spécifique et contrôlée, Ces expériences ont fourni des connaissances très précieuses sur la sensibilité du système auditif de l'être humain , comme les propriétés acoustiques telles que l'intensité et la fréquence .

Le traitement à des niveaux supérieurs dans le cerveau ,est présenter la figure (2-6) présenté schématiquement une séquence de traitement central avec de multiples représentations suivies d'un certain type de reconnaissance de formes [44]



M. B. Sachs et al.

Figure(2-6) : Schéma équivalente de système auditif.[46]

L'intensité du son est déterminée par l'intensité des ondes sonores. Les ondes sonores très intenses causent une plus grande vibration de la membrane basilaire, ce qui entraîne une augmentation de la fréquence des influx nerveux qui atteignent l'encéphale, Il est possible qu'un plus grand nombre des cellules ciliées soient également stimulées par des sons plus forts.

Georg Von Békésy (Prix Nobel de physiologie et de médecine en 1961), il démontra que lorsque la cochlée était stimulée par une onde sonore, la membrane basilaire était déplacée suivant une onde qui se propageait à partir de la base de la cochlée, L'amplitude de cette onde augmente au fur et à mesure qu'elle se propage, passe par un maximum et décroît ensuite très rapidement, La position du maximum d'amplitude de l'onde propagée dépend de la fréquence du son, pour les aigus il sera près de la base de la cochlée et pour des sons graves vers son apex. c'est -a- dire que La membrane basilaire est plus étroite mais plus rigide à la base de la cochlée, les sons de haute fréquence induisent des vibrations maximales dans cette région Vers l'apex de la cochlée, et la membrane basilaire est plus large mais plus flexible. les sons de basse fréquence entraînent une vibration maximale de la membrane basilaire dans cette région. La figure (2-7) présente une onde qui se propageait à partir de la base de la cochlée [47].

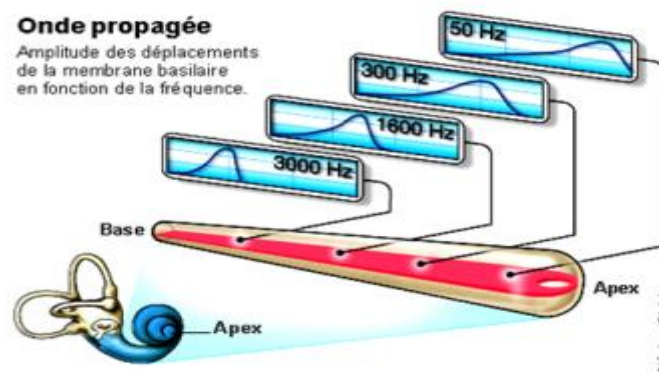
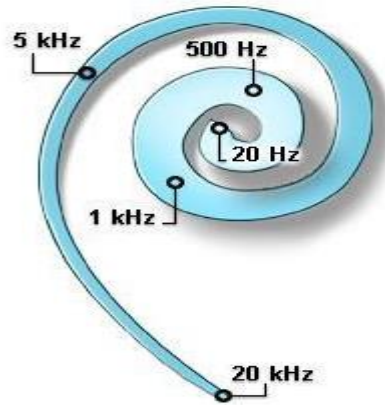


Figure (2-7) : Tonotopie passive d'après Békésy (prix Nobel en 1961) [47]

La figure (2-8) représente la distribution des fréquences le long de la membrane basilaire d'une cochlée humaine. En fonction de sa fréquence, donc le gradient d'élargissement de la membrane basilaire depuis la base (20 Hz) jusqu'à l'apex (20 kHz). La vibration a un effet maximal (résonance) en des points différents de la membrane basilaire, c'est la tonotopie passive, et permettent de remarquer que cette distribution est non-linéaire (la tendance est logarithmique).



Figure(2-8) : Distribution des fréquences le long de la membrane basilaire.[45]

L'un des facteurs clés de la perception des sons et de la parole est l'intensité sonore, elle est la qualité perceptuelle liée à la propriété physique du niveau de pression acoustique. L'intensité sonore est une relation entre le niveau de pression acoustique réel d'un son pur (en dB par rapport à un niveau de référence standard) et l'intensité sonore perçue du même son (dans une unité appelée phons) sur la plage d'audition humaine (20 Hz–20 kHz). La figure (2-9) montre que la perception de l'intensité sonore dépend de la fréquence, et indique le « seuil audible » du niveau de pression acoustique requis pour qu'un son à une fréquence donnée soit audible.

Des mesures minutieuses montrent qu'une tonalité de 50 dB 1000 Hz et la tonalité 60 dB 100 Hz sont censées avoir un niveau sonore de 50 phons.

Les courbes de niveau d'intensité sonore pour l'audition humaine de la figure (2-9) montrent que le système auditif est le plus sensible pour des fréquences allant d'environ 100 Hz à environ 6 kHz. C'est presque précisément la gamme des fréquences occupée par la plupart des sons de la parole. [44]

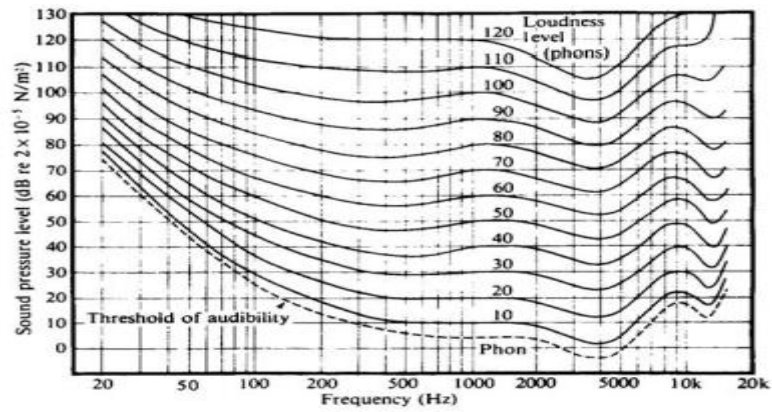
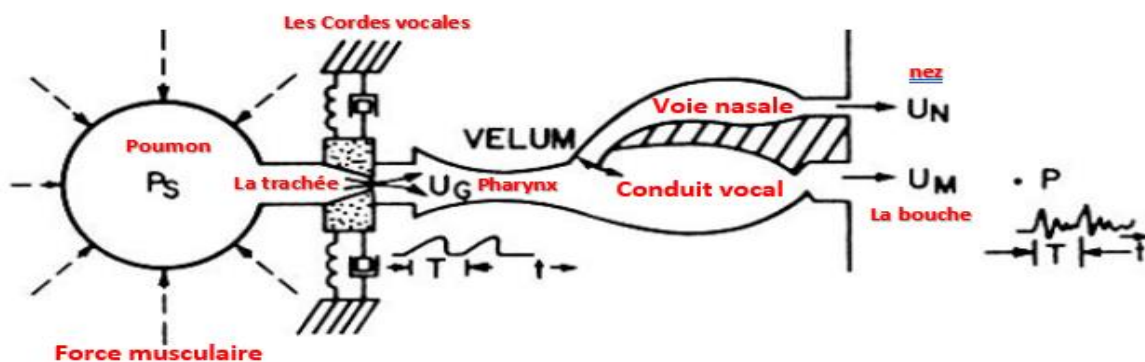


Figure (2-9) : Niveau d'intensité sonore pour l'audition humaine[48].

II-11- Physiologie du système de production de la parole

La figure (2-10) , présente Modèle schématique du mécanisme de l'appareil vocal humain , Il montre modèle approximatif raisonnable pour les longueurs d'onde des sons de la parole, et présente les caractéristiques physiques essentielles qui entrent dans les étapes du processus de production de la parole . et Cette figure montre le chemin de l'air sortant et fourni par les poumons , qui est acheminé par la trachée et passe via le larynx , cette larynx il à Le rôle de la production d'une excitation, qui à une forme quasi-périodique , et passe via le conduit vocal dans le cas des sons , qui est similaire comme un tube de section transversale non uniforme qui est délimité à une extrémité par les cordes vocales . c'est à-dire que Le passage d'une bouffée d'air via la glotte entraîne un mouvement vibratoire des cordes vocales et la création du son (signal acoustique). [44]



Figure(2.10) : Modèle Schéma de principe du système vocal humain . (After Flanagan et al.) [460].

II-12- La Production de la Parole

La structure physiologique du système de production de la parole se reflète dans les propriétés du signal de parole . la production d'un signal de parole $s(n)$ peut être considéré comme un résultat de l'excitation $u(n)$ du conduit vocal par un train d'impulsion ou un bruit blanc avec un gain G donnant lieu respectivement aux sons voisés et non voisés et par un mélange des deux pour les sons mixtes , Dans ces sons voisés, l'excitation est une vibration périodique des cordes vocales suite à la pression exercée par l'air provenant de l'appareil respiratoire .Le signal de parole n'est pas stationnaire , et Pour cette raison la modélisation de signal parole est difficile , et il faut travailler sur des tranches successives de temps de durée limitée de 30ms à 20ms décalées de 10ms avec recouvrement , c'est une procédure classique consiste à calculer un modèle .[50]

II-13- Modèle de production de la parole

Il existe deux approches de production de la parole on distingue :

1-Les approches articulatoires

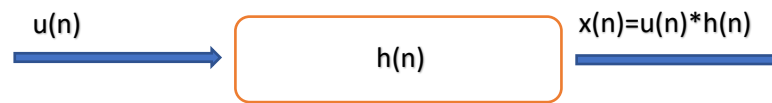
Ce type dépend de notre visualisation du conduit vocal liée aux paramètres articulés, dont la position et la forme de la langue, l'ouverture des lèvres, le larynx, la forme des conduits et des cordes vocales .En théorie, ces modèles nous permettent de générer un grand nombre de sons en modifiant simplement un ensemble des paramètres articulatoires.

2-Les approches électriques

Ce modèle proposé , utilise l'équivalents électriques du mécanisme de la production de la parole, on y décrit la parole comme le signal produit par un système constitué de générateurs et de filtres numériques, Dans cette catégorie on retrouve deux modèles : les modèles AR (Auto-Régressif) et les modèles ARMA (Auto-Régressif à Moyenne Ajustée) . [51]

II-14- Le Modèle auto régressif

Le terme de modélisation autorégressive vient du fait qu'un signal est dit autorégressif lorsqu'il est engendré par récurrence . le mécanisme de production de signal parole par le système Figure (2-11) peut être assimilée à l'excitation d'un filtre par une source pseudopériodique ou aléatoire . En d'autres part , nous pouvons considérer le signal vocal comme étant la convolution de la source $u(n)$ et de la réponse impulsionnelle du filtre représentant le conduit vocal $h(n)$, comme présenté sur la formule (2-5). on peut dire aussi , que Pour un signal voisé l'excitation est un train périodique d'impulsions d'amplitude , par contre pour les sons non voisés, l'excitation est un bruit blanc de moyenne nulle et de variance unité.



La figure (2-11) : Modèle simple de la production de la parole

$$2-4 \quad u(n) = \sum_k \delta(n - kp) \quad , \quad p : \text{periode fondamental}$$

$$2-5 \quad x(n) = u(n) \otimes h(n)$$

D'autre part en utilisant la transformée en Z de $x(n)$ on peut écrire par la formule suivante (2-6):

$$2-6 \quad X(z) = U(z)H(z)$$

La fonction $H(z)$ est la transfert du conduit vocal , est un filtre tous-pôles d'ordre p , et G représente un gain de la source d'excitation , cette équation est exprimée par la formule suivante (2-7) .

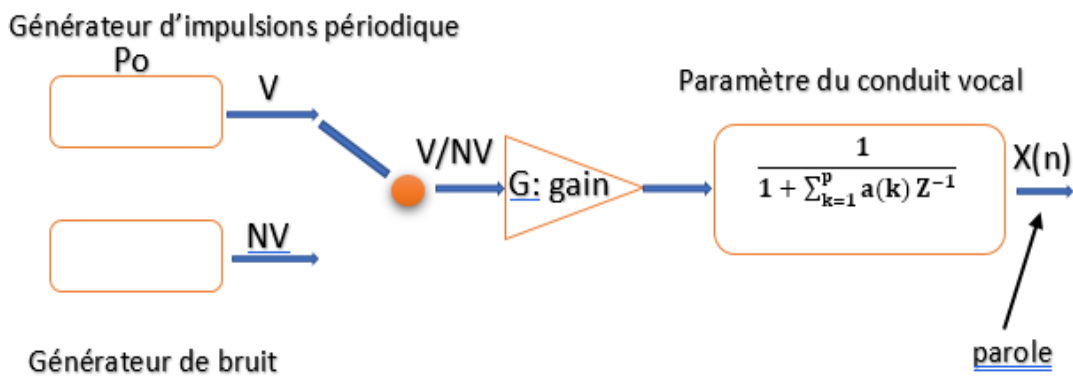
$$2-7 \quad H(z) = \frac{G}{1 - \sum_{k=1}^p a(k)z^{-k}}$$

Les Coefficients G et $a(k)$ dans (2-7) changent avec le temps, Le système linéaire dans le modèle est communément appelé simplement le système de "voie vocale" , et la réponse impulsionnelle correspondante est appelée la "réponse impulsionnelle de la voie vocale" , Pour les systèmes linéaires omnipolaires, représentés par (2-7), l'entrée et la sortie sont liées par une équation aux différences de la formule (2-9) , Puisque que chaque échantillon peut être déterminé à partir des échantillons précédents , Donc l'équation de modélisation de la production de parole est exprimée par la formule suivante (2-8) , et par la formule (2-9) dans le domaine temporel .

$$2-8 \quad X(z) = \frac{G U(z)}{1 + \sum_{k=1}^p a(k)z^{-k}}$$

$$2-9 \quad x(n) = \sum_{k=1}^p a(k) \cdot x(n-1) + G u(n)$$

En général l'excitation $u(n)$ est soit une suite d'impulsions périodiques, soit un bruit blanc, donc on peut prévoir l'échantillon $x(n)$ à partir d'une combinaison linéaire des p échantillons qui le précèdent. Les coefficients $a(k)$ du filtre représentent les caractéristiques du conduit vocal. Ces coefficients sont d'ailleurs appelés les coefficients de prédiction, car en cas d'excitation nulle, chaque échantillon $x(n)$ pourrait être prédit à partir des p échantillons précédents, associés aux coefficients de prédiction [52]. C'est-à-dire que l'un de ces modèles utilise une représentation plus détaillée de l'excitation en termes des générateurs de source séparés pour la parole voisée et non voisée, comme le montre la figure (2-12). Dans ce modèle, l'excitation non voisée est supposée être une séquence de bruit aléatoire, et l'excitation voisée est supposée être un train d'impulsions périodiques avec des impulsions espacées par la période de hauteur (p_0) arrondi à l'échantillon le plus proche. La figure (2-12) représente le modèle autorégressif de production de la parole (AR) d'ordre (p) relatif à la définition du signal autorégressif. On lui donne souvent le nom "tout pôle" puisque toutes les racines de la fonction de transfert sont des pôles.



La figure (2-12) : représente le modèle autorégressif de production de la parole (AR).

Il est utile de visualiser le spectre d'un signal synthétique obtenu au moyen du système décrit par la figure (2-12) du modèle autorégressif, et de mieux comprendre l'intérêt de ce modèle, il est en effet une approximation du spectre du signal, qui est généré à travers d'un conduit vocal dont on aurait déterminé les caractéristiques. Cette approximation étant d'autant meilleure que l'ordre du filtre de prédiction est élevé. Le problème de l'estimation d'un modèle AR, revient à déterminer les coefficients d'un filtre récursif, dont on connaît le signal de sortie mais pas l'entrée, et il faut adopter un critère pour choisir entre une infinité de solutions possibles. Le critère le plus courant est la minimisation de l'énergie de l'erreur de prédiction.

II-15- Modèle Auto-Régressif à Moyenne Ajustée (ARMA)

Le modèle ARMA est un modèle autorégressif stationnaire dans lequel les variables indépendantes suivent des tendances stochastiques et le terme d'erreur est stationnaire. En d'autres termes, le modèle ARMA intègre l'autocorrélation et le modèle de moyenne mobile dans sa régression. Le modèle est connu sous le nom autorégressif à moyenne ajustée "ARMA" ou aussi modèle de Box-Jenkins qui parfois utilisé, ils sont les principaux des séries temporelles. Le modèle ARMA est un outil pour comprendre et prédire, composé de deux parties : une part autorégressive (AR) et une part moyenne-mobile (MA). Le modèle est généralement noté ARMA(p,q), où p est l'ordre de la partie AR et q l'ordre de la partie MA. Il se distingue par la présence à la fois des pôles et des zéros. [53].

II-16- La biométrie

1- Histoire de la biométrie

La biométrie n'est pas une technologie moderne, mais elle remonte au début du XIV^e siècle lorsque les Chinois utilisaient les empreintes plantaires et elle s'est beaucoup développée en tant que domaine de recherche. Il y a eu tellement de développements au cours des XIX^e et XX^e siècles qu'il serait fou de tous les énumérer.

En 1788, J.C.A Meyer, l'anatomiste allemand, écrivait que la disposition des bords de la peau ne se répète jamais chez deux personnes, et a été le premier qui déclarer cette théorie. [71]

En 1882 le système appelé bertillonage est la premier système biométrique, d'après son propriétaire, Alphonse bertillon, ce système est basé sur la Science qui étudie les caractères anatomiques et biologiques de l'espèce humaine comme la mesure des longueurs de la main ou la distance entre les yeux, la taille, la longueur d'un pied, le tableau (2-2) présente un modèle de méthode des Mesures bertillonages (détail de fiche) en 1894. Cela n'a pas duré longtemps car il manquait la précision des données, ainsi que le temps nécessaire pour des mesures complètes. La méthode Bertillon, ou «bertillonage» est un ensemble de procédures qu'il a lui-même codifiées, Elle se base sur quatorze mensurations. Cela n'a pas duré longtemps car il manquait la précision des données, ainsi que le temps nécessaire pour des mesures complètes, une année plus tard en 1823 le premier à classer les empreintes digitales en 9 modèles et ne les a pas utilisées pour l'identification, le physiologiste tchécoslovaque, Johannes Evangelista Purkinje (1787 - 1869). [71]

Haute	163.5	Oreille dr. Tête.	Long: 22.7	Pied g. 27.8	Coudée de l'iris g.	2	Agé de 28
Voûte			Larg: 19.2	Médius g. 9.2		Aur. M	né le 2 Mars 1859
Enverg	73.2		Long: 8.3	Auric. g. 10.3		Père Mard	à Paris
Buste	63.4		Larg: 5.2	Coudée g. 43.7		Part.	dép.

le tableau (2-2) :présente un modèle de méthode de Mesures bertillonnage (détail de fiche) en 1894.[54]

En 1900, le système de Henry a été officiellement adopté en Grande-Bretagne, qui utilise les empreintes digitales pour identifier les criminels, car jusqu'à récemment, il était largement utilisé et l'est encore aujourd'hui dans une moindre mesure.

En 1969, les empreintes digitales et la reconnaissance faciale étaient largement utilisées pour identifier les personnes. Dans les années 1980, l'Institut national des normes et de la technologie a développé un groupe de reconnaissance vocale pour étudier et faire progresser les opérations de la technologie de reconnaissance vocale. Ces études sont à la base des commandes vocales et des systèmes de reconnaissance que nous utilisons aujourd'hui. En 1994, le premier algorithme de reconnaissance de l'iris a été breveté. En 1991, le développement de la technologie de détection faciale a rendu possible la reconnaissance faciale. Malgré les erreurs qu'elle a entachées à l'époque, cette technologie a accéléré son développement. Et à partir de l'an 2000, des nombreux algorithmes de reconnaissance et d'identification ont obtenu un brevet et ont été largement utilisés et commercialisés.[78]

En 2008 Les États-Unis ont commencé à utiliser des algorithmes pour mesurer la qualité de l'image et la qualité de l'iris et du doigt sur la base des données vitale. En 2010, le National Security Service américain a utilisé la biométrie pour identifier les terroristes, notamment dans les événements de Septembre 2011.

Les fabricants de téléphones portables ont équipé leurs smartphones de scanners d'empreintes digitales au grand public Touche ID, une fonction de reconnaissance d'empreintes digitales conçue et publiée par Apple Inc.

2- Définition biométrie

Le terme « biométrie » vient de deux mots grec bio (vie) et métrie (mesure) . La biométrie consiste à vérifier ou déterminer l'identité d'un individu . Autrement La biométrie est une reconnaissance automatique et facile des personnes basées sur la mesure des plusieurs caractéristiques physiques ou comportementales d'un individus , Autrement dit , ce sont des traits distinctifs utilisées pour l'identifier ou pour vérifier l'identité un individu , Chaque traits biométriques personnels doit varier d'une personne à l'autre , ces caractéristiques ne variant pas avec le temps , enregistrables et mesurables , durable et robuste . [56],[77],[78]

II-17- les Modalités biométriques les plus importantes

On peut aussi définir la biométrie comme suit : c'est le domaine technologique de la vérification et de l'identification pour déterminer l'identité d'un individu, est basé sur les caractéristiques physiologiques ou comportementales mesurables d'un individu. Il existe plusieurs appondues et modalités biométriques utilisées dans divers secteurs tels que biologiques (comme l'ADN, Odeur, sang, urine,), comportementales (comme la voix et la signature et le modèle de frappe au clavier,), ou morphologiques (comme l'empreinte digitale, la géométrie de la main, la reconnaissance du visage, l'iris,). Ces caractéristiques sont stables, ne varient pas avec le temps et l'âge de l'un individu. [79], [80]

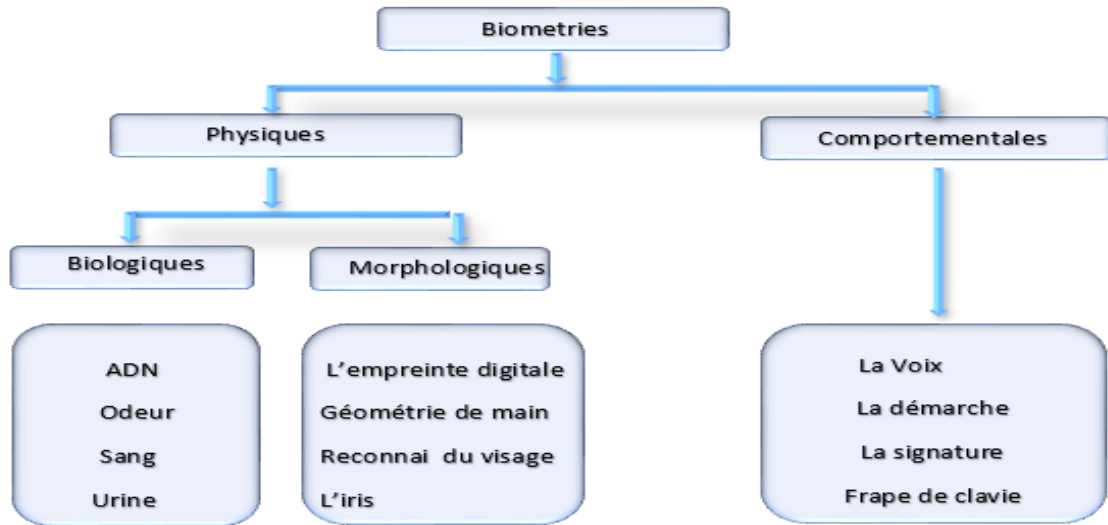


Figure (2-13-a): modalités biométriques



Figure (2-13-b) : Certain modalités biométriques (Biologiques , Morphologique , Comportementales) [57]

Le tableau suivant fournit une comparaison des modalités biométriques les plus répandus , citées leurs propriétés avec des degrés mais est plus ou moins adaptée à des applications .[81]

Biométries	Univ	Uniq	Perm	Enreg	Accep	Perfor
voix	moyen	faible	moyen	moyen	haute	faible
Démarche	moyen	faible	faible	haute	haute	haute
Signature	faible	faible	faible	haute	haute	faible
Frappe clv	faible	faible	faible	moyen	moyen	faible
ADN	haute	haute	haute	faible	faible	haute
IRIS	haute	haute	haute	moyen	faible	haute
Visage	haute	faible	moyen	haute	haute	faible
Géométrie de la main	moyen	moyen	moyen	haute	moyen	moyen
Empreinte digitale	moyen	haute	haute	moyen	moyen	moyen

Univ :(universelles) , Uniq : (niques) , Perm : (permante) , Enreg : (enregistrable) , Accep : (acceptabilité) , Perfor : (performance) .

Tablau (2-3) : Comparaison modalités biométriques

II-18- Fonctionnement d'un système biométrique

Un système biométrique typique peut être représenté par quatre partie principaux :

La première partie est la partie responsable de la capture des informations ou des données biométriques de l'individu .La deuxième partie , est la partie chargée d'extraction les caractéristiques des données obtenues à partir de capture , suivant des normes différentes selon chaque type de système de mesure des données biométriques et selon chaque individu. La troisième partie est l'unité chargée de la comparaison entre les caractéristiques extraites de la deuxième étape , et la base des données du système biométrique . cette unité détermine le degré de similitude (ou de contradiction) entre les deux selon chaque type du système de mesure des données biométriques et en fonction de chaque individu .La quatrième partie est la partie de décision vérifie l'identité , pour détermine l'identité d'une personne basée sur le degré de similitude entre les caractéristiques extraites et le(s) modèle(s) stocké(s) [57] .

II-19- Evaluation des Systèmes biométriques

Le type d'application est le seul moyen qui détermine les performances d'un système biométrique, car chaque modalité biométrique a ses propres forces et faiblesses, en d'autres termes, aucun système biométrique n'est « optimal ». De nombreuses études ont été menées pour évaluer les performances des systèmes biométriques. (IBG) La société américaine - International Biometrics Group - a noté sur la base de quatre critères d'évaluation :

1-l'existence d'un contact direct entre le capteur utilisé et l'individu à reconnaître.

- 2- la qualité d'environnement dans lequel l'utilisateur se trouve (fiabilité), Ce critère influe sur la reconnaissance de l'utilisateur par le système.
- 3-le coût de système biométrique. ne doit pas être relativement coûteuse pour établir une base de données.
- 4-Facilité lors de la saisie des données biométriques des utilisateurs, qui doivent être réduites au minimum.

	la voix	le visage	l'empreinte digitale	l'iris	la signature	ré tine	la frappe sur le clavier	la main
contact direct	Faible	Faible	Faible	Faible	Faible	Faible	Faible	Faible
fiabilité	haut	haut	haut	haut	Faible	haut	Faible	Faible
coût	Le moins coût	plus coûteuses	Le moins coût	plus coûteuses	Le moins coût	plus coût	Le moins coût	plus coût
Facilité	Le facile	Le facile	difficile	Le facile	facile	difficile	Le facile	difficile

Le tableau (2-4) : présente les performances de certaines modalité biométriques

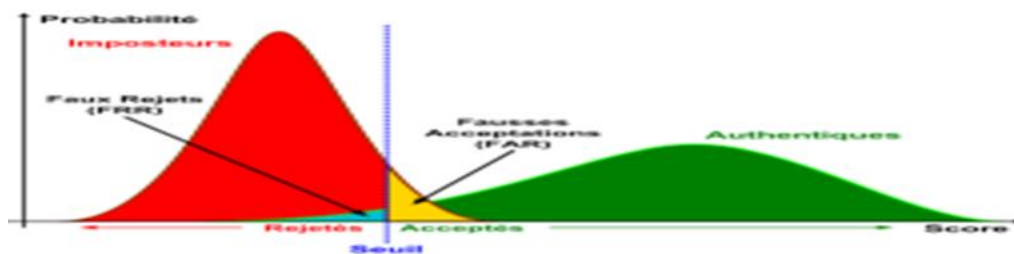
II-20- Evaluation des performances des Systèmes biométriques

L'évaluation des performances du système biométrique dépend de la prise d'autres critères , dont l'exactitude et la taille des données stockées, et selon le système à évaluer, qu'il s'agisse d'un système d'identification ou de vérification d'identité après la comparaison de deux signatures biométriques , et jugé performant s'il présente une efficacité et une robustesse en termes de précision et de vitesse et la précision de reconnaissance des individus,et afin de comprendre comment déterminer la performance d'un système biométrique, il nous faut définir clairement trois critères principaux [55],[56].

- **FRR** (False Rejection Rate) : Dénommé Taux de Faux Rejets , Correspond au pourcentage de personnes légitimes, rejetées par erreur . est le pourcentage des données de test authentiques qui ont été rejetées , c'est-à-dire le pourcentage de personnes censées être reconnues mais qui sont rejetées par le système [77] .
- **FAR** (False Acceptance Rate) : Taux de Fausses Acceptations , Correspond au Pourcentage d'imposteurs acceptés en tant que personnes légitimes. est le pourcentage des comparaisons imposteurs qui ont été acceptées, c'est -à-dire représente le pourcentage de personnes censées ne pas être reconnues , mais sont acceptées par le système . [ISO 2006]. [62]
- **EER** (Equal Error Rate) : Taux d'Egale Erreur , correspond au cas particulier où les deux courbes des évolutions du FAR et FRR se croisent. Il correspond alors à un seuil qui est indépendant de l'utilisation du système, est calculé par les deux critères FRR et FAR sont égaux, c'est-à-dire le meilleur compromis entre les faux rejets et les faux acceptations [58]

Selon le domaine d'utilisation du système biométrique, les valeurs des erreurs sont directement liées au choix du seuil d'acceptation et de rejet, et il faut noter que plus FAR est faible et plus FRR est élevé et vice versa. Le nombre de faux rejets et celui de fausses acceptations sont inversement proportionnels.

Dans le cas d'applications avec un niveau de sécurité très élevé, la valeur de FAR est faible et aucune fausse acceptation n'est tolérée. Par contre pour des applications moins sécurisées, la valeur de FAR est élevée et un seuil d'acceptation beaucoup plus faible, c'est-à-dire le système acceptera des utilisateurs clients et des imposteurs si le seuil de décision est faible, et rejettera des imposteurs et des utilisateurs clients si le seuil de décision est élevé, et malheureusement il est impossible de varier le seuil de décision de faire diminuer les deux types d'erreurs en même temps, la (Figure 1.4) représentée La relation entre FAR et FRR [59]. La figure (2-14) illustre le FRR et le FAR à partir de distributions des scores authentiques et imposteurs tandis que l'EER est représenté sur la figure (2-15).



la figure (2-14)

Des appareils biométriques prennent le seuil de sécurité désiré comme paramètre du processus de décision. et d'autres prise par l'application elle-même sur la base d'un score (borné), En général, si le score est plus grand ou égal au seuil, l'utilisateur va être accepté et, si le score est plus petit, il sera rejeté [10]. habituellement (FAR) et (FRR) est exprimé en un pourcentage par rapport au nombre total de tentatives d'accès autorisés, Selon la nature (authentification ou identification) du système biométrique, il existe deux façons de mesurer la performance en termes de Taux Faux d'acceptation (FAR), et le taux de faux rejet (FRR), défini comme suit [59]:

Taux de faux rejet :

$$FRR = \frac{\text{nombre faux rejets}}{\text{nombre de tests total}}$$

Taux de fausses acceptation :

$$FAR = \frac{\text{nombre de fausses acceptations}}{\text{nombre de tests total}}$$

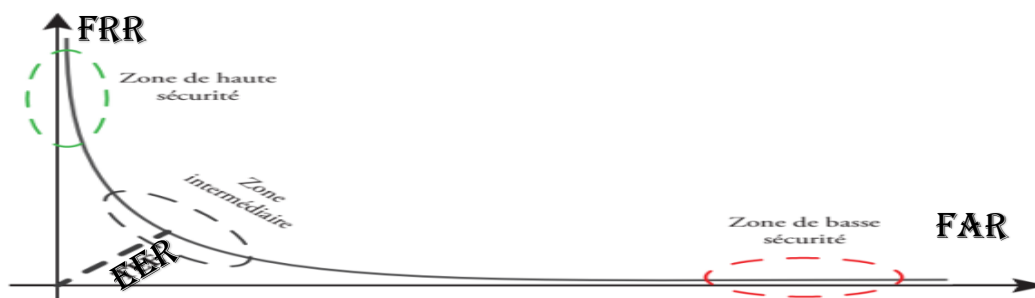
II-21- La précision d'un système biométriques

La précision d'un système biométriques est évaluée par le taux d'égalité d'erreurs (EER) qui est obtenu graphiquement, en s'appuyant sur la courbe ROC (Receiver Operating Characteristic).

Il existe deux façons de mesurer la performance d'un système biométrique, selon la nature (authentification ou identification). on utilise appelle une courbe ROC (Receiver Operating Characteristic).ou bien une courbe CMC (Cumulative Match Characteristic).

On utilise une courbe ROC ("Receiver Operating Characteristic" en anglais) , si le système en mode authentification , ce courbe trace (FRR) le taux de faux rejet en fonction du (FAR) le taux de fausse acceptation [Perr 02]. Plus cette courbe tend à épouser la forme du repère, plus le système est performant, c'est-à-dire possédant un taux de reconnaissance global élevé, la figure (2-15) présente La courbe ROC.

En mode identification, on utilise une courbe CMC ("Cumulative Match Characteristic" en anglais) , est que autre manière d'afficher la performance d'un système biométrique et peut également être calculée à partir du FAR et du FRR [57].



Figure(2-15) : Courbe ROC

II-21- Modalité empreinte digitale

L'empreintes digitales est la plus ancienne utilisée , Généralement, les systèmes d'empreintes digitales capturent des images d'empreintes digitales, extraient des caractéristiques des images , La force de cette méthode réside dans le fait que le dessin formé par les empreintes digitales est unique à chaque personne, telles que les bifurcations de crêtes, lac, les terminaisons, le centre delta ou impasse.... etc , Ces éléments sont appelés minuties (caractéristiques distinctes) , Statistiquement il est impossible de trouver douze points identiques chez deux personne . Cette modalité d'identification utilisée dans de nombreux domaines à ce jour, par exemple dans les aéroports, les passeports et les documents biométriques en général, la figure (2-16) présente l'empreintes digitales. Sachant qu'en France, la methode douze points similaires valident une identification , et parcontre Il en faut huit aux Etats-Unis [61].

Plus 95% de la population mondiale partage trois formes d'empreintes digitales, 60 % personnes sur un total de 95 % ont une forme d'empreintes digitales en boucle , et 30% des formes sont appelées verticilles en tourbillons , tandis que la forme d'arc rare des empreintes digitales

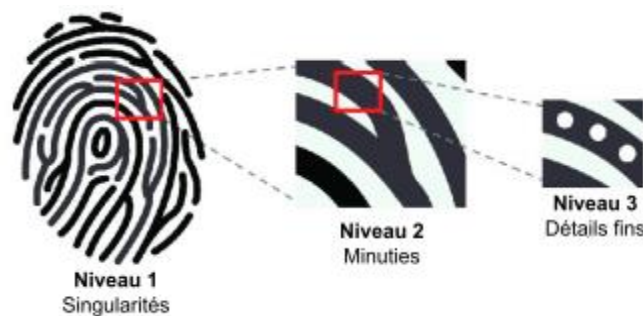
représente 5% des totaux , le reste de la population appartient à une catégorie plus complexe d'agencements avec de multiples boucles [62] .

Le système américain AFIS C'est un base de données utilisé pour identifier les empreintes digitales et le pourcentage de congruence entre deux cas par des algorithmes .AFIS C'est un système automatique d'identification des empreintes digitales, datant de 1999.



Figure (2-16) : représente une extraction des informations fournies par les empreintes digitales. [62]

La figure (2-17) représente une méthode d'extraction des informations fournies par trois niveaux de précision que nous obtenons à partir d'échantillons d'empreintes digitales par lesquels une personne peut être identifiée.



Figure(2-17) : Les niveaux de précision fournis par les empreintes digitales [64]

Le doigt d'un imposteur (Un doigt non enregistré) : Une attaque contre les systèmes d'authentications par un imposteur avec ses propres données biométriques est appelée falsification sans effort. Généralement, la précision de l'authentification des systèmes d'empreintes digitales est évaluée par le taux de faux rejets (FRR) et le taux de fausses acceptations (FAR). Le FAR est un indicateur important pour la sécurité contre une telle méthode comme avec un doigt non enregistré .

II-23- Modalité visage

1- La fonctionnement de la reconnaissance faciale

Un système de reconnaissance faciale est une application logicielle de reconnaître une personne grâce à son visage de manière automatique Basé sur 4 étapes . Ce système consiste à reconnaître l'utilisateur à partir d'un visage détecte et localise peut être capturée par une caméra ,la détection de ce visage peut être seul ou dans une foule avec modèle 2D ou modèle 3D . La plupart des technologies de reconnaissance faciale utilisent la 2D plutôt que la 3D, car il est plus pratique de comparer une image en 2D à des photos ou aux images d'une base de données dans un but de sécurité ou enregistrée . [67]

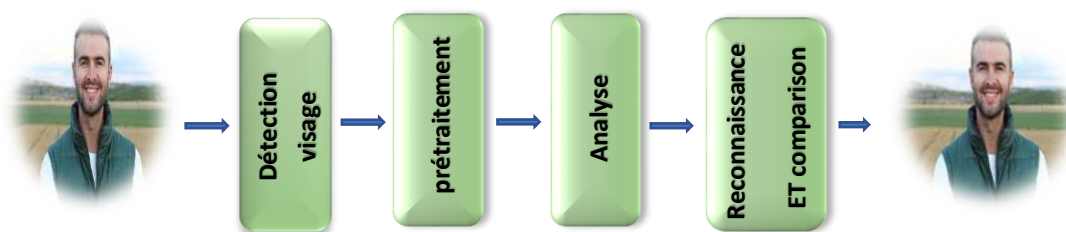


Figure (2-18) : chaîne de système de reconnaissance de visage

Il existe trois algorithmes sont dominants pour la reconnaissance faciale : Il s'agit de l'algorithme d'Analyse en Composantes Principales (PCA), la deuxième algorithme de Discrimination Linéaire (LDA) et la troisième algorithme (Elastic Bunch Graph Matching (EBGM)). Ces algorithmes se répartissent en deux grandes catégories, la première est basée sur les traits du visage et est géométrique, et la seconde photométrique , qui est basé sur la vue. [68]

2- L'histoire du système de reconnaissance faciale

La construction du premier système de reconnaissance faciale est due à Kelly et Kanade . les années 80 du siècle dernier ont connu de nombreux systèmes , le premier système réussi dans la technologie de reconnaissance faciale ,qui a été développé par Sirovich et Kirby en 1987 est basée sur l'apparence , utilisant l'analyse des composants principaux, et a été mis en œuvre par Matthew Turk et Alex Pentland (du Massachusetts Institute of Technology) dans l'identification et La vision artificielle pour résoudre le problème de reconnaître les visages humain. En 2011 C'est l'ordinateur qui sélectionne les points à comparer grâce au deep learning , est une méthode d'apprentissage automatique s'appuyant sur des réseaux de neurones artificiels , Plus on prend de photos, meilleur est l'apprentissage . La technologie de reconnaissance faciale s'est développée en 1993 jusqu'à ce que le taux d'erreur atteigne 0,003% en 2011, selon l'American Standards and Technologies Agency (NIST) .[69]

La plupart de ces systèmes utilisent 80 points nodaux sur le visage humain, qui sont des points nodaux utilisés pour mesurer les variables faciales humaines, telles que la longueur ou la largeur du nez, la profondeur des orbites et la forme du visage.

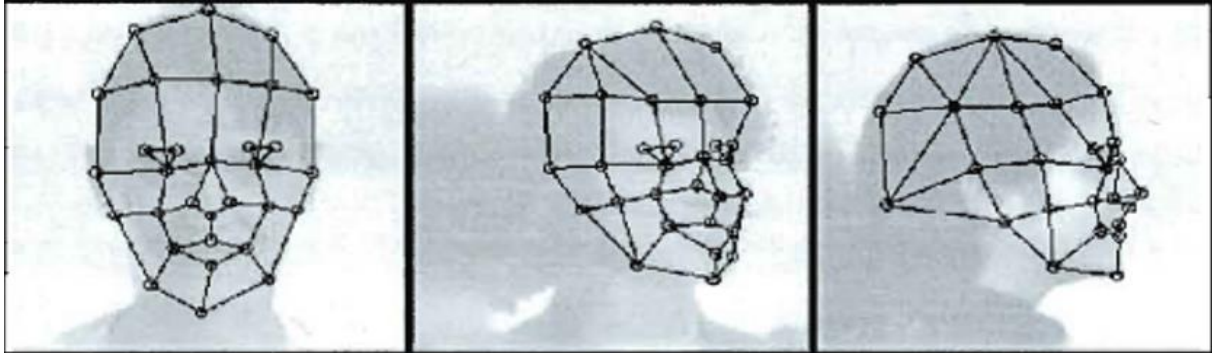


Figure (2-19) : nodaux sur un visage humain[68]

II-24- Modalité de reconnaissance par iris

1- Historique iris

La première idée de l'utilisation de l'iris comme moyen d'authentification et de la reconnaissance d'une personne était proposée et remarquée en 1936 par l'américain l'ophtalmologue Frank Burch .

En 1989, le professeur John Daugman (Né le 17 février 1954) de l'Université d'Harvard a présenté des algorithmes d'identification de l'iris, et ces algorithmes ont été brevetés en 1994 (Brevet n°WO09409446) , et constituent la base de tous les systèmes d'identification de l'iris , et ces algorithmes biométriques pour identifier les personnes de manière fiable et rapide en utilisant la texture aléatoire visible dans l'iris d'un œil.

Aujourd'hui ,cette utilisation ou système biométrique de reconnaissance d'iris reste liée aux aéroport notamment les zones de haute sécurité et dans tous les domaines et Chaque jour, ses algorithmes effectuent des centaines de milliards de comparaisons croisées d'iris en raison de la nécessité de rechercher des identités en double lors de la création de bases des données . [79][80]

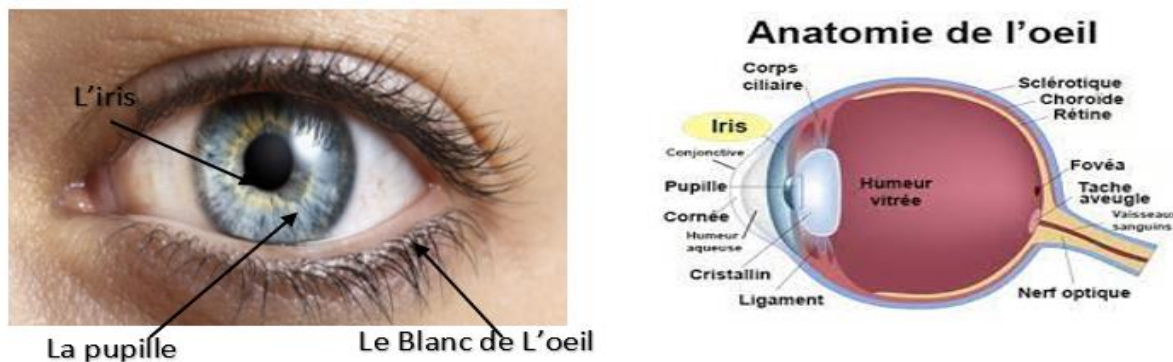
2- Définition

Iris, mot qui vient du grec iris, désigne la partie colorée de l'œil humain .L'iris humain est un attribut biométrique particulièrement discriminant entre individus . La reconnaissance de l'iris est considérée comme le système d'identification biométrique la plus fiable et le plus efficace est considérée comme étant parmi les meilleures technologies biométriques est l'un des moyens les plus performants pour identifier une personne, est une technique de biométrie permettant De reconnaître Une personne par l'observation de son iris. des études biologiques ont montré que les profils et les courbes présents dans un iris garantissent son

unicité et sont propres à chaque individu et stables dans le temps . Par ailleurs la reconnaissance par l'iris est considérée comme impossible à frauder, La probabilité de trouver deux iris identiques est 1 sur 10 puissances 72 , est très faible taux d'erreur selon Daugmann . La reconnaissance de l'iris est le seul système pouvant être utilisé sur un grand nombre de personnes en identification complète [81].

Un système biométrique permet l'identification automatique d'un individu repose sur ses caractéristiques unique et aux propriétés de l'iris. La plupart des systèmes commerciaux de reconnaissance par l'iris utilisent l'algorithme breveté et développé par Daugman. Cet algorithme est capable de produire des taux de reconnaissance parfaits .

L'identification de personnes par l'iris pose plusieurs problèmes parmi eux, De nombreux chercheurs se sont intéressés à cette problématique , Les méthodes proposées se distinguent les unes des autres par la phase de segmentation et la phase d'analyse et de caractérisation. [81].



La Figure (2-20) : présente La zone de l'iris dans l'œil [82]

Les algorithmes de codage et de reconnaissance pour les modèles d'iris ont été les programmes exécutables utilisés dans tous les systèmes de reconnaissance d'iris déployés commercialement ou dans les tests à ce jour , et Malgré sa petite taille (11 mm),l'iris a le grand avantage mathématique que la variation de ses motifs entre différentes personnes est énorme .

3- Architecture générale de système reconnaissance d'iris

Parmi les meilleures techniques biométriques pour identifier les personnes, il y a la technique d'identification d'une personne à travers une image de l'iris de l'œil, et cela est dû aux caractéristiques et caractéristiques de l'iris de l'œil. La plupart des systèmes commerciaux de reconnaissance par l'iris utilisent l'algorithme breveté et développé par Daugman . Cet algorithme est capable de produire des taux de reconnaissance parfaits.

Un système de vérification par reconnaissance d'iris peut se décomposer d'un certain nombre de sous-système suivant la même méthode et la même étape de reconnaissance de l'iris , généralement est constitué sur la principe :

1-a - L'acquisition :

La capture d'image

1-b - La segmentation :

Un système de segmentation automatique, qui est basé sur la transformée de Hough, qui représente un prétraitement des données ; il est capable de localiser la région de pupille, de l'iris et des paupières, et d'éliminer les informations inutiles (paupières, cils, et la sclérotique, etc) en ne conservant que la région de l'iris. Un seuillage a également été employé pour isoler les cils et les reflets. [81]

1- c - La normalisation :

La normalisation est utilisée pour la conversion de l'iris circulaire en forme rectangulaire avec des dimensions constantes pour tenir compte des incohérences d'images sous forme de dimensions fixes en utilisant le modèle pseudo- polaire de Daugman. [83].

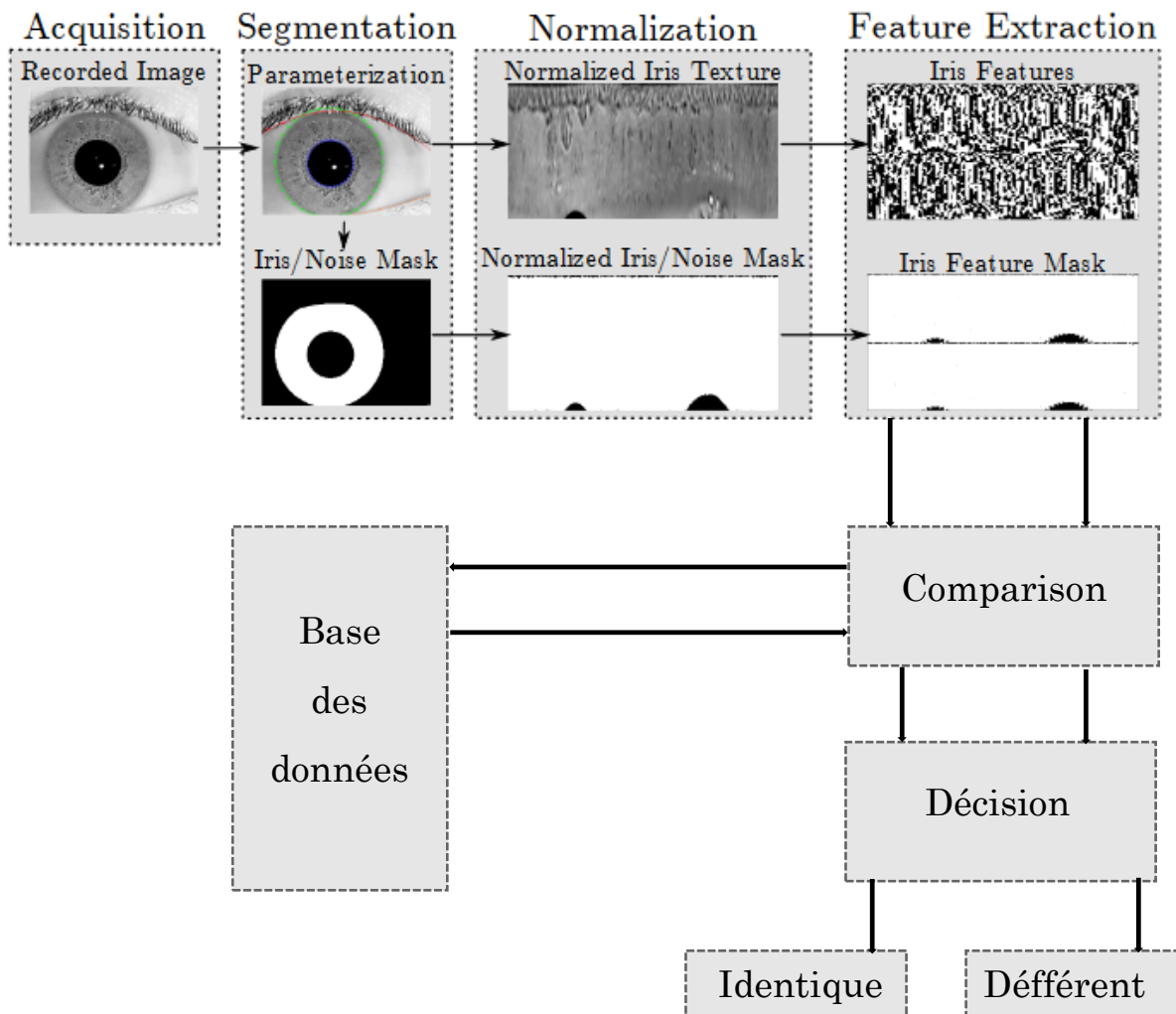


Figure (2-21) : presente Architecture générale de système reconnaissance d'iris

II-25- Modalité ADN

1- Définition

L'acide désoxyribonucléique (ADN) est une molécule biologique formant les chromosomes , et composée d'une double hélice à 2 chaînes spiralées . La double hélice d'ADN est recombinaée en deux copies identiques lorsqu'elle est divisé , et est le processus par lequel les traits génétiques sont transmis des parents à la progéniture. ADN est le support l'information génétique l'ensemble des caractères héréditaires d'une cellule et de l'information génétique , et le lieu où cette information génétique d'un individu est stockée , parmi les systèmes d'identification biométrique, l'ADN est intrinsèquement numérique ,et la plus fiable et immuable pendant la vie d'un être humain et même après sa mort. [84]

2- Historique

L'image connue sous le nom d'Image 51 est l'image la plus connue prise d'une molécule d'ADN par cristallographie par diffraction des rayons X. est une image prise par l'étudiant Raymond Gosling (né le 15 juillet 1926 à Wembley et mort le 18 mai 2015), avec Maurice Wilkins , et Rosalind Franklin sur la découverte et la caractérisation de la structure de l'ADN, en mai 1952 ,au King's College de Londres , c'ette image a été étiquetée "photo 51" parce que c'était la 51e photographie de diffraction que Franklin et Gosling avaient prise [85].

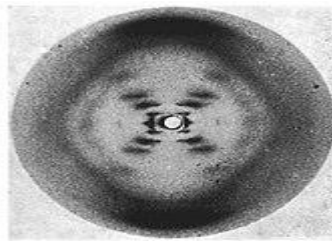
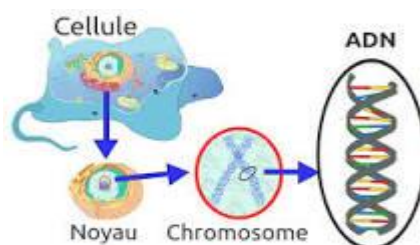


Photo 51, showing X-ray diffraction pattern of DNA

Photo 51, montrant le schéma de diffraction aux rayons X de l'ADN [86]



Figure(2-22)[87] :ADN

3- Analyse de l'ADN :

L'analyse des empreintes génétiques est une méthode extrêmement précise, unique et permanente durant toute la l'analyse de l'ADN pour déterminer l'identité de la personne, Il est impossible de trouver deux personnes qui ont le même ADN. Cet ADN est codé avec un algorithme binaire de 160 bits, utilisant la fonction de hachage pour protéger la confidentialité , et parmi les complexités entourant la question de la biométrie de l'ADN : l'accès à l'ADN et aux données, stockage de l'ADN , et du temps d'extraction , et de traitement [88].

4- Les Avantages et ses Inconvénients de l'ADN :

Parmi les avantages biométriques de l'ADN : Il est unique dans la précision des mesures et de l'identification des personnes. Et ses inconvénients : Lent à obtenir des résultats , et leur coût très élevé .

II-25- Modalité par voix

1- L'Historique

Dans les années 1940, des chercheurs des Bell Téléphone Laboratoires , ont mis au point un appareil spectrographe qui prend un enregistrement et trace le intensité du son dans l'enregistrement à différentes fréquences bandes. Dans les années 80, plusieurs entreprises développèrent des systèmes de reconnaissance de la voix .et dans les années 90 , les entreprises qui travaillent sur la voix comme IBM , mettre au point un système pour le marché commercial .Selon La société américaine - International Biométrie Groupes (IBG) , cette technologie biométrique détenait 4,3 % du marché en 2000 . [1]

2- La Biométrie comportementale

La Modalité Biométrique de la voix est L'Analyse des caractéristiques quantitatives : fréquences, harmoniques, puissance sonore, etc....., et est La technologie qui traite des données , et qui proviennent à la fois de facteurs physiologiques dépendants de l'âge, du sexe, l'accent , et de facteurs comportementaux comme la vitesse et le rythme , et permet de reconnaître un individu à distance est La Biométrie comportementale . Cette biométrie en général très bien acceptée car la voix humaine est trait unique de chaque individu et un signal naturel à produire, et se constituer des composantes physiologiques, et n'est pas un attribut permanent , et cette technique est très facilement falsifiable et elle nécessite une excellente d'enregistrement , Et elle rester une empreinte d'usage limité .

3- les Avantages et inconvénients de la modalité voix

1-a- Les Avantages :

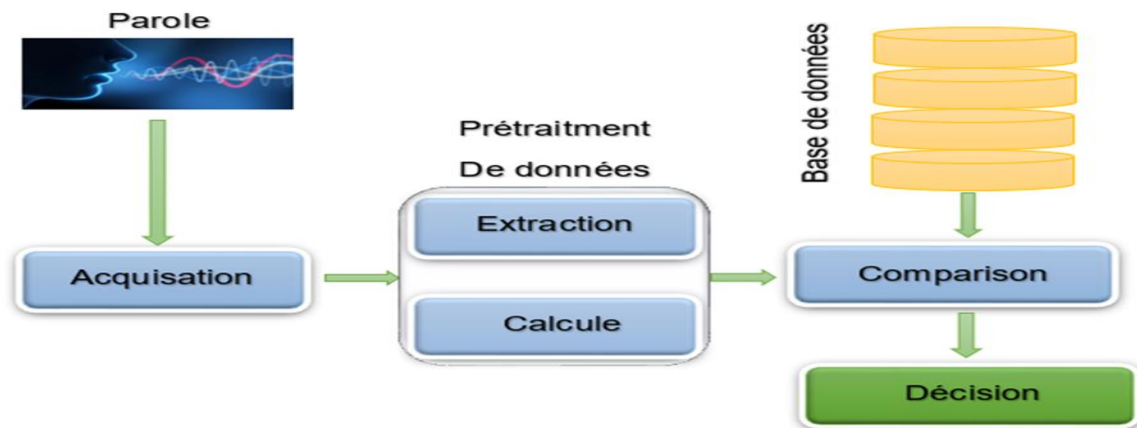
Aucun contact avec le capteur, Et la Sensibilité à l'état physique et émotionnel d'un individu, Et leur Capteur généralement disponible (microphones).

1-b- Les Inconvénients :

La Sensibilité aux conditions d'enregistrement du signal de parole par le bruit ambiant et la qualité du capteur (microphone) utilisé, et la Fraude possible par enregistrement de la voix de la personne.

II-26- Le Système biométrique

La biométrie de voix est un examen indolore. Elle est réalisée à distance de l'appareil. Les mesures sont prises en quelques secondes après avoir parlé. Comme les autres systèmes biométriques, le système biométrique par la voix repose sur les mêmes principes, en commençant par la prise de voix, et passant par l'analyse de ce voix, l'extraction des données, la comparaison avec la banque des données, puis la prise de décision.



Figure(2-23) : Architecture d'un système biométrique basé sur la voix

II-27- Les autres modalités biométriques :

Il existe d'autres méthodes biométriques qui n'ont pas été abordés et qui ont été utilisés dans les systèmes d'authentications biométriques, dépend des caractéristiques comportementales associées à une personne, basé sur les veines de la main, la rétine, l'odeur corporelle, la forme de l'oreille, les empreintes palmaires, la forme des lèvres, le rythme de frappe sur un clavier, La signature (une forme largement utilisée d'identification et de vérification), la démarche, et d'autres méthodes biométriques associées aux caractéristiques psychologiques. Il est important de noter qu'il n'existe aucune caractéristique biométrique idéale.

II-28- Les techniques de la transmission sécurisée de l'information

Selon leurs objectifs premiers, nous avons trois technologies de protection et de sécurité de l'information qui sont, La cryptographie, La stéganographie et le Tatouage, qui sont des technologies qui traitent de la protection et de la sécurité de l'information. La figure (2-24) suivant représente les différentes Techniques de la sécurité de l'information.

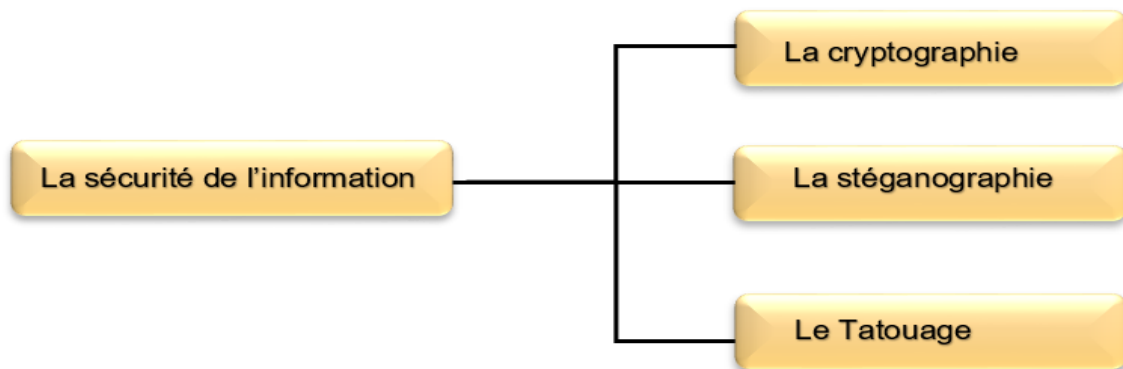


Figure (2-24) : Techniques de la sécurité de l'information

II-29- Cryptographie

Le mot cryptographie est un mot d'origine du grec ancien et se compose de deux mots : Le premier mot est "cryptos" pour dire caché et le deuxième mot est "graphein" pour dire écriture. La cryptographie est une des disciplines de la cryptologie, Elle est utilisée depuis l'Antiquité, mais certaines de ses méthodes les plus modernes, comme la cryptographie asymétrique, datent de la fin du XXe siècle, elle est s'attachant à protéger des messages, en s'aidant souvent d'utilisation des codes secrets ou de clés des décryptages, qui fait passer inaperçu un message dans un autre message, La cryptographie traditionnelle est l'étude des méthodes permettant de transmettre des données de manière confidentielle, et la cryptanalyse, est l'inverse. [89]

La cryptographie est la pratique de la protection des informations par l'utilisation d'algorithmes codés, de hachages et de signatures. Les informations peuvent être au repos comme un fichier sur un disque dur, ou en transit comme une communication électronique échangée entre deux ou plusieurs parties, ou en cours d'utilisation lors de l'utilisation des données, alors que la cryptographie rend un message supposément inintelligible à autre que qui-de-droit. La cryptographie est principalement utilisée pour protéger un message jugé confidentiel, on peut être utilisé dans des domaines très divers, comme le monde militaire, l'informatique, la protection de la vie privée, La cryptographie a quatre objectifs principaux [90] :

- **Confidentialité** : les informations ne sont accessibles qu'aux utilisateurs autorisés.
- **Intégrité** : garantit que les informations n'ont pas été manipulées.
- **Authentification** : confirme l'authenticité des informations ou l'identité d'un utilisateur.
- **Non-répudiation** : empêche un utilisateur de nier des engagements ou des actions antérieurs.

1- Comment fonctionne la cryptographie

Certains sont considérés comme basiques (on décale par exemple la lettre de l'alphabet d'un nombre déterminé de rang vers la droite ou vers la gauche), d'autres proposent un niveau de sécurité presque absolu.[91]

II-30- La stéganographie

La stéganographie est un mot du grec ancien et se compose de deux mots : Le premier mot est (steganós), Ce mot signifie " couvert ou caché " et le deuxième mot (graphé), et Ce mot signifie " écriture ". La stéganographie est un domaine où l'on cherche à dissimuler discrètement de l'information au sein d'un autre message , typiquement un signal de type texte, son, image, vidéo, etc.... , de façon qu'on ignore même l'existence du message secret et seule personne connaissant l'astuce est apte à extraire le message caché. Elle se distingue de la cryptographie qui est traditionnellement utilisée pour dissimuler des messages aux yeux de certains utilisateurs , et cherche à rendre un contenu inintelligible à autre que quide-droit.[93]

II-31- Le tatouag (watermarking)

1- Historique de tatouage (watermarking)

L'idée d'utiliser des tatouages est venue initialement en Italie , et en Europe pour identifier la marque ou le fabricant de papier, puis est venue déterminer la taille, la qualité et la couleur du papier. Le tatouage numérique est une discipline très récente ,et Sa naissance remonte au début des années 90 .

En 1990, un article publié par k.Tanaka traitait pour la première fois le tatouage d'images numériques . En 1993, A.Tirkel il a publié et discuté dans son article scientifique , la possibilité de coder le tatouage numérique "indélectable" sur une échelle de gris de 8 bits image.[92]

À partir de l'année 1995, le tatouage numérique a suscité beaucoup d'attention et s'est développé très rapidement , Il existe de nombreux sujets et recherches et des méthodes de travail et des systèmes pratiques ont été développés dans ce domaine , et A travers la création d'ateliers et de conférences dédiés à la sécurité de l'information, la sécurité, la dissimulation des données et la sécurité multimédia, le magazine EURASIP a traité en 2006 d'un sujet sur la sécurité de l'information, qui a mis en lumière le dynamisme du domaine. [92]

2- Définition de tatouage numérique (watermarking)

Il existe de nombreuses définitions proposées par différents auteurs dans le domaine informatique, électronique ou autre , et puisqu'il n'y a pas une définition unifiée du terme « tatouage numérique », Nous pouvons définir le tatouage numérique comme suit :

Le tatouage consiste à insérer une une marque invisible et et parfois visible , cette tatouage est appelée signature elle est indélébile dans une image ou dans d'autres documents numériques , le but de l'introduire d'une marque, le nom ou le logo de l'auteur, dans une image , ou une image de niveaux de gris , elle doit être connue uniquement par le propriétaire ou par l'emeteur .

Pour de la protéger contre les copies et divers buts tel que la lutte contre la fraude , le piratage informatique et la protection des droits d'auteur , et signifie l'incorporation d'une information numérique dans un contenu multimédia, comme une vidéo, un audio ou une image de telle manière que l'information insérée doit être imperceptible pour un observateur humain , et est inséparable des données et qu'il est résistant à de nombreuses opérations, sans dégrader la qualité du document initial ou original .Le tatouage est difficile à

remarque, est résistante aux attaques malveillantes, contient de nombreuses informations, peut coexister avec d'autres tatouages et nécessite peu de calculs pour être inséré ou détecté.[94]

3- Système de tatouage (watermark)

D'un point de vue généralisé, le système de tatouage (watermarking) peut être composé de trois étapes principales : première étape est l'intégration ou l'insertion de la marque tatouage c'est-à-dire l'insertion de l'information à cacher dans le signal hôte, et la transmission par le canal du signal tatoué, et la troisième étape est la récupération de tatouage, c'est-à-dire la récupération de l'information cachée à partir des données, tout les systèmes de watermarking prend la forme donnée sur la figure (figure 2-25). [95]

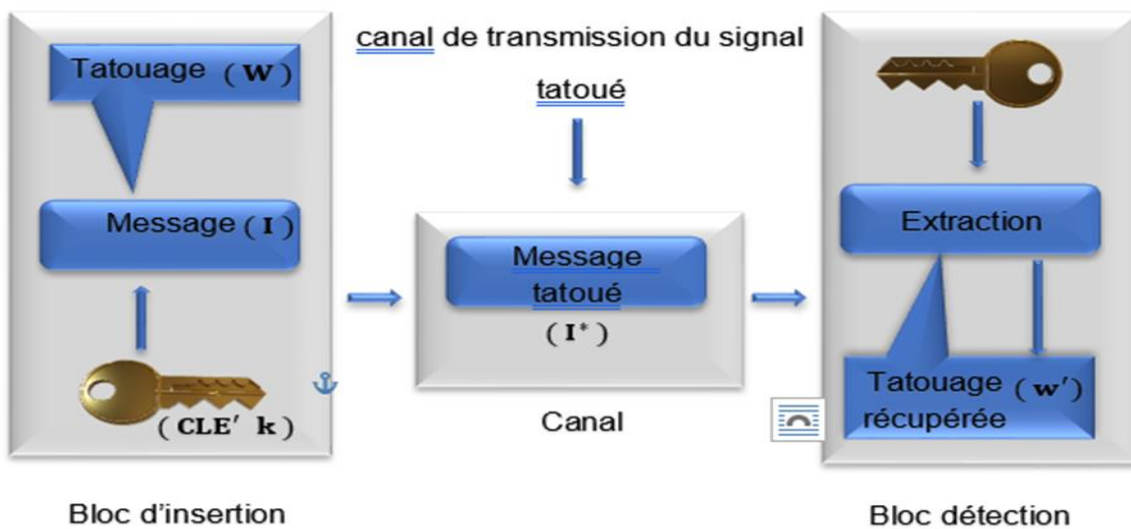
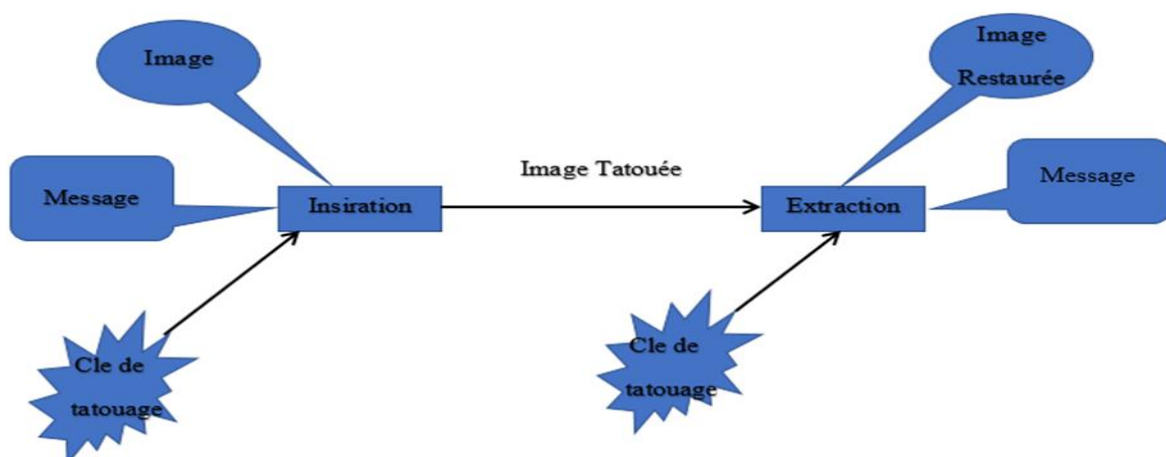


Figure (2-25) : Principe du système du tatouage.



Figure(2-26) :schemas de Principe du système du tatouage.

4- Les Propriétés d'un système de tatouage (watermarking)

Il n'y a pas d'ensemble unique de propriétés que tous les systèmes de tatouage doivent satisfaire, puisque l'importance relative de ces caractéristiques dépend de l'application comme la protection du droit d'auteurs, l'annotation cachée, l'authentification, et la communication secrète invisible. Chacune de ces applications a ses propres exigences il ya Certaines propriétés principales des systèmes de tatouage qui sont la fidélité, la robustesse, la capacité, la sécurité, et la complexité. il est difficile de concevoir un système de tatouage en respectant toutes ces propriétés au plus haut degré, Généralement tout dépend des exigences de l'application du système de tatouage, Certaines propriétés sont présentées suivantes [96].

5- Compromis entre robustesse, imperceptibilité et capacité de tatouage

Les techniques de tatouage numérique doivent trouver un compromis entre la capacité de tatouage, la robustesse et l'imperceptibilité. Différentes techniques ont leurs avantages et inconvénients, Le choix de la technique de tatouage dépend des exigences spécifiques de l'application. Les techniques de tatouage numérique doivent être à la fois robustes pour éviter la perte d'informations et imperceptibles pour ne pas altérer la qualité du média original. La technique de tatouage par modification de la fréquence est plus robuste mais peut entraîner une perte de qualité de l'image. Il existe différents compromis entre la capacité de tatouage, la robustesse et l'imperceptibilité en fonction des techniques de tatouage numérique utilisées. Enfin, la technique de tatouage basée sur la transformation de domaine offre une bonne imperceptibilité et robustesse, mais a une capacité de tatouage limitée. Le choix de la technique de tatouage dépend des exigences de l'application. [97].

6- Classificaion du tatouage numérique

Chaque type d'application possède ses propres techniques de tatouage, Les techniques de tatouage ont différents types de classifications en fonction de la nature de son application, on peut être classé plusieurs catégories selon différentes caractéristiques résumé comme suit :

1-a- Domaine Tatouage (Temporel , Fréquentiel)

Le tatouage numérique est Fusionner ou bien Masquer des secretés informations dans les médias, c'est-à-dire est une technologie en cours de développement pour assurer et faciliter l'authentification des données, la sécurité et la protection des droits d'auteur des médias numériques tels que l'image, l'audio et la vidéo. en utilisant un algorithme fort et approprié, à l'aide de cette algorithme ne peut pas être facilement détecté Le tatouage numérique. A moins qu'il connaisse l'algorithme, si non il sinon il devrait savoir le tatouage. l'insertion de la marque est effectuée dans un domaine délicate dans la conception du système de tatouage avec Différents critères régissent le choix d'un domaine adapté Ce domaine peut classées en deux catégories :

a-a- Les techniques dans le domaine spatial (temporel)

Ont une faible robustesse aux attaques géométriques, et plus simples, et plus coûteuses en temps d'exécution, et une complexité de calcul réduite.

a-b- Les techniques dans le domaine fréquentiel offrent

Une meilleure robustesse à la compression et moins sensibles aux attaques géométriques, et une complexité de calcul plus élevée, ce domaine est réalisé après une décomposition par

transformé telle que DCT (Discrete Cosine Transform)[22], DWT (Discrete Wavelet Transform), DFT (Discrete Fourier Transform) , SVD(Singular Value Decomposition).[105]

a-c- Media (Video , Image , Audio , ou texte) :

Classification des tatouages en fonction des médias (l'image, la vidéo,l'audio,texte) , Pour ce qui est de techniques de tatouage vidéo actuelle traitent les images vidéo comme une séquence d'images fixes , et tatouent chacune d'elles , Ansi que L'œil humain est également incapable de distinguer plus d'environ 20 images par seconde. Par conséquent, le tatouage vidéo numérique est résumé dans une approche image par image Le tatouage de vidéo numérique se résume souvent à des approches image par image comme ci-dessous [100]:

$$2-10 \quad \check{f}_t = f_t + \alpha w_t$$

Avec

f_t : est la trame vidéo originale à l'instant t

\check{f}_t : version tatouée de trame vidéo originale à l'instant t

α : la force de tatouage

w_t : le signal de tatouage

Les techniques de tatouage audio est plus difficile , Comparé au tatouage image et vidéo , en raison du manque d'impulsions répétées dans le signal audio , et c'est ce qui distingue la sensibilité élevée du système auditif humain (HAS), qui est supérieure à celle du système visuel humain (HVS).

7- Les applications du tatouage numérique

Le tatouage a été développé comme outil pour la protection des droits d'auteur, mais rapidement d'autres applications sont apparues . généralement Le terme de tatouage regroupe en fait un ensemble des techniques qui n'ont pas toutes les mêmes objectifs et ne visent pas le même type d'applications , et les Différentes applications déterminent la durabilité requise pour un tatouage , et Très souvent, il est nécessaire d'insérer des informations supplémentaires dans un document sous forme numérique, telles que de la musique, un fichier texte, une vidéo ou une image . comme exemple L'application protection du copyright elle-même regroupe plusieurs techniques (traçage, preuve de propriété, protection de copie, contrôle automatique de diffusion) dont les contraintes et les contextes d'utilisation sont très différents. on peut classer Les principales applications de tatouage numérique comme suit[101] :

8- Protection des droits d'auteur

La protection des droits d'auteur a été une des premières applications étudiée en tatouage .et l' utilisation du tatouage pour un copyright, ne doit pas forcément contenir beaucoup d'informations car son existence prouve que le document est protégé , et indique le propriétaire du document , et Le contenu numérique peut être intégré avec des tatouage représentant des métadonnées identifiant les titulaires des droits d'auteur .Ce service reste cependant toujours d'actualité et concerne encore la majorité des publications.[102]

a- Identification :

Le tatouage numérique peut être utilisé pour stocker des informations d'identifications personnelles, telles que le nom, l'adresse, etc.

b- Sécurité :

Le tatouage numérique peut être utilisé pour sécuriser les documents confidentiels, etc.

c- Protection contre la copie**d- Empreintes digitales****e- Signature****f- Vérification du contenu****g- Authentification****h- Vérification de l'intégrité****j- Masquage d'informations****k- Surveillance de la diffusion****m- Système****x- Renforcement****II-31-2- Les attaques menaçant le tatouage****1- La sécurité**

La technologie Watermark a connu de nombreux développements rapides au cours des dernières années successives, concerne le comportement du système de tatouage face à des attaques intentionnelles qui visent à rendre le tatouage inutilisable, et donc la définition de la sécurité d'un système de tatouage est différent selon les applications. et donc on peut dire que Toutes transformations, volontaires ou involontaires, ayant une influence directe sur le tatouage, sont appelées des attaques, et le points fort d'un tatouage efficace réside dans sa robustesse. Néanmoins, certaines transformations basiques peuvent effacer le marquage. [103]

2- L'attaque par tatouage

On peut être décrite Une attaque comme tout traitement qui contourne les données, et détruire volontairement ou involontairement le tatouage, comme suite :

2-a- Dynamique :

La modification de l'amplitude, et d'atténuation fait partie du groupe d'attaques dynamiques qui affectent la caractéristique d'intensité sonore, des données d'un signal audio tatoué. Tels que, la compression, Re-quantification, et la modification de la fréquence.

2- b- Filtre :

Des filtres passe-bas ou passe-haut sont utilisés pour ne laisser passer qu'une partie du signal, ou une partie spécifique du contenu fréquentiel et bloquer le reste, le changement de spectre en fonction de ce type de filtre entraîne une détérioration le tatouage et suppression le tatouage caché dans ce partie.

2-c- Bruit :

L'ajout de bruit détruit le tatouage , et l'ajout de tout ce qui provoque des changements indésirables dans le signal hôte à partir de différentes sources est une attaque courante. Exemple : bruit aléatoire , et bruit gaussien blanc additif.

2-d- Conversion :

Les signaux audio sont soumis à différents formats , c'est-à-dire la Conversion d'une formats de signal audio (échantillonnage ou rééchantillonnage et inversion.).

2-e- Compression avec perte :

Compressée le signal , qui basé sur le effets psychoacoustiques de l'audio.

2-f - Sample Permutation :

Les attaques qu'ils rendre le tatouage intégré inutile (modifier ou supprimer).

2-g- Time stretch (également connu sous le nom de pitch shift) :

Time stretch (également connu sous le nom de pitch shift) , ils font l'une des suivant : faites varier la durée de l'audio sans affecter la hauteur ou vice versa, ce qui est fait pour affiner ou ajuster l'audio dans des fenêtres temporelles. [104]

CONCLUSION DE CHAPITRE

Ce chapitre traite de plusieurs sujets liés au traitement du signal, aux caractéristiques du signal de la parole, au système auditif, à l'appareil vocal de l'homme et aux organes responsables du son et de l'ouïe. Il présente également le modèle équivalent de l'appareil auditif ainsi que l'appareil équivalent vocal.

Dans ce chapitre, le signal de la parole est défini et les techniques utilisées pour améliorer la qualité du son et extraire des informations utiles de la parole sont présentées. Le système auditif humain se compose de plusieurs parties, dont l'oreille externe, l'oreille moyenne et l'oreille interne, qui sont responsables de la conversion du signal sonore en signal nerveux via les cellules sensorielles de l'oreille interne. La biométrie est également définie, la différence entre les types des biométries et les utilisations les plus importantes des types des biométries actuellement utilisés par l'homme sont discutées. L'utilisation de techniques de codage de la parole pour convertir la parole parlée en données numériques pour le stockage et l'analyse est également abordée. Ces technologies comprennent le codage audiovisuel, les technologies de reconnaissance vocale et la traduction automatique , La cryptographie, La stéganographie ,Le Tatouage .

En somme, le traitement du signal, les propriétés du signal vocal, le système auditif et l'appareil vocal des humains comprennent de nombreuses technologies importantes pour comprendre et améliorer le son et la communication.

CHAPITRE (III)
ETUDE , EXPERIENCES et DISCUSSION

III-1 Introduction de chapitre

Le troisième chapitre est un chapitre appliqué qui traite de quatre parties d'expériences et de travaux qui sont au cœur de notre recherche. Avant d'aborder la diversification de la protection du signal de parole, il est nécessaire d'étudier la transformée en ondelettes discrète, et l'ondelette mère est la meilleure dans le traitement du signal de parole, et les niveaux d'analyse en ondelettes, ainsi que l'étude de l'énergie retenue dans les niveaux de transformée en ondelettes discrète et Comparaison le taux d'énergie d'un signal du parole de niveau trois à l'aide de la transformée en ondelettes discrète (DWT) de l'ondelette mère Haar, Meyer, biorthogonal, coiflets, symlets, daubechies. Dans la partie troisième, nous avons présenté une nouvelle approche, et un nouveau critère pour déterminer le sexe du locuteur, que nous avons nommé « Taux D'énergies d'ondelettes » (WER), c'est une nouvelle approche pour l'identification et la vérification du sexe du locuteur à l'aide de l'énergie de premier niveau DWT et du passage à zéro. Dans la partie quatrième du chapitre Trois, nous avons introduit une nouvelle approche du processus de protection du signal vocal grâce à de nouvelles méthodes qui rendent difficile la pénétration et la détection du signal vocal par les pirates et les attaques extérieures, selon les expériences, les chiffres, et les données qui sont présentés à ce chapitre.

III-2 historique de Transform d'ondelette:

En 1807, Joseph Fourier développe une méthode de représentation d'un signal basée sur la fonction d'analyse. En 1909 Alfred Haar introduit la forme en ondelettes dans sa thèse de doctorat. Et en 1946 le scientifique Dennis Gabor (Dennis Gabor), le mathématicien hongrois a proposé une technique de fenêtres de signal similaire à la transformée de Fourier. En 1985, Meyer a introduit l'analyse par ondelettes, qui a permis aux chercheurs, et aux mathématiciens de construire leurs propres familles d'ondelettes à l'aide de paramètres dérivés. En 1989, Stéphane Mallat a développé un algorithme pour l'application de la transformée en ondelettes discrètes à l'aide des filtres. En 1998, Ingrid Daubechies a utilisé la théorie de l'analyse des ondes polycycliques pour construire sa propre famille d'ondelettes faciles à mettre en œuvre, qui sont des ondelettes orthogonales appelées "ondelettes de Daubechies", et sont devenues la pierre angulaire des applications des ondelettes plus tard [105]

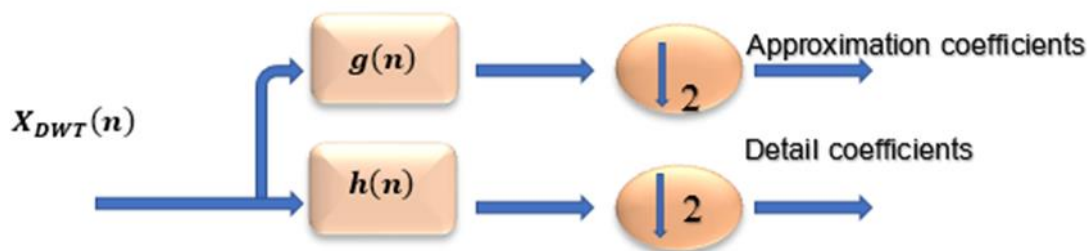
III-3 Transformée en ondelettes discrète (DWT)

La transformée en ondelettes discrète (DWT) remonte à 1983, elle est basée sur le codage en sous-bande, il est développé ensuite pour être une technique de codage hiérarchique pour être une technique similaire au codage en sous-bande, est identifié par la relation suivante, mentionnée par de relation (3-1) [106].

$$3-1 \quad X_{DWT}(n) = x(n) * h(n) = \sum_{k=-\infty}^{\infty} x(k) * h(n - k)$$

La transformée en ondelettes discrète (DWT) est une technique d'analyse de signaux qui utilise différentes de bandes de fréquences avec différents degrés des précisions pour décomposer le signal d'origine $x(n)$, en coefficients approximatifs, et détaillés. Les coefficients approximatifs, ont les amplitudes les plus élevées, et les composantes de basse fréquence, tandis, et possèdent la majeure partie de l'énergie, que les coefficients détaillés, ont les amplitudes les plus basses et les composantes de fréquence les plus élevées, qui peuvent être associées au bruit présent dans le signal d'origine, alors que le bruit du signal est concentré dans la haute fréquence [119], [106].

Comme mentionné précédemment, le DWT utilise des techniques de filtrage numérique, où le signal à analyser est passé à travers des filtres ayant différentes fréquences, et différentes échelles. Le signal d'origine est divisé, en coefficients de détails, et en coefficients d'approximations, souvent désignés respectivement par a_k et d_k , où d est une lettre désignant les coefficients détaillés, et a désigne les coefficients approximatifs, tandis que k désigne le niveau analytique. Les deux paramètres sont le résultat de la convolution du signal d'origine $X(n)$ avec le filtre, ce qui signifie que les paramètres approximatifs sont le résultat de la convolution du filtre passe-bas $g(n)$ avec le signal d'origine $X(n)$, et que les paramètres détaillés sont le résultat de la convolution du filtre passe-haut $h(n)$ avec le signal d'origine $X(n)$. [107]. c'est -a-dire que la DWT dépend de la convolution.



figure(3-1) : Analyses en coefficients approximatifs, et en coefficients détaillés (Discrete Wavelet Transform DWT) [108]

La figure (3-1) représente le modèle d'analyse de forme d'ondelettes (DWT), et les deux relations (3.2-a) et (3-2-b) ci-dessous, représente la convolution du signal $X(n)$ avec les filtres passe-bas, et passe-haut. [109]

$$3-2-a \quad a_1 = \sum_{k=-\infty}^{\infty} X(k)g[2n - k]$$

$$3-2-b \quad d_1 = \sum_{k=-\infty}^{\infty} X(k)h[2n - k]$$

Dans l'analyse de forme d'onde, le signal est analysé et synthétisé par étapes qui commencent par diviser le signal $x(n)$ jusqu'à la fin du niveau requis, et le reconstruire du dernier niveau auquel nous nous sommes arrêtés, jusqu'à début du premier niveau du la quel nous commencé, c'est-à-dire vice versa.[119]

III-4 Transformée en ondelettes continue (CWT):

La transformée en ondelettes continue (CWT) , est similaire à la transformée de Fourier en termes d'intégration sur tout le temps du signal, mais diffère en dimensions, en ce que le signal est projeté sur un plan d'échelle de temps bidimensionnel, contrairement à la transformée de Fourier qui est dimensionnelle dans le domaine fréquentiel. Le CWT peut extraire les caractéristiques du signal à un certain niveau d'échelle de temps, et il est défini comme l'intégrale sur l'ensemble du signal multipliée par différentes formules pour l'échelle (s) et la transition (τ) de la fonction d'ondelette parente. [110]

La transformée en ondelettes continue (CWT) est utilisée dans sa relation mathématique avec la corrélation, et elle effectue une analyse multi-solutions en réduisant , et en développant des fonctions d'ondelettes [111], qui est exprimée par l'équation (3-3)et (3-4). [106]

3-3

$$X_{CWT}(\tau, s) = \frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} x(t) \Psi^* \left(\frac{t-\tau}{s} \right) dt$$

(*) complexe conjugué de la fonction Ψ

translation

Constant normalisation

signal

Echelle (scale)

3-4

$$\psi_{(s,\tau)}(t) = \frac{1}{\sqrt{s}} \psi \left(\frac{t-\tau}{s} \right) , S > 0 , \sqrt{S} > 0 , \tau \in R$$

III-5 Etude et discussion de la premier partie

Le traitement , et l'analyse du signal vocal peuvent être effectués de diverses façons, telles que le filtrage, l'optimisation, et la suppression du bruit, chacune ayant son propre modèle. La transformée en ondelettes, soit la Discrete Wavelet Transform (DWT) ou la Continuous Wavelet Transform (CWT) .

La transformée en ondelettes est l'un des outils les plus performants dans le traitement du signal image et des autres signaux et basée sur la transformation STFT de Fourier court-temps, où l'on modifie la largeur de la fenêtre pour obtenir des ondelettes dont la fréquence varie en fonction de la largeur de la fenêtre utilisée , elle fonctionne en analysant le signal d'original selon un groupe de petites ondelettes de longueur limitée et de valeur moyenne égale à zéro comme représenté par la figure (3-2), et qui sont les fonctions de base des ondelettes , et elle peut utilisée pour l'image ou la parole dans des nombreuses applications. nous utilisant une ondelette mère pour analyse de signal d'original en plusieurs niveaux dans les domaines fréquentiels , et temporels, selon différents types et durées [112]

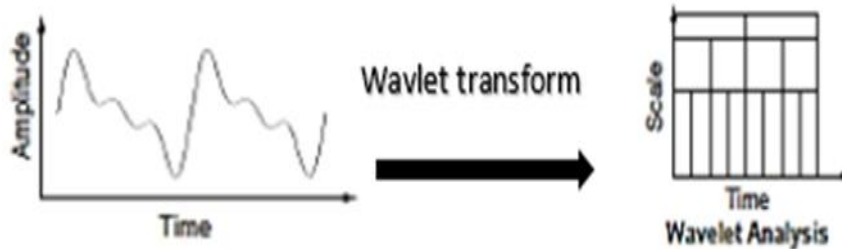


Figure (3-2) : Transformée d'ondelettes d'un signal

1- Les Caractéristiques de la transformée en ondelettes :

La transformée en ondelettes se caractérise par sa flexibilité dans la conception de la forme de l'impulsion, sa moindre sensibilité à la distorsion, et aux interférences de canal, et l'efficacité de son spectre, et de ses algorithmes dans les ondelettes utilisées dans la compression des données.

La transformée en ondelettes est divisée en deux branches : la transformée en ondelettes continue (CWT), et la transformée en ondelettes discrète (DWT). Il existe des nombreuses ondelettes mères, chacune avec certaines caractéristiques, et critères de base, notamment :

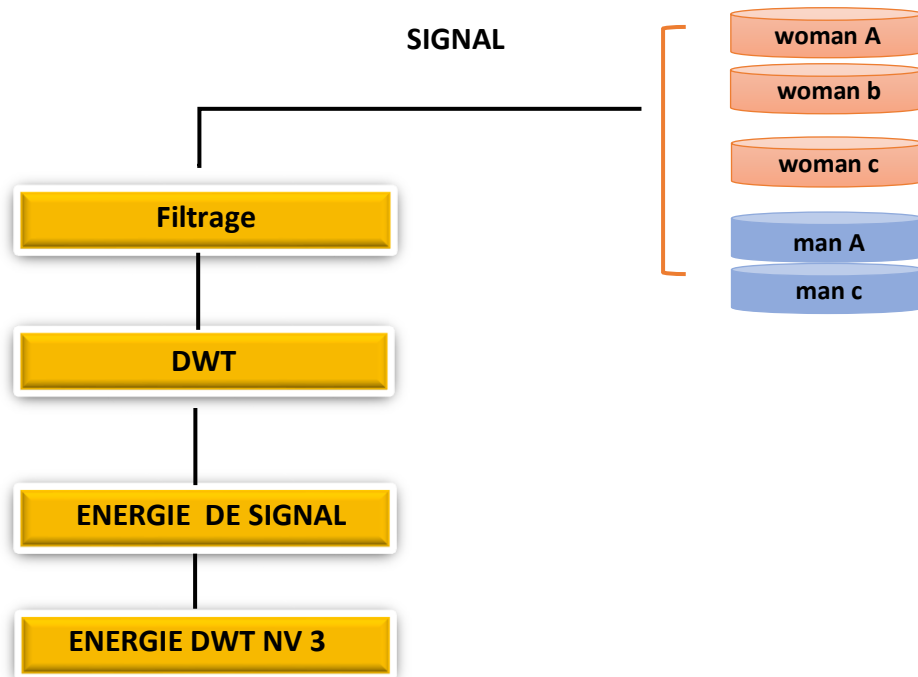
1- La caractéristique standard du degré de douceur et de régularité, et c'est important dans la reconstruction des signaux.

2- La propriété de la vitesse de convergence de la fonction ondelette $\psi(t)$ vers zéro lorsque le temps t dévie vers l'infini en dehors du domaine de $[a, b]$, qui détermine la position de l'ondelette dans le temps.

3 - La propriété de symétrie afin d'éviter les changements de phase dus à la présence d'ondelettes symétriques et asymétriques, [112]

Dans cette étude de première partie de chapitre troisième, nous avons cherché à trouver l'ondelette mère la plus appropriée pour extraire le taux d'énergie retenu dans le troisième niveau après l'analyse du signal de parole (wav) par DWT, en suite en comparant les différentes branches d'ondelettes mères pour choisir la meilleure. Cela se fait selon la quatrième étapes, comme le montre la figure (3-3).

2- Organigramme de la première partie



La figure (3-3) : l'organigramme quantité d'énergie retenue dans le troisième niveau du coefficients approximatifs de la transformée en ondelettes discrète (DWT) .

3- discussion de la première partie

Dans cette partie du chapitre, nous nous sommes concentrés sur un critère que nous avons appelé le symbole E_{ra3} , qui consiste à trouver la quantité d'énergie retenue dans le troisième niveau du coefficients approximatifs de la transformée en ondelettes discrète (DWT) comparé aux à la l'énergie totale du signal étudié. c'est-à-dire trouver l'énergie en pourcentage, Où l'on analyse le signal donné en trois niveaux par la transformée en ondelettes discrète, et on néglige les coefficients détaillés, car ils ne sont pas importants, tandis que l'on garde les coefficients approximatifs, Puis nous l'analysons à nouveau de la même manière jusqu'à atteindre le troisième niveau, et à ce niveau on calcule l'énergie dans les coefficients approximatifs après avoir calculé l'énergie dans le signal d'original, il faut préciser que l'énergie calculée, que ce soit dans le signal d'origine ou dans les coefficients approximatifs pour le troisième niveau, est de l'énergie dans le domaine temporel.

En rappelant que l'énergie dans le domaine temporel est équivalente à l'énergie dans le domaine fréquentiel, selon l'équation de L'égalité de Parseval, Après avoir calculé l'énergie pour les deux signaux, on calcule le taux d'énergie retenue E_{ra3} , Puis nous le divisons les coefficients approximatifs de le troisième niveau de la transformée en ondelettes discrète E_{a3} , sur l'énergie du signal d'original $E_{X(n)}$, et multiplions le quotient par 100, selon l'équation suivante (3-5) :

3-5

Le taux d'énergie retenue

L'énergie de 3 -ème niveaux DWT

$$E_{ra3} = \frac{E_{a3}}{E_{x(n)}} \times 100$$

l'énergie du signal

La figure (3-4) représente l'extraction d'énergie après analyse du signal d'origine filtré par un filtre Butterworth en signal approximatif de trois niveaux d'une transformée en ondelettes discrètes à l'aide de la fonction d'ondelette mère biorthogonal (bior3.9) , le signal parole analysé est la phrase " Do you speak English " , qui a été prononcée par une femme. [04].

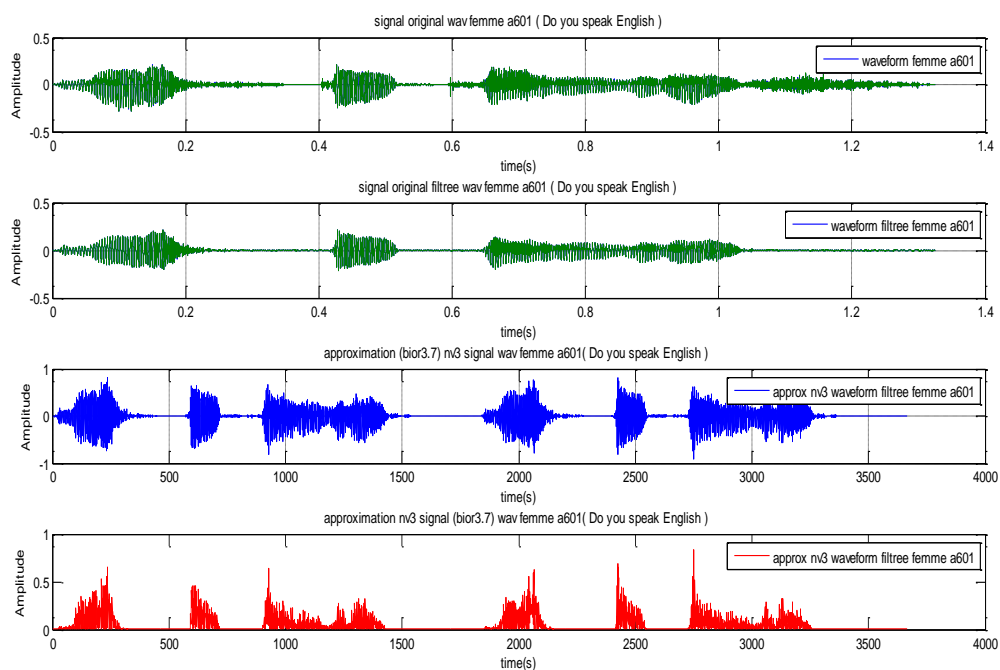


Figure (3-4) : Représentation de la division d'un signal wav en trois niveaux DWT et représentation en energie .

III-6 Etude et discussion de la deuxième partie

Le signal de parole, est un signal non stationnaire, complexe et variable dans le temps . De plus, la fréquence la plus élevée du signal de parole, est de 5 kHz. Les différences dans les caractéristiques linguistiques des êtres humains sont caractérisées par la variance des énergies pour les deux sexes . généralement le signal parole à trois caractéristiques : le Hauteur , Fréquence fondamentale , Energie .cette dernière , est sont petites chez les femmes par rapport les hommes .

L'objectif de ce travail de la partie deuxième de troisième chapitre , est de Comparé le taux d'énergie d'un signal du parole de troisième niveau de la transformée en ondelettes discrète (DWT) de l'ondelette mère Haar , meyer , biorthogonal , coiflets , symlets , daubechies représenter dans le tableau (3-2) , pour déterminer une gamme des valeurs de deux sexe du locuteur. ce signal du parole est filtré par un filtre Butterworth d'ordre 8 .

A partir de la transformée en ondelettes discrète (DWT) de l'ondelette mère , nous calculant l'énergie , du troisième niveau de la DWT , de chaque ondelette mère pour les deux sexes , puis en comparant les valeurs d'énergie des hommes , et des femmes du signal parole de la même personne , pour la même phrase., et de la même phrase pour les deux sexe Ce comparaison a été testé sur plusieurs bases de données hommes , et femmes avec différentes phrases répétées pour la même personne ou pour les deux sexes , cette Comparaison il a donné des résultats acceptable sur lequel nous pouvons travailler.

1- Loi de Shannon

Les opérations effectuées sur le signal du parole , telles que l'échantillonnage sans distorsion, et autres, nécessitent des fréquences conformément à la loi de Shannon , que la fréquence d'échantillonnage , est plus de deux fois la fréquence du signal d'origine étudié [118].

$$f_e = 2 f_{max}$$

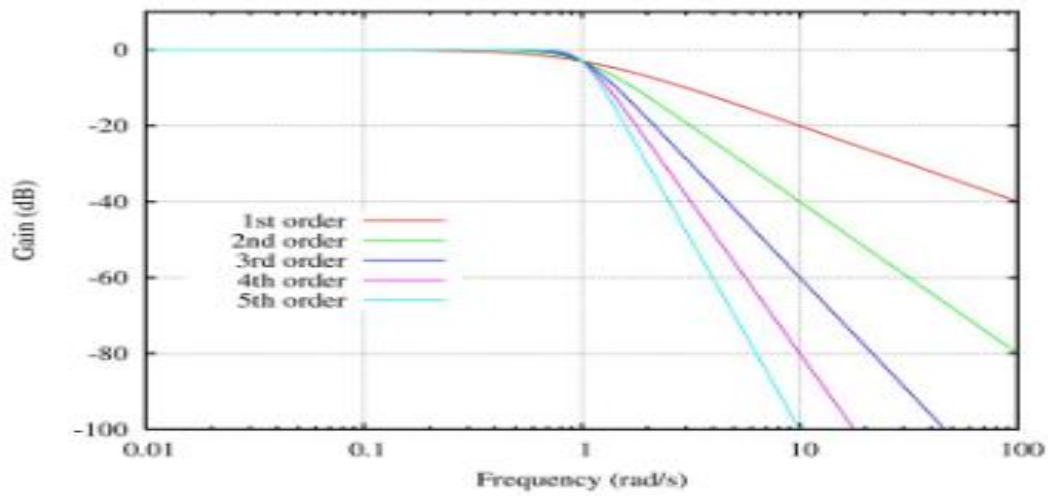
f_e : la fréquence d'échantillonnage

f_{max} : la fréquence du signal d'origine étudié

2- Filtrage

L'étape de filtrage est accomplie par une transformation mathématique appliquée sur le signal de parole à l'étude , en utilisant un filtre passe-bas avec une fréquence de coupure $f_c = 600 \text{ hz}$, Ce filtre passe-bas réduit la haute fréquence du signal à $-10 \log_2 10$, c'est-a-dire de -3dB , ce qui signifie diminuer l'énergie de sortie du signal à 71% du signal de parole d'original [121].

Parmi les filtres les plus largement utilisés , figurent les filtres linéaires Butterworth, qui sont de forme similaire avec une différence dans la plage de fréquences de coupure , Ce filtre a la plus grande amplitude , et plus stable avec une fréquence dans la plage de passage , conçu pour posséder un gain aussi constant que possible dans sa bande passante , et dans la région de transition avec une réduction modérée , ce filtre est sélectionné en fonction de la précision de l'amplitude [350]. Sur un diagramme de Bode logarithmique, cette réponse décroît linéairement vers $-\infty$, L'ordre du filtre représente le nombre de colonnes dans la région de passage du filtre, par exemple, un filtre d'ordre (n) a un taux de réduction de $(-6*n) \text{ dB/octave}$ sera $(-20*n)/\text{decade}$, lorsque n=8 sa réduction est de -48dB/octave sera de -160dB/decade à partir de $F_c=1$ [120] .



Figure(3-5) :présente l'ordre de filtre Butterworth

L'équation générale (3-6) pour filtrer un signal wav à l'aide d'un filtre Butterworth pour extraire les propriétés du locuteur est :

$$3-6 \quad y(n) = x(n) - G * h(n)$$

G : est un facteur d'échelle qui garantit que le signal de sortie a l'amplitude souhaitée est généralement égal à 1 ou 0.99 [113].

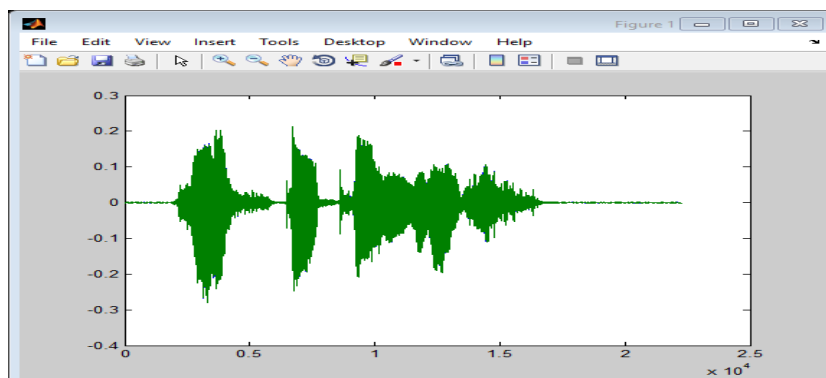
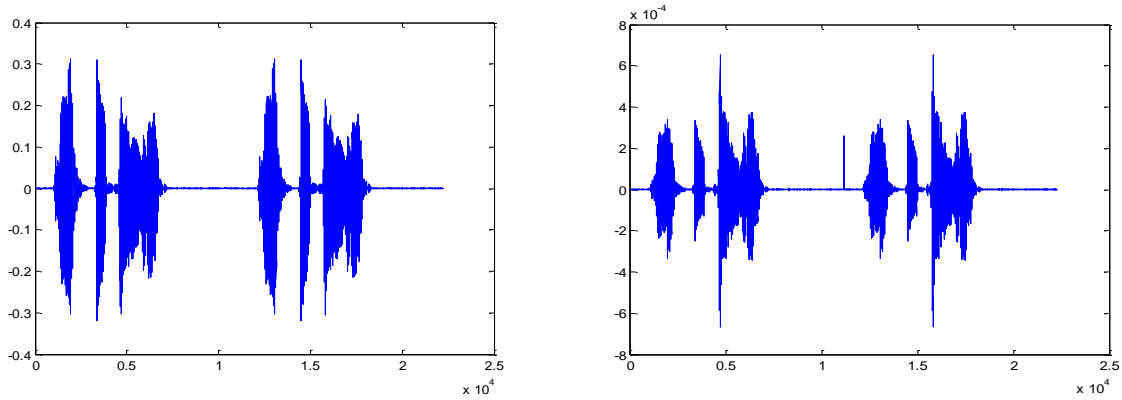


Figure (3-6): signal original wav filtre



Partie approximation (a1) du signal wav

Partie détaillée du signal wav (d1)

Figure (3-7) : Décomposition DWT du signal de premier niveau après filtrage

3- Caractéristiques du signal vocal

Le signal de parole est un support audio d'informations complexes et diverses, et possède des caractéristiques telles que la caractéristique de fréquence fondamentale, qui est souvent notée f_0 . D'autres caractéristiques sont également l'énergie, et le spectre de fréquence. La vibration, ou le cycle d'ouverture/ la fermeture des cordes vocales représente la fréquence fondamentale f_0 . Cette fréquence ne caractérise que les segments voisés, est évolué lentement dans le temps

a- Fréquence fondamentale

La fréquence fondamentale, également appelée fréquence fondamentale de la voix, est la fréquence la plus basse d'un son périodique qui est perçu comme la hauteur de la voix. Dans le cas de la voix humaine, la fréquence fondamentale correspond à la fréquence de vibration des cordes vocales lors de la production du son. Elle est mesurée en Hertz (Hz), et est généralement représentée par la lettre f_0 . La fréquence fondamentale est un paramètre important dans l'analyse acoustique de la voix et est utilisée dans de nombreuses applications telles que la synthèse vocale, la reconnaissance de la parole et l'analyse des émotions [119]. La fréquence fondamentale du locuteur varie selon son âge, et son sexe. Elle s'étend approximativement de 60 Hz à 150 Hz pour les hommes, et de 150 Hz à 450 Hz pour les femmes. L'extraction de la fréquence fondamentale n'est pas une tâche facile car la périodicité de vibration des cordes vocales n'est pas toujours parfaite.

b- Amplitude

L'amplitude est une mesure de l'énergie d'une onde, telle que le son ou la lumière. Pour une onde sonore, l'amplitude est liée à la pression acoustique de l'onde, qui correspond à la variation de pression de l'air autour de son état d'équilibre. L'amplitude peut être utilisée pour déterminer le niveau de volume perçu. Plus l'amplitude est grande, plus le son est fort. L'amplitude peut également être utilisée pour déterminer la qualité du son, en particulier pour les instruments de musique. Dans le cas d'un signal parole, l'amplitude du signal de parole change considérablement dans le temps, et dans la parole audible, cette amplitude est beaucoup plus grande que dans la parole inaudible, qui nous reflète les changements d'énergie dans le signal, généralement mesurée en unités telles que le pascal (Pa) pour les ondes sonores, et le volt (V) pour les ondes électromagnétiques, et en dB pour le signal de parole [119], [122].

c- Energie

L'une des caractéristiques du signal de parole est la caractéristique énergétique, généralement symbolisée par E. L'énergie d'un signal peut être utilisée pour déterminer si le signal, d'une nature transitoire (signal impulsionnel) ou d'une nature périodique (signal sinusoïdal). Un signal transitoire a une énergie finie tandis qu'un signal périodique a une énergie infinie.

L'énergie d'un signal est également utile pour déterminer la puissance moyenne d'un signal. La puissance moyenne d'un signal est la quantité d'énergie transmise par unité de temps. Elle est calculée en divisant l'énergie du signal par la durée de la période de temps T. La puissance moyenne d'un signal est souvent utilisée pour caractériser les signaux non-périodiques tels que les signaux audio et les signaux électriques. L'énergie est très élevée dans les sons audibles par rapport aux sons inaudibles dans lesquels l'énergie est faible, [123] and [122], elle est calculée selon l'équation (3-7) .

En résumé, l'énergie d'un signal est une mesure importante pour caractériser la quantité d'énergie contenue dans un signal dans le domaine temporel. Elle peut être utilisée pour distinguer les signaux transitoires des signaux périodiques, et pour déterminer la puissance moyenne d'un signal.

$$3-7 \quad E_x = \sum_n^{N-1} (X(n))^2$$

L'énergie d'un signal est une mesure de la quantité d'énergie contenue dans le signal. Elle peut être utilisée pour caractériser un signal dans le domaine temporel.

	femme a 402	femme b 203	femme d302	homme a 102
haar	2.3536e+05	1.1466e+05	1.0610e+06	2.8128e+06
coif 5	3.0122e+05	1.4552e+05	1.3415e+06	3.1079e+06
sym8	2.9938e+05	1.4396e+05	1.3270e+06	3.0941e+06
bior 3,9	6.6329e+05	3.3010e+05	2.9330e+06	4.3590e+06
db45	3.0773e+05	1.5067e+05	1.3833e+06	3.1641e+06
max	6.6329e+05	3.3010e+05	2.9330e+06	4.3590e+06

Tableau (3-1) : comparaison d'énergie avec des chiffres globale d'une phrase (Do you speak English) de deux sexe on utilisant Le troisième niveau DWT

	Haar	meyer	daubechies	symlets	biorthogonal	oiflets
haar	haar					
Meyer		Myer				
db			db45			
sym				Sym8		
bior					bior3.9	
coif						Coif5

Tableau (3-2) : comparaison de la plus grande d'énergie avec tous les branches des ondelettes, on utilisant Le troisième niveau DWT, pour la même phrase (Do you speak English).

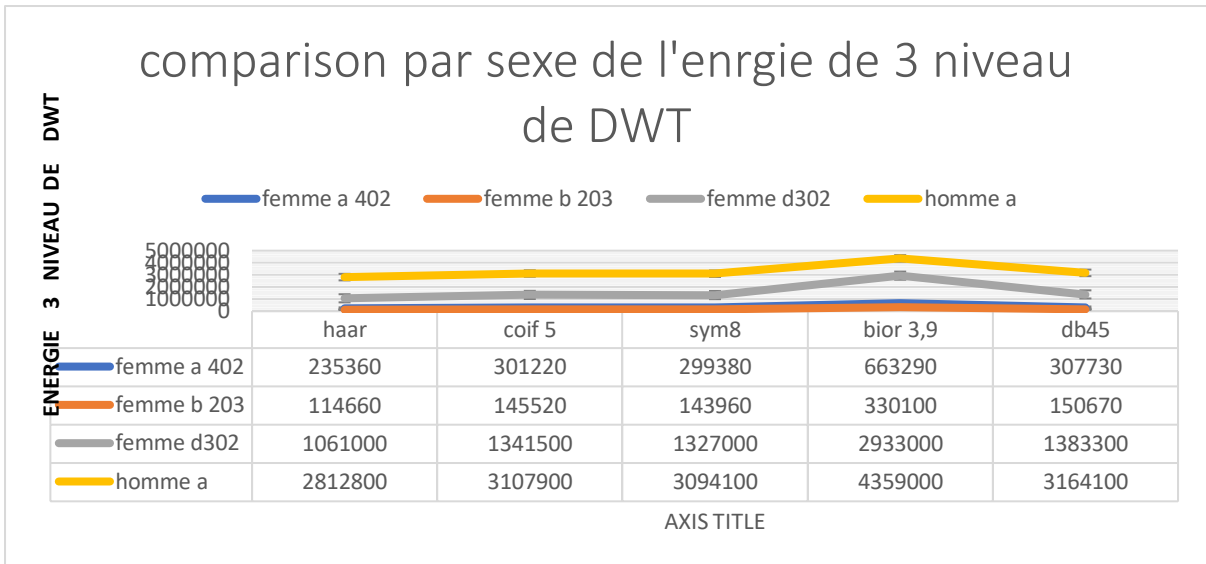


Figure (3-8) : comparaison globale d'une phrase (do you speck english) de sexe 3 femmes (a ,b ,d) on utilisant DWT

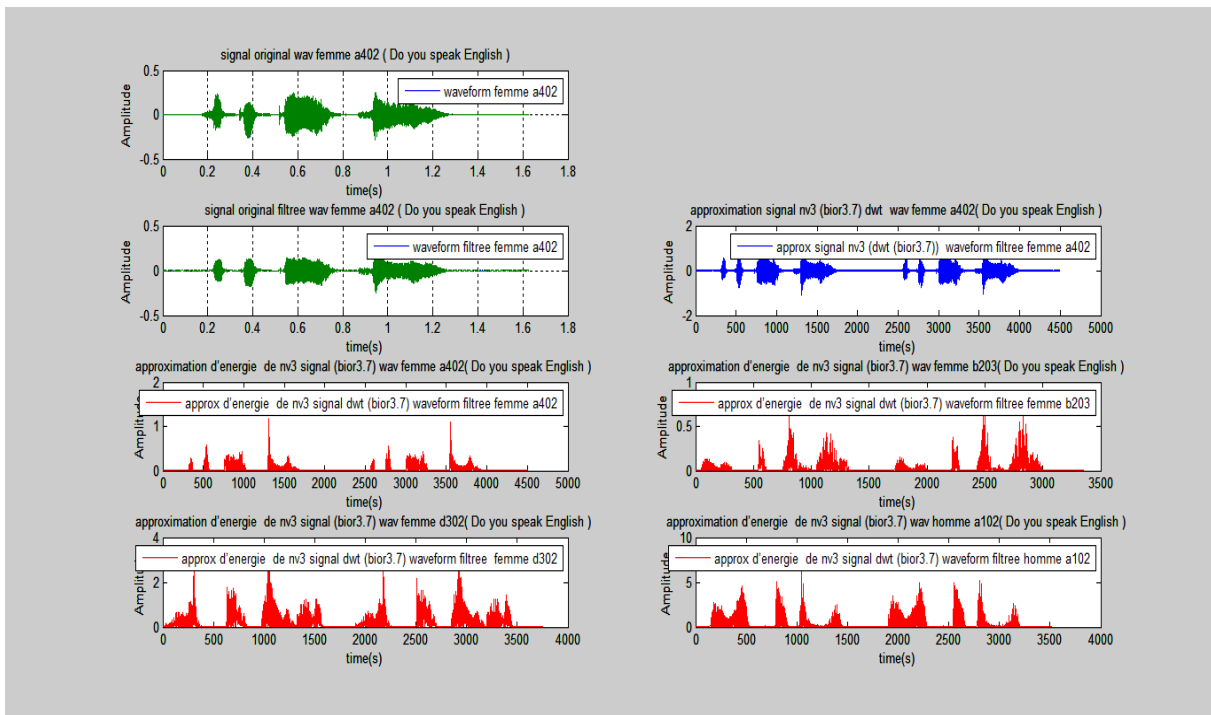


Figure (3-9) : comparaison globale d'une phrase (do you speck english) de sexe 3 femmes (a ,b ,d) on utilisant DWT

III-7 Etude et discussion de la troisième partie

L'objectif de la troisième partie de chapitre trois est de trouver un nouveau critère pour déterminer une gamme de valeurs afin de déterminer le sexe du locuteur. A partir de la transformation en ondelettes discrète (DWT) de l'ondelette mère 'Daubechies, db7' , en calculant l'énergie , et le passage par zéro du premier niveau de la DWT , suivi en calculant les valeurs du critère pour les deux sexes , puis en comparant avec la valeur de la fréquence de fondamentale de signal du parole des hommes , que pour les femmes , ou le même personne , pour la même phrase. Ce critère a été testé sur plusieurs bases des données hommes , et femmes avec différentes phrases répétées pour la même personne ou pour les deux sexes , il a donné un résultat acceptable sur lequel nous pouvons travailler.

Dans l'ensemble, ce travail propose une nouvelle méthode pour déterminer le sexe d'un locuteur à l'aide de la DWT, comparer les valeurs calculées avec la gamme de fréquences fondamentale pour les hommes , et les femmes. La méthode a été testée sur différentes bases de données , et a montré des résultats prometteurs . D'autres recherches pourraient explorer le potentiel de cette méthode dans d'autres applications , telles que la reconnaissance , et l'authentification du locuteur.

Les différences dans les caractéristiques linguistiques des êtres humains sont caractérisées par la variance des fréquences pour les deux sexes. Les fréquences fondamentale des femmes sont petites par rapport aux fréquences fondamentales des hommes . Habituellement, le son se compose de vibrations de différentes longueurs , et hauteurs . l'oreille humaine ne peut pas entendre toutes les vibrations, puisque certaines sont audibles, certaines sont douloureuses pour l'oreille humaine, et certaines sont inaudibles, et cela se situe dans la plage de : 20 Hz à 20 kHz. En dessous de cette plage, les fréquences sont inaudibles , et au-delà de cette plage, elles sont douloureuses pour l'oreille humaine. Bien sûr, l'être humain, en vieillissant, a une plus petite limite d'audition, parfois jusqu'à 17 kilohertz [129]. Le signal de parole est non stationnaire, complexe et variable dans le temps. De plus, la fréquence la plus élevée du signal de parole est de l'ordre de 5 kHz. Comme les opérations effectuées sur ce signal, telles que l'échantillonnage et autres, nécessitent des fréquences conformément à la loi de Shannon présente a la première partie de ce chapitre , que la fréquence d'échantillonnage est plus de deux fois la fréquence du signal d'origine étudié [118].

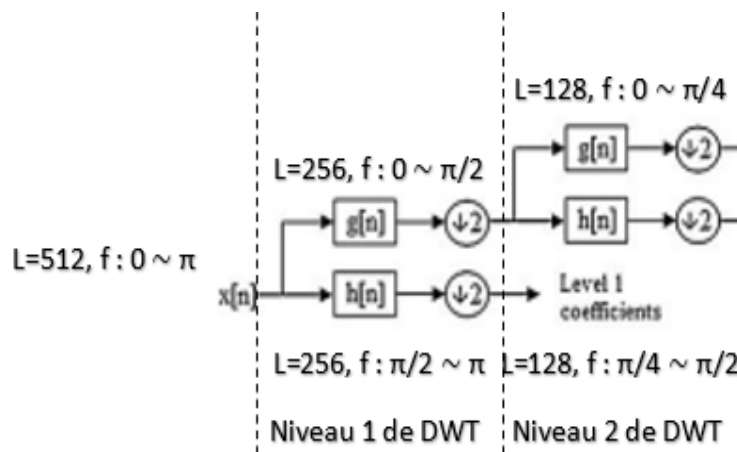
De nombreux travaux de recherche ont autrefois porté sur l'étude et la détermination de la plage de valeurs de la fréquence fondamentale f_0 comme dans [06] , qui traitait de l'extraction de la fréquence de fondamentale dans les deux domaines du temps , et de la fréquence (HPS) , ainsi que des ondelettes (ACEP) . De plus , le travail de [125] porte sur une comparaison entre les méthodes d'estimation pour le calcul du fréquence fondamental de nouveau-né . d'autre part Dans [128], ils ont présenté des techniques pour déterminer le rôle de la fréquence dans le temps ,d'autres travaux de recherche tels que : [128], [126] et [127] [04] , sont intéressés à ce sujet . les recherches se développent encore puisqu'il n'est pas possible de confirmer définitivement tel ou tel critère qui détermine le sexe du locuteur selon le domaine fréquentiel de fondamentale .

Selon le schémas illustré à la Figure (3-11), nous symbolisons la norme appelée << Wavelet Energy Rate >> (WER), qui est appliquée au premier niveau de La transformée en ondelettes (DWT) du signal de parole à l'étude. Ce signal, est une phrase prononcée par deux sexe de locuteurs , contenant des phrases répétées , et différentes , Ce sont des échantillons d'une base de données que nous avons préparée spécifiquement pour cette étude, ils sont des vidéos de phrases pour apprendre l'anglais de format mp4 pour les deux sexes . Ces vidéos sont téléchargées , ensuite converties au format wav , Chaque phrase est échantillonnée séparément

selon la fréquence d'échantillonnage $f_e=11025\text{Hz}$, qui est une fréquence supérieure à deux fois la fréquence du signal de parole 5kHz comme mentionné précédemment sur la base de la loi de Shannon.

Dans l'analyse de forme d'onde, le signal est analysé et synthétisé par étapes qui commencent par diviser le signal $x(n)$ jusqu'à la fin du niveau requis, et le reconstruire du dernier niveau auquel nous nous sommes arrêtés, jusqu'au début du premier niveau duquel nous avons commencé, c'est-à-dire vice versa.

Par exemple, si le signal d'origine $x(n)$ contient 512 échantillons et une fréquence de 0 à π , alors chacun des coefficients approximatifs, et détaillés du premier niveau, est de 256 échantillons, et la moitié ($1/2$) de la fréquence du signal pour chacun d'eux, c'est-à-dire $\pi/2$, et au deuxième niveau, il y a 128 échantillons, et la demi-fréquence du niveau du signal, c'est-à-dire $f = \pi/4$ ou $1/4$ de la fréquence du signal d'original pour les coefficients approximatifs, et détaillés. La figure 5 montre le processus d'analyse en ondelettes d'un signal original $x(n)$ contenant 512 échantillons, et la fréquence $f_0 = \pi$.



Figure(3-9) :

Figure (3-10) : Signal decomposition into level2 [109]

Comme mentionné ci-dessus, la transformée en ondelettes a plus de flexibilité dans la conception de la forme d'impulsion et moins de sensibilité à la distorsion du signal.

Le DWT analyse le signal d'origine $x(n)$ dans différentes bandes de fréquences avec différents degrés de précision en analysant le signal en coefficients approximatifs et en coefficients détaillés. Les coefficients approximatifs ont l'amplitude et les composants de basses fréquences les plus élevés et possèdent la majeure partie de l'énergie du signal, On sait que le signal de parole est continu dans lequel la fréquence fondamentale d'une voix d'homme descend à 50 Hz , c'est-à-dire sa période $T = 20\text{ ms}$, et la fréquence de fondamentale d'une voix de femme descend à 180 Hz , c'est-à-dire sa période $T = 250\text{ ms}$. De plus, la fréquence fondamentale pour la voix de femme monte à 500 Hz , tandis que pour la voix d'homme, elle monte à 200 Hz .

2- Organigramme de troisième partie

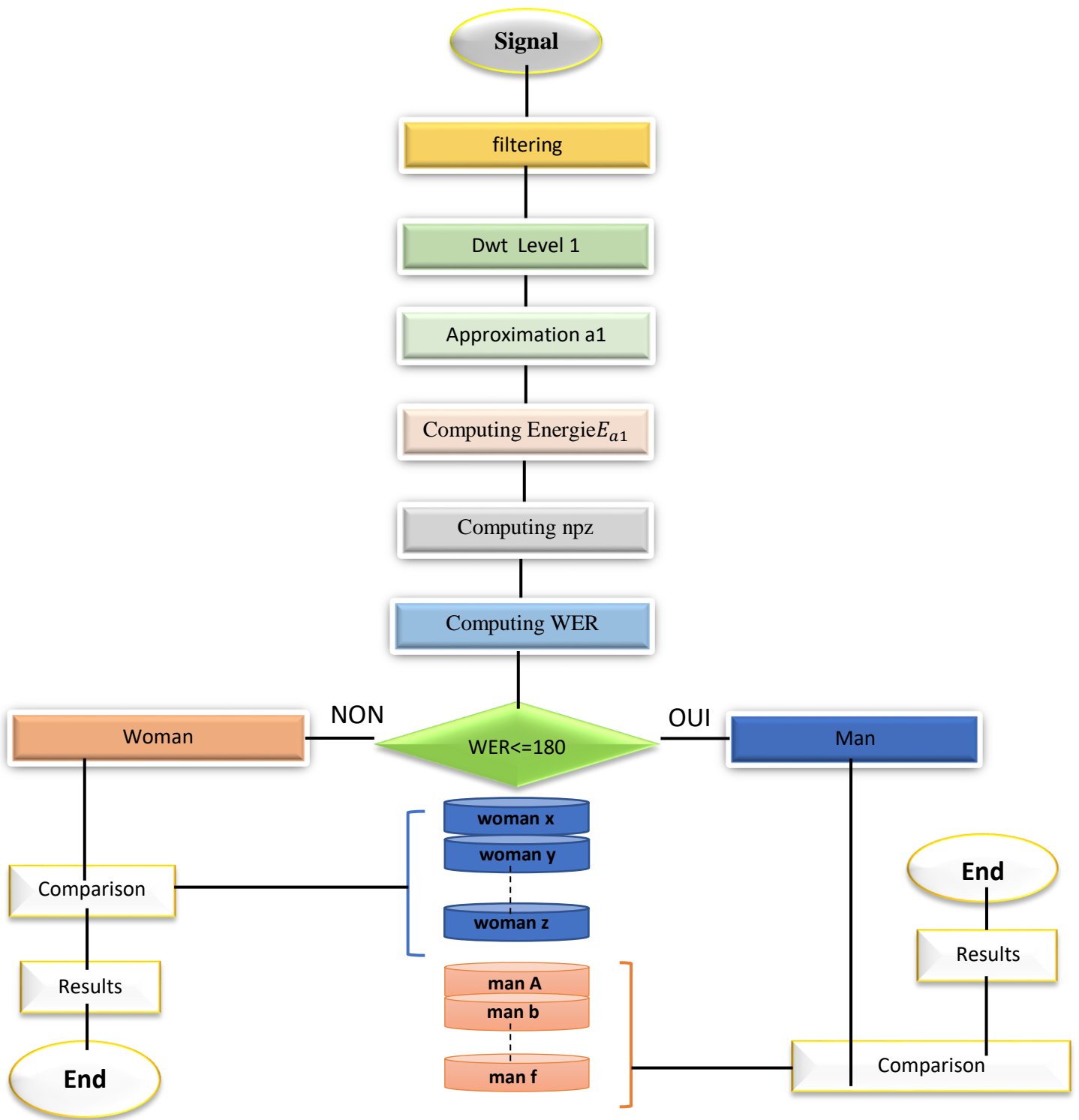


Figure (3-11): Organigramme de calule WER

3- Le passage par zéro

Le passage par zéro, est le nombre de fois que le signal $x(n)$ passe par zéro dans un certain intervalle de temps. C'est une indication de la fréquence à laquelle l'énergie est concentrée dans le spectre du signal, car l'énergie est concentrée aux basses fréquences dans la parole audible, tandis que dans la parole inaudible, la majeure partie se trouve dans les hautes fréquences [114]. À partir de là, le plus grand nombre de passage par zéro est aux hautes fréquences, et le plus bas aux basses fréquences.

En conséquence, nous pouvons dire que la parole audible a le nombre de passage par zéro, est nulle [310]. Ensuite, nous pouvons dire que le passage du signal par zéro a une forte relation avec la distribution de l'énergie avec la fréquence, et le nombre de passage par zéro est calculé pour le signal selon l'équation (3-8), [340] :

$$3-8 \quad npz = \frac{1}{2N} \sum_{n=0}^{N-1} |\text{sgn}[x(n)] - \text{sgn}[x(n-1)]|$$

$$3-9 \quad \text{sgn}[x(n)] = \begin{cases} +1 & x(n) \geq 0 \\ -1 & x(n) < 0 \end{cases}$$

4- Taux d'énergie d'ondelettes proposé (WER)

Le critère Taux d'énergie d'ondelettes (WER), est calculé selon l'équation (3-10), où ce nouveau critère permet de déterminer le sexe du locuteur à partir des valeurs étudiées, et il a été nommé comme précédemment mentionné WER, qui est le quotient de l'énergie E_{a1} obtenue à partir de le coefficient d'approximation du premier niveau de la DWT appliqué sur le signal $x(n)$ selon l'équation (3-11) divisé par le nombre de passage par zéro pour le même signal multiplié par T_{pz} le rapport passage par zéro pour le signal d'origine :

$$3-10 \quad \text{WER} = \frac{E_{a1}}{npz} \times T_{pz}$$

$$3-11 \quad E_{a1} = \sum_{n=0}^{N-1} (a_1(n))^2$$

a_1 : la coefficient d'approximation du premier niveau de la DWT

E_{a1} : L'Energie de la coefficient d'approximation du premier niveau de la DWT

T_{pz} : Taux de passage par zéro de signal original

npz : Nombre de passage par zéro la coefficient d'approximation du premier niveau de la DWT

Les valeurs du critère WER, changent selon une fourchette spécifiée allant de 50 à 180 pour les hommes, et de 180 à 500 pour les femmes, Ces valeurs sont des valeurs similaires à la fréquence fondamentale du signal de parole f_0 , qui détermine le sexe du locuteur dans un premier temps avant d'identifier le locuteur à l'aide de la base de données par comparaison avec d'autres valeurs de notre base des données existante préalablement préparée comme référence pour les personnes parlantes de deux sexes.

III -8 Etude et discussion de la quatrième partie

1- définition

Pour identification et vérification d'un locuteur , et extraire ses informations personnelles susceptibles de faire l'objet d'attaques externes ou des piratages, plusieurs systèmes ont été conçus utilisant plusieurs algorithmes dont le plus important est mfcc et lpc et d'autre . Le processus d'extraction de caractères passe par les étapes suivantes :

- 1 – Traitement du signal primaire
- 2 - Filtrer le signal
- 3 – Appliquer la transformée discrète en ondelettes
- 4 - Couper le signal avec trames
- 4 – Fenêtre (windowing)
- 5- Autocorrélation
- 6- Lpc
- 7- Extraction de caractéristiques par algorithmes lpcc
- 8- Comparaison des données

Ci-dessous, nous proposons un schéma de traitement du signal de parole et de reconnaissance et de vérification du locuteur à l'aide d'une nouvelle approche basée sur des calculs matriciels et la traçabilité matricielle.

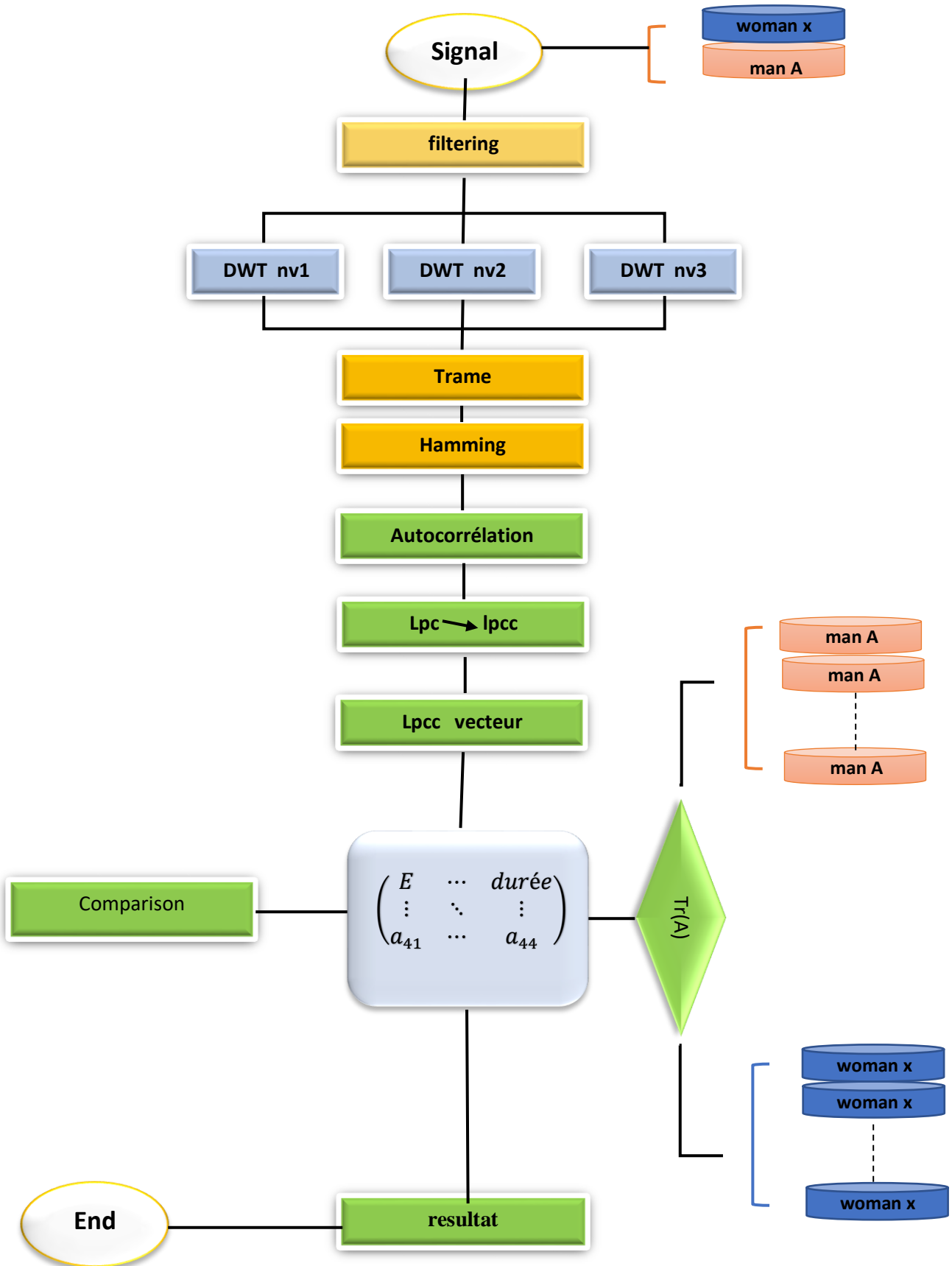
2- Traitement du signal primaire

La fréquence la plus élevée du signal de parole chez l'homme est de 5 khz , comme cela a été discuté dans les premier , et deuxième chapitres . puisque le processus de deviser le signal nécessite deux fois sa fréquence , donc le processus de deviser le signal du parole nécessite une échantillonnage de fréquence au moins le doubles , c'est a dire de 11,025 kHz ou 11025hz , qui fait partie des valeurs connus dans les processus appliqués au signal de parole soit pour l'enregistrement des signaux audio ou d'autres applications sur le signal parole , et sont connus avec des valeurs 11025 Hz, 22050 Hz et 44100Hz.

3- L'étape de filtrage

Dans cette expérience et selon l'équation d'un filtre Butterworth de degré $n=8$ a été appliqué au signal de parole avec une fréquence de coupure égale à $f_c=600$ Hz pour supprimer les fréquences hauts qui est lie au bruit , et d'une part de éliminer des fréquences qui dépasse 600 Hz car, comme on le sait, la fréquence fondamentale la plus élevée est Pour les femmes, qui ne dépasse pas cette valeur dans tous les cas .

4- Organigramme



la figure (3-12) : schéma proposé pour traiter le signal parole et Pour identification ,
vérification d'un locuteur

5- Segmentation de signal

Le signal de parole est découpé en échantillons audio constitués de taille de fenêtre dans une plage de durée temps allant de 20ms à 30ms ou 40ms qui est supposé stable pour une cette durée avec un recouvrement majoritairement de 10 ms. Ceci pour résoudre le problème du signal de parole, qui est un signal variable et non-stationnaire avec échantillonnage de fréquence $f_e=11025\text{hz}$.

Puisque nous prenons la taille de fenêtre avec une durée de 20ms ou 0,02 s, et que nous avons la fréquence d'échantillonnage du signal parole $f_e=11025$, alors le nombre d'échantillons sur lesquels nous travaillons dans le signal du parole sera égale à 220 échantillons pour chaque trame, l'équation suivante donne comment calculer :

$$\text{Nbrs éch} = f_e * \text{durée (sec)}$$

$$\text{Nbrs éch} = 11025\text{hz} * 0.03 = 330 \text{ ech (pour 30ms)}$$

6- Fenêtrage (windowing)

En traitement du signal lors de découpage d'un signal parole en trames de longueur (N) d'échantillons est généralement entre 10 ms à 30ms tel que les trames successives peuvent être séparées par une valeur (k) de leur durée, pratiquement le recouvrement est de 10ms ou bien 75 de la taille de trame. voir la figure(3-13).

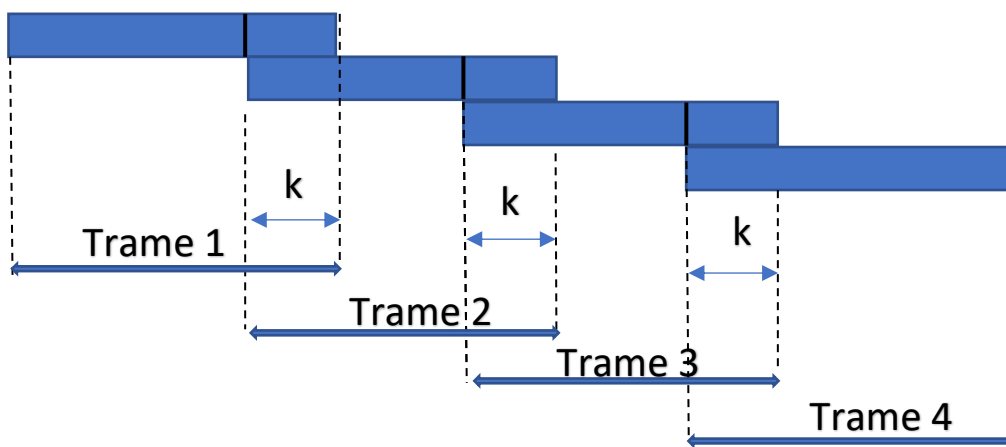
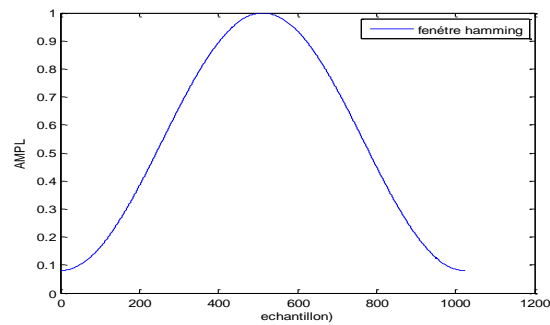


Figure (3-13) : découpage d'un signal parole

L'application d'une fenêtre est pour réduire la discontinuité à la fin de trame et pour réduction progressive des bords. Donc on pondère le signal de cette trame avec la fonction (fenêtre) de pondération. Généralement en traitement de signal parole la fenêtre la plus utilisée est (**Hamming**) défini par l'équation :

$$3-12 \quad w(n) = \begin{cases} 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right), & 0 \leq n < N - 1 \\ 0 & \text{ailleurs} \end{cases} \quad \text{si non}$$

Et sous forme d'onde et sa réponse fréquence . voir figures (3-14) .



Figure(3-14) : domaine temporelle fenêtre Hamming

Dans le domaine temporel il utiliser la multiplication de signal de trame de longueur (N) $\{s(0).....s(N-1)\}$ par une fenêtre pondération . voir figure (3-15).

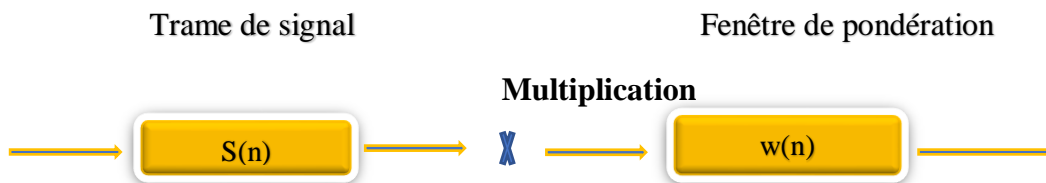
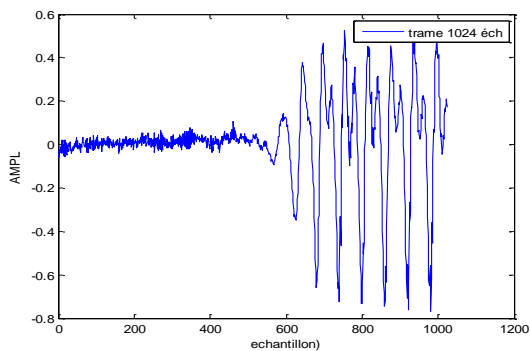
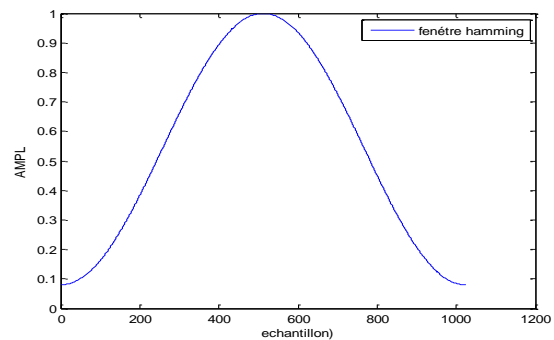


Figure (3-15) : trame pondérée multiplie par fenêtre

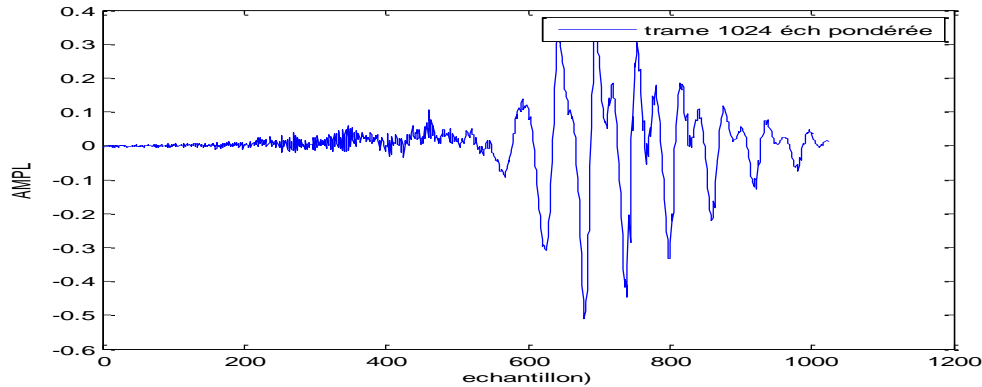
Le résultat d'une pondération de trame d'un signal parole échantillonner de 1024 échantillons et défini : voir figure (3-16- c) .



Figure(3-16-a) : trame de signal temporel 1024 éch



Figure(3-16- b) : fenêtre hamming domaine temporelle



Figure(3-16-c) : trame de signal pondérée

Parmi les fenêtres les plus utilisées pour pondération de signal parole on a [51]:

Rectangulaire :

$$w(n) = \begin{cases} 1 & \text{pour } 0 \leq n < N - 1 \\ 0 & \text{si non} \end{cases}$$

Hanning :

$$w(n) = \begin{cases} 0.5 - 0.5 \cos\left(\frac{2\pi n}{N-1}\right) & \text{pour } 0 \leq n < N - 1 \\ 0 & \text{ailleurs si non} \end{cases}$$

Hamming :

$$w(n) = \begin{cases} 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right), & 0 \leq n < N - 1 \\ 0 & \text{ailleurs si non} \end{cases}$$

Après la pondération de trame on calcule la Transformée de Fourier discrète .

7- La fonction d'autocorrélation

La fonction d'autocorrélation est de déterminer à quel point deux échantillons successifs d'un signal se ressemblent . La fonction d'autocorrélation est une fonction mathématique qui mesure la corrélation entre les différentes valeurs d'un signal en fonction du décalage de temps . . La fonction d'autocorrélation est périodique avec la même période pour un signal périodique de période P :

$$3-13 \quad \phi(k) = (k + P)$$

Dans le cas d'un signal ergodique et stationnaire, la fonction d'autocorrélation est définie comme la fonction mesure corrélation entre deux valeurs du signal à des instants temporels différents est représenté par l'équation (.3-14.) [114] .

$$3-14 \quad \phi(k) = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=-N}^N x(n)x(n+k)$$

La fonction d'autocorrélation est une fonction paire en k , elle atteint son maximum pour $k = 0, \pm P, \pm 2P, \pm 3P \dots$ ou les valeurs égale l'Energie du signal . La fonction d'autocorrélation possède les caractéristiques suivante :

- La fonction est max : $\phi(k) \leq \phi(0)$ quelque soit k
- La fonction est paire : $\phi(k) = \phi(-k)$
- $\phi(0)$ est égale l'énergie du signal :

$$3-15 \quad \phi(0) = \sum_{n=-\infty}^{\infty} x^2(n)$$

Dans le traitement court-terme du signal vocal sur un nombre constant et fini de N échantillons sur un trame (segment), la fonction d'autocorrélation peut être calculée par [114] :

$$3-16 \quad \phi(k) = \frac{1}{N-k} \sum_{n=0}^{N-k-1} x(n)x(n+k)$$

8- Extraction et extraction de caractéristiques

Il existe de nombreuses façons d'identifier et de vérifier le locuteur, selon la variété des algorithmes utilisés, comme l'algorithme Lpcc (linear predictive cepstral coefficients), qui utilisent l'analyse de prédiction linéaire, qui est qui dépend du mécanisme de production de la parole, et aussi l'algorithme pLp, Qui dépend de la technique de la relation entre l'effet physique et les perceptions, et similaire à l'algorithme Lpc (Linear Predictive Coding), qui est un algorithme de codage à prédiction linéaire utilise une technique qui suppose que chaque échantillon audio, est linéairement lié avec un ensemble d'échantillons qui l'ont précédé. ainsi que l'algorithme mfcc (Mel Frequency cepstral coefficients) dont les filtres sont sensibles au signal audio, qui fait l'objet de notre étude dans cette quatrième partie du troisième chapitre, et sur lequel nous sommes appuyés sur cet algorithme pour d'extraire les caractéristiques pour identifier, et de vérifier le locuteur après avoir déterminé le sexe du locuteur dans la deuxième partie du troisième chapitre appliqué.

En reconnaissance de la parole, le rôle de La paramétrisation du signal parole est de fournir et d'extraire des informations caractéristiques du signal précis pour minimiser le temps de calcul, et produire une représentation moins redondante du signal. parmi les méthodes d'extraction des paramètres on peut citer le codage prédictif linéaire LPC, et le cepstre où on tente simplement de déconvoluer la source et le conduit.

9- La méthode d'analyse par prédiction linéaire

C'est une représentation directe du signal vocal sous la forme d'un nombre limité de paramètres. Sa puissance provient du fait qu'elle est fondée sur un modèle simple de production de la parole qui s'approche du système phonatoire.

Pour identification et vérification d'un locuteur, et extraire ses informations personnelles susceptibles de faire l'objet d'attaques externes ou de piratage, plusieurs systèmes ont été conçus utilisant plusieurs algorithmes dont le plus important est mfcc et lpc et d'autres. Le LPC (Linear Predictive Coding) généralement utilisé pour l'analyse des formants ou trames de (20 à 30 ms), Lpc est une technique utilisée dans le traitement du signal numérique pour modéliser l'enveloppe spectrale d'un signal vocal, et calcule un spectre de puissance logarithmique du signal, et conçu pour prédire l'échantillon actuel du signal vocal comme une combinaison linéaire des échantillons précédents. Il est largement utilisé dans le codage, la synthèse et la reconnaissance vocale [114]. Dans ce travail nous avons choisi la méthode lpc de paramétrisation. On distingue deux modèles de prédiction, les modèles auto régressifs AR représentés par l'équation (3-17) et (3-18), et ceux appelés auto régressifs à moyenne ajustée ou ARMA représentés par l'équation (3-19).

$$3-17 \quad x(n) = u(n) * h(n)$$

$x(n)$: signal sortie (simple $x(n)$ prévue)

$h(n)$: réponse impulsionnelle du filtre

$u(n) = \sum_k \delta(n - kp)$ est signal d'excitation pour :

- Sons voisés est un train périodique d'impulsions d'amplitude unité .
- Sons non voisés est un bruit blanc de moyenne nulle et de variance unité.
- p période fondamentale

Dans le domaine temporel :

$$3-18 \quad x(n) = \sum_{k=1}^p a(k)x(n-k) + \sigma u(n)$$

Avec :

- $a(k)$ les coefficients du filtre (les coefficients de prédiction)

On peut représenter le modèle (AR) par la figure (3-17) :

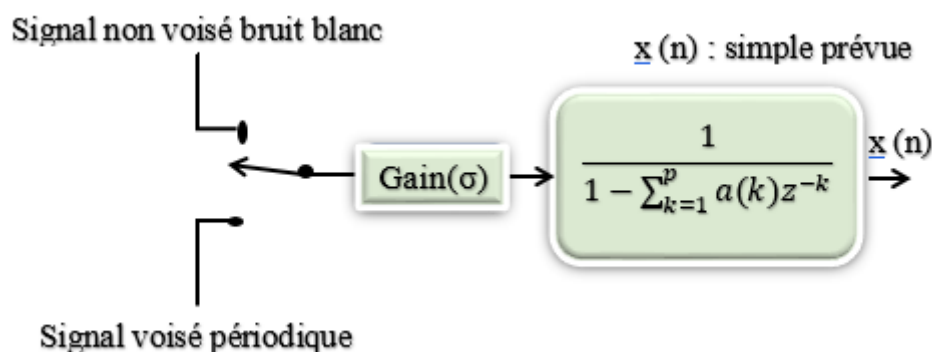


Figure (3-17) : modèle ARM

ARMA Dans le domaine temporel :

$$3-19 \quad x(n) - \sum_{k=1}^p a(k)x(n-k) = \sigma \sum_{i=0}^q c(k)u(n-k)$$

Avec

- (pq) ces l'équations d'ordre
- $c(0) = 1$

Donc on dire que l'équation (3-20) de prédiction d'un échantillon $x(n)$:

$$3-20 \quad x(n) = \sum_{k=1}^p a_k x(n-k)$$

10- Les coefficients cepstraux prédictifs linéaires LPCC

LPCC (Les coefficients cepstraux prédictifs linéaires) sont une variante du LPC qui utilise une transformation des coefficients LPC pour produire un ensemble de coefficients cepstraux. Les coefficients cepstraux sont obtenus en prenant la transformée de Fourier inverse du logarithme du spectre de puissance du filtre LPC. elle sont généralement utilisés comme caractéristiques pour les systèmes de reconnaissance vocale et d'identification du locuteur . c'est-à-dire que LPCC est obtenus par la dérivation a partir les coefficients LPC .[115] :

$$3-21 \quad \log\left[\frac{1}{A(z)}\right] = \sum_{n=1}^{\infty} c_n z^{-n}$$

$$A(z) = 1 - \sum_{k=1}^p a_k z^{-k}$$

$$\sum_{n=1}^p n a_n z^{-n} = \sum_{n=1}^{\infty} n c_n z^{-n} - \sum_{k=1}^p \sum_{n=1}^{\infty} n c_n a_k z^{-(n+k)}$$

$$3-22 \quad c_n = a_n + \frac{1}{n} \sum_{k=1}^{(p, n-1)} a_k (n-k) c_{(n-k)}$$

La figure (3-12) représente le mécanisme de fonctionnement de l'algorithme Lpcc qui dépend de la analyse prédictive linéaire , qui consiste à estimer les coefficients d'un filtre linéaire qui modélise l'enveloppe spectrale du signal de parole.

Le filtre est conçu pour prédire l'échantillon actuel du signal vocal comme une combinaison linéaire des échantillons précédents. En fait, les paramètres de prédiction réels ne sont pas

utilisés dans la reconnaissance vocale car ils présentent généralement une variance élevée, de sorte que ces paramètres de prédiction sont convertis en un autre ensemble de paramètres qui sont ('cepstral coefficients') [116].

On peut résumer le processus pour déterminer les coefficients LPCC par le schéma (3-12). Nous avons appliqué une nouvelle approche dans l'utilisation de la technologie LPCC, où nous avons choisi les 4 premières caractéristiques de vecteur LPCC, en utilisant les trois niveaux de la transformée en ondelettes discrète DTW, où les 4 caractéristiques sont inclus pour que le premier niveau de la transformée en ondelettes devienne la deuxième ligne d'une matrice à dimensions 4×4 , et 4 caractéristiques de deuxième niveau DWT devienne la troisième ligne de la même matrice 4×4 , et de même pour le troisième niveau de la transformée en ondelettes pour devenir la quatrième ligne de la matrice, tandis que la première ligne de la matrice représente 4 caractères dans le domaine temporel, ce sont respectivement l'énergie du signal (E), le nombre de passage par zéro pour le signal (NPZ), le rapport de de passage par zéro (TPZ) pour un même signal, la (durée) de temps du signal audio, pour que le nombre total devienne 16 caractères, qui est le nombre d'éléments de la matrice A de dimension 4×4 , qui sont les suivants :

$$3-23 \quad A = \begin{bmatrix} E & \dots & \text{durée} \\ \vdots & \ddots & \vdots \\ & \dots & \end{bmatrix}_{4 \times 4}$$

Après avoir formé la matrice 4×4 des caractéristiques, nous calculons $\text{tr}(A)$ comme indiqué dans l'équation suivante :

Soit, $A = \text{tr}(a_{i,j}) \in M_n(k)$ on appelle trace de A, et en notre $\text{tr}(A)$ le scalaire $\sum_{i=1}^n a_{i,j}$.

$$3-24 \quad \text{tr}(A) = \sum_{i=1}^n a_{i,j}$$

$M_n(k)$: matrice carré a : élément de matrice A

11- processus de comparaison et de performance

Afin de connaître le taux de reconnaissance du locuteur, on calcule les deux relations (3-25) et (3-26) qui sont respectivement FAR (Taux de fausses acceptation), et FRR (Taux de faux rejet), que nous avons référencées au chapitre deux comme suit :

a- Taux de faux rejet :

FRR (False Reject Rate) : représente le pourcentage des tests rejetés par le système, Il est égale au nombre de faux rejets divisé par le nombre total des tests.

$$3-25 \quad FRR = \frac{\text{nombre faux rejets}}{\text{nombre de tests total}}$$

b- Taux de fausses acceptation :

FAR (False Acceptation Rate) : représente le pourcentage des tests accepter par le système , Il est égale au nombre de fausse acceptation divisé par le nombre total des tests.

$$3-26 \quad FAR = \frac{\text{nombre de fausses acceptations}}{\text{nombre de tests total}}$$

c- Taux de précision de système

Taux de précision de système (System Accuracy Rate) , c'est le taux d'erreurs égales un compromis entre les fausses acceptations et les faux rejets, autrement dit c'est le point de mesure sur lequel FAR = FRR . A partir de processus d'évaluation , et le pourcentage d'identification du locuteur se déroule selon la relation suivante (3-27):

3-27

$$SAR = 1 - \frac{\text{nombre de fausses acceptations} + \text{nombre faux rejets}}{\text{nombre de tests total}} * 100$$

Le tableau de comparaison (4-9) montre les résultats après les avoir appliqués à la même phrase " DO YOU SPEAK ENGLISH " qui a été précédemment utilisée dans toutes les étapes appliquées , pour les mêmes locuteurs des deux sexes , et pour les trois niveaux de transformée en ondelettes discret DWT . Le tableau résume l'application de ces relations au signal du locuteur, qui est celui d'une locuteur de sexe Femme .

III -9 Etude et discussion de la cinquième partie

Dans cette cinquième partie du troisième chapitre , nous effectuons le processus de cryptage du signal wav par tatouage numérique, en utilisant une nouvelle méthode pour le processus d'encodage utilisant des opérations mathématiques sur les matrices et leurs propriétés, y compris le calcul de trace pour la matrice dimension $4 * 4$, qui a été produite à partir des coefficients de prédiction linéaire extraits des caractéristiques du signal wav dans les trois niveaux de transformée discrète en ondelettes , et des caractéristiques du signal d'original dans le domaine temporel précédemment formé , et traité .

Dans la première étape avant l'application de watermark, nous appliquons l'algorithme des coefficients cepstraux prédictifs linéaires (LPCC) qui est une technique de traitement de la parole utilisée pour extraire des caractéristiques , et pour caractériser l'enveloppe spectrale des signaux paroles pour obtenir le vecteur de caractéristiques LPCC .

La prédiction linéaire consiste à modéliser un signal vocal comme une combinaison linéaire de ses valeurs passées. Ce vecteur de caractéristiques est couramment utilisé pour la reconnaissance vocale, la reconnaissance du locuteur et d'autres applications de traitement de la parole. Cela passe par les étapes suivantes :

Traitement du signal : Après échantillonnage du signal à une fréquence de 11,025kHz, qui est une fréquence 2 fois supérieure au signal de parole, qui est généralement égal à 5kHz.

Nombre d'échantillons :

$$\text{Nbrs éch} = 11025 \text{ hz} * 30 \text{ ms} = 11025 \text{ hz} * 0.030 \text{ s} = 330,75 \text{ éch} \approx 330 \text{ échantillons}$$

- Le processus de filtrage du signal wav avec un filtre Butterworth d'ordre $n=8$, et une fréquence de coupure égale à $f_c=600$ Hz.



Figure(3-18) : le processus de filtrage du signal

a : facteur de multiplication (0.97 : 1) [117].

$X(n)$: signal d'entrée ; $y(n)$: signal de sortie après traitement

-- Calcul de l'Énergie du signal original dans le domaine temporelle, et calcul la durée du signal wav, le passage par zéro, et le rapport de passage par zéro avec zéro.

L'énergie d'un signal est une mesure de la quantité d'énergie contenue dans le signal. Elle peut être utilisée pour caractériser un signal dans le domaine temporel. L'énergie d'un signal est calculée en intégrant le carré de l'amplitude du signal sur une période de temps donnée[119]. $x(t)$ est un signal dans le domaine temporel, l'énergie E du signal sur la période de temps T est définie comme suit[119] :

$$3-28 \quad E = \int_{-0}^{+T} |y(t)|^2$$

E : energie

$|X(t)|^2$: le carré de l'amplitude du signal à chaque instant t et l'intégrale est prise sur la période de temps T .

Il convient de noter que pour les signaux avec une énergie finie, comme la plupart des signaux physiques, l'énergie dans les deux domaines est finie et peut être calculée à l'aide du théorème de Parseval, qui indique que l'énergie dans le domaine temporel est égale à l'énergie dans le domaine de fréquence. Cependant, cela n'est généralement pas vrai pour les signaux avec une énergie infinie, comme le bruit blanc.

$$3-29 \quad E = \sum_{i=1}^n |Y(n)|^2$$

-- Application de la transformée en ondelettes discrètes à trois niveaux du signal étudié DWT, en utilisant l'ondelette mère

$$3-30 \quad a_1 = \sum_{k=-\infty}^{\infty} y(k)g[2n - k]$$

a_1 : coefficient approximatifs

-- Divisez le signal auquel la transformation en ondelettes discrète a été appliquée en courtes fenêtres de recouvrement de bien 75% de la taille de trame. , généralement la trame est de 10 à 30 ms de long multipliée par une fonction de fenêtre , telle que la fenêtre de Hamming pour réduire la perte spectrale .

$$3-31 \quad W(n) = 0.54 - 0.46 \cos\left(2\pi \frac{n}{N}\right) \quad 0 \leq n \leq N$$

-- Autocorrélation : L'idée de l'utilisation de la fonction d'autocorrélation est de déterminer à quel point deux échantillons successifs d'un signal se ressemblent . En calculant la fonction d'autocorrélation, nous pouvons identifier le degré auquel le signal est corrélé avec lui-même à différents décalages temporels . Si la fonction d'autocorrélation est élevée à un décalage particulier, cela suggère que le signal est fortement corrélé avec lui-même à ce décalage, indiquant un modèle ou une régularité dans les données qui se répètent à cet intervalle. D'autre part, si la fonction d'autocorrélation est faible ou proche de zéro, cela suggère qu'il n'y a pas de corrélation entre le signal et sa version retardée à ce décalage .

$$3-32 \quad \phi(k) = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=-N}^N y(n)y(n+k)$$

-- En calcule les coefficients lpcc , et nous prenons les quatre premières valeurs du vecteur LPCC pour les trois niveaux de la transformée en ondelettes discrète , et ceci pour former les trois lignes de la matrice de comparaison, que l'on appelle (A). qui sont les deuxième, troisième, et la quatrième ligne , Quant la première ligne, elle est composée des valeurs du signal dans le domaine temporel, et ce sont respectivement l'énergie , le nombre de passages par zéro , le rapport de passage par zéro , la durée du signal , Ces quatre lignes forment une matrice de 4*4 .

-- En convertir le signal wav après l'avoir traité, qui est la phrase "DO YOU SPEACK ENGLISH" prononcée par un locuteur , nous renvoie les données audio sous forme de vecteur y et le taux de fréquence d'échantillonnage égale de $F_e=11025$ hz .

-- Ensuite en mesuré la longueur de watermark , nous modifions la taille de la longueur du watermark à la même longueur que le signal wav , Convertir le watermark en valeurs -1/+1 , et nous modifions la taille de la longueur du watermark à la même longueur que le signal wav . Avant de l'ajouter au signal wav pour réduire le risque que le watermark provoque une distorsion audible , nous multiplions watermark par un facteur de 0,01, et par la trace (A) de matrice qui nous avons calculer .

Calcule la matrice (A) de taille (4*4) :

3-33

$$\mathbf{A} = \begin{vmatrix} E & NPZ & TPZ & dr \\ a_{1NV1} & a_{2NV1} & a_{3NV1} & a_{4NV1} \\ a_{1NV2} & a_{2NV2} & a_{3NV2} & a_{4NV2} \\ a_{1NV3} & a_{2NV3} & a_{3NV3} & a_{4NV3} \end{vmatrix}$$

-- La trace : Nous calculons la trace de matrice (A) défini par l'équation () pour une utilisation ultérieure dans le codage du signal.

Calcule la trace de matrice (A) :

Soit $A = (a_{i,j}) \in M(k)_n$ donc la trace de matrice (A) le scalaire noté $\text{Tr}(A)$ est :

3- 34

$$\text{T}(A) = \sum_{i=1}^n a_{i,i}$$

Danc $\text{Tr}(A) = a_{11} + a_{22} + a_{33} + a_{44} = \text{Tr}(A) = a_{1NV1} + a_{2NV2} + a_{3NV3} + a_{4NV4}$

III-10 Watermark

La watermark est un outil efficace pour encoder le signal de toute nature, que ce soit dans le domaine du son ou de l'image. La watermark peut être un moyen d'identifier le propriétaire du contenu et de le l'indiquer . il Ya plusieurs type de watermark varie également selon la durabilité et l'utilisation pour contrer les attaques et les types de menaces . En ce qui concerne la watermark utilisé dans notre recherche , il s'agit d'une phrase de texte comme image carrée de type gris avec des dimensions (128*128*3) converti en une matrice binaire , nous pouvons également le changer avec un autre watermark plus grand ou différent selon nous besoins , afin de l'utiliser plus tard avec un signal de parole .

Ensuit :

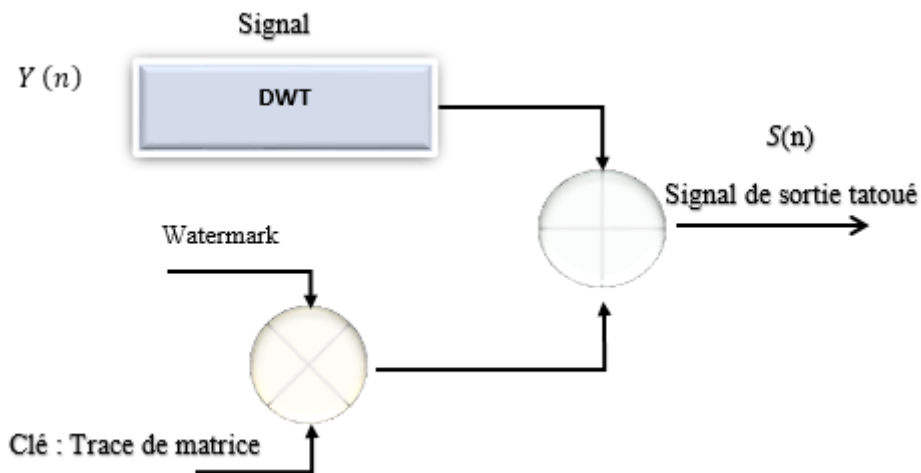
3-35
$$e(n) = 0.01 * w_{\text{watermark}} * \sum_{i=1}^n a_{i,i}$$

$e(n)$: signal tatoué normalisé avec facteur = 0.01

$w_{\text{watermark}}$: vecteur le watermark

3-36
$$S(n) = (0.01 * w_{\text{watermark}} * \sum_{i=1}^n a_{i,i}) + Y(n)$$

Enfin, nous combinons le signal wav , et le watermark , et écrivons le signal combiné résultant dans un nouveau vecteur (fichier WAV) comme présenter dans la figure (3-19) . IL existe plusieurs façons d'intégrer un watermark dans un signal wav , dont certaines peuvent être plus sécurité que d'autres .



Figure(3-19) : présente le système watermark

III-11 Conclusion de chapitre troisième

Dans ce chapitre appliqué, nous avons présenté cinq parties consistant en de nouvelles et diverses méthodes séquentielles pour ajouter un watermark à un signal de parole d'un locuteur des deux sexes et le comparer avec d'autres locuteurs, et ce dans le but d'assurer une protection adéquate de ce signal lors de la vérification et de la identification. Dans la première partie de ce chapitre troisième, nous avons introduit une nouvelle approche pour trouver Les meilleures ondelettes mères, et leurs meilleures branches conservent les caractéristiques du signal wav étudié. Comme pour la troisième partie de ce chapitre, nous avons présenté une nouvelle approche, et un nouveau critère de détermination du sexe du locuteur. Il s'agit d'un critère similaire à la fréquence fondamentale f_0 , et il peut déterminer le sexe du locuteur sur la base d'une plage de valeurs de critère qui nous avons appelé (WER) << taux d'énergie d'ondelettes >>. Le but de ce critère est de réduire le temps de comparaison entre les locuteurs, car au lieu de comparer le signal avec tous les signaux, la comparaison est limitée à une catégorie des deux sexes.

Nous avons également présenté dans ce chapitre, dans sa quatrième partie, une nouvelle approche qui s'appuie sur les opérations mathématiques des matrices et leurs caractéristiques comme la trace de matrice pour vérifier et identifie le locuteur, c'est une nouvelle méthode difficile à pénétrer et à détecter par les attaquants, et cela est dû à la difficulté de déterminer sa matrice de base utilisée dans le processus de vérification et de identification du locuteur.

Où nous avons développé une nouvelle façon de sécuriser le signal et de le protéger de toutes les attaquants, en utilisant le watermark et des matrices numériques et la trace de matrice pour sécuriser ce signal du fait de la complexité des calculs mathématiques utilisés en sécurisation de signal avec watermark.

CHAPITRE (IV)
RESULTATS et DISCUSSIONS

IV Introduction de chapitre :

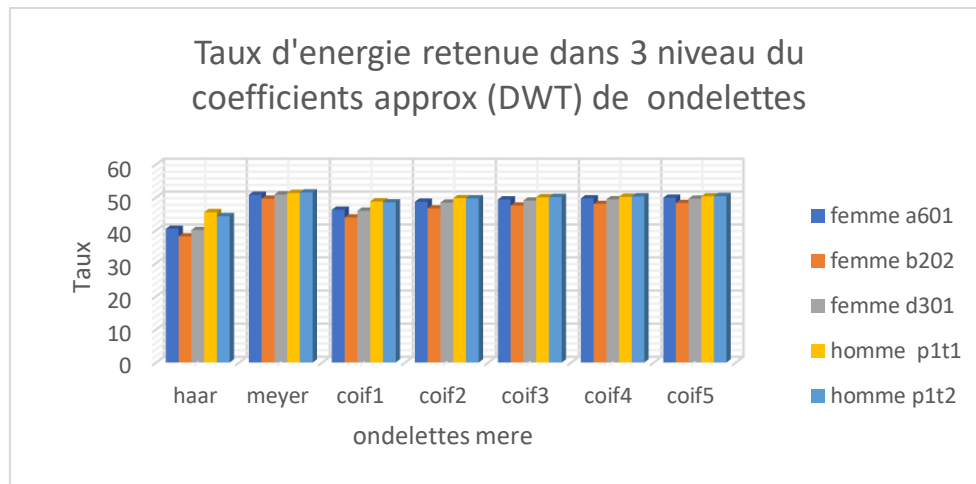
Le quatrième chapitre présente les résultats obtenus à partir d'expériences et d'approches des cinq parties qui composent le troisième chapitre, qui ont été appliquées à différents signaux audio pour différents locuteurs de sexes différents, qui sont l'objet et le but de la recherche. Il présente également des tableaux, des figures et des graphiques pour toutes les étapes de la recherche. Ces résultats sont dans l'ordre comme suit :

IV -2 Résultats de la première partie :

Lors de l'application, et l'étude de la troisième niveau de la transformée en ondelettes discrètes (DWT), en utilisant les fonctions de nombreuses ondelettes mères dans leurs différentes branches, nous avons extrait les résultats selon les tableaux, et les données ci-dessous, à travers 570 expériences sur la transformée en ondelettes discrète (DWT), pour trois niveaux du signal vocal (wav). Cette phrase est "Do you speak English", a été prononcée par des hommes, et des femmes, et de nombre de 5 signaux. L'accent a été mis sur le troisième niveau et appliqué à 6 ondelettes mères à 38 branches. La même phrase a été choisie pour que l'on puisse remarquer la différence entre les locuteurs. Après l'application du DWT, on obtient les tableaux, et graphiques suivants ont été extraits .

	ondelette (coiflets)						
	Haar	meyer	coif1	coif2	coif3	coif4	coif5
femme a601	40,4373	50,7579	46,2438	48,6751	49,3566	49,6957	49,8980
femme b202	38,1499	49,5898	43,885	46,6369	47,5175	47,9907	48,2760
femme d301	40,0709	50,8582	45,9261	48,3948	49,0162	49,3896	49,5960
homme p1t1	45,494	51,328	48,7443	49,7414	50,024	50,1741	50,2880
homme p1t2	44,2925	51,4776	48,4763	49,6875	50,1015	50,2762	50,3890

Tableau N° (4-1) : Comparaison des ondelettes de la mère (coiflet , meyer , haar)



figure(4-1) : Comparaison des ondelettes de la mère(coiflet , meyer, haar)

Tableau n° (4-2) : représente une comparaison des valeurs du taux d'énergie retenue après application de la transformée en ondelettes discrète sur le signal de parole de 3 femmes , et 2 hommes, et ceci pour la même phrase à travers l'ondelette mère (haar et meyer) , et l'ondelette mère (coiflet) avec ses 5 branches (coif1, coif2, coif3, coif4, coif5).

Après la comparaison, il a été constaté que le pourcentage d'énergie le plus élevé se situe au niveau de la fonction d'ondelette mère (Mayer), que ce soit chez les hommes ou chez les femmes, et cela est également illustré par la figure (4-2) .

	ondelette (symlets)							
	sym1	sym2	sym3	sym4	sym5	sym6	sym7	sym8
femme a601	40,4373	46,0673	46,0673	48,5393	48,9584	49,2333	49,4272	49,5726
femme b202	38,1499	43,7146	45,5327	46,4352	47,0071	47,3554	47,6421	47,8377
femme d301	40,0709	45,7798	47,4704	48,2163	48,6234	48,9111	49,0651	49,2438
homme p1t1	45,494	48,6484	49,3688	49,6623	49,8231	49,9126	49,9913	50,0542
homme p1t2	44,2925	48,3537	49,3167	49,6942	49,874	49,996	50,0815	50,1471

Tableau (4-2) : Comparaison entre les 8 branches de l'ondelette mère (symlet) application (DWT) au troisième niveau

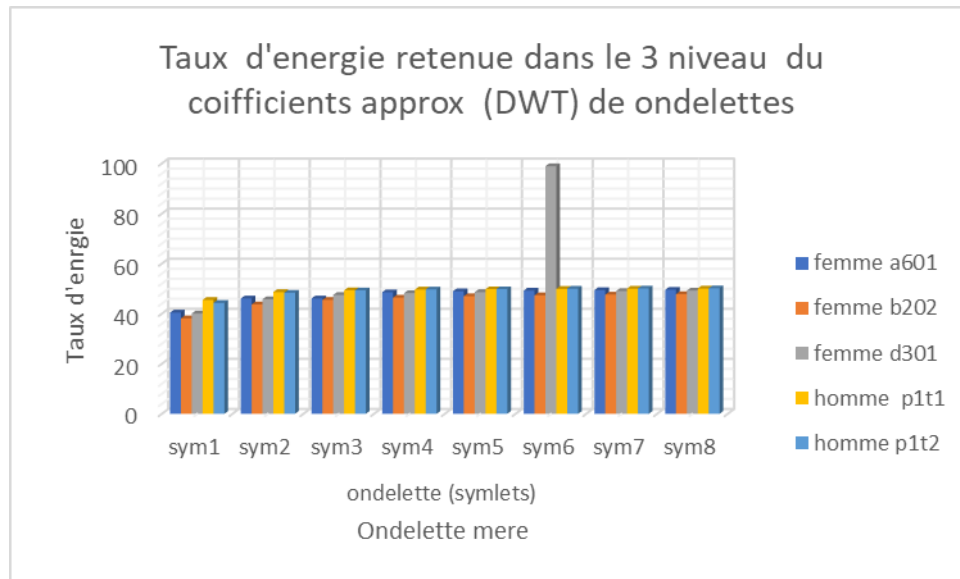


figure (4-2) : Comparaison entre les 8 branches de l'ondelette mère (symlet) application (DWT) au troisième niveau

Tableau n° (4-3) : représente une comparaison des valeurs du pourcentage d'énergie retenue après application de la transformée en ondelettes discrète (DWT) sur le signal de parole de 3 femmes , et 2 hommes, et ceci pour la même phrase à travers l'ondelette mère (symlet) avec ses 8 branches (sym1, sym2, sym3,, sym8).

Après la comparaison, il a été constaté que le pourcentage d'énergie le plus élevé se situe au niveau de la fonction d'ondelette mère (sym8), que ce soit chez les hommes ou chez les femmes, et cela est également illustré par la figure (4-3) .

	ondelette (biorthogonal)							
	bior1,3	bior1,5	bior2,8	bior3,1	bior3,5	bior3,9	bior4,4	bior5,5
femme a601	52,5461	56,629	75,4599	83,2203	94,7508	98,9292	42,3131	27,3732
femme b202	50,1448	54,6487	77,5758	87,9758	102,8502	108,7093	41,0657	25,8135
femme d301	52,3396	56,4064	75,5836	84,638	96,1609	100,6246	42,0043	26,8250
homme p1t1	52,173	58,8573	61,5874	65,1766	69,2421	70,5172	45,5213	37,1874
homme p1t2	52,8991	55,1263	64,8149	14,0105	74,2075	75,5755	44,3996	33,8534

Tableau (4-3) : Comparaison entre les 10 branches de l'ondelette mère (biorthogonal) application (DWT) au troisième niveau

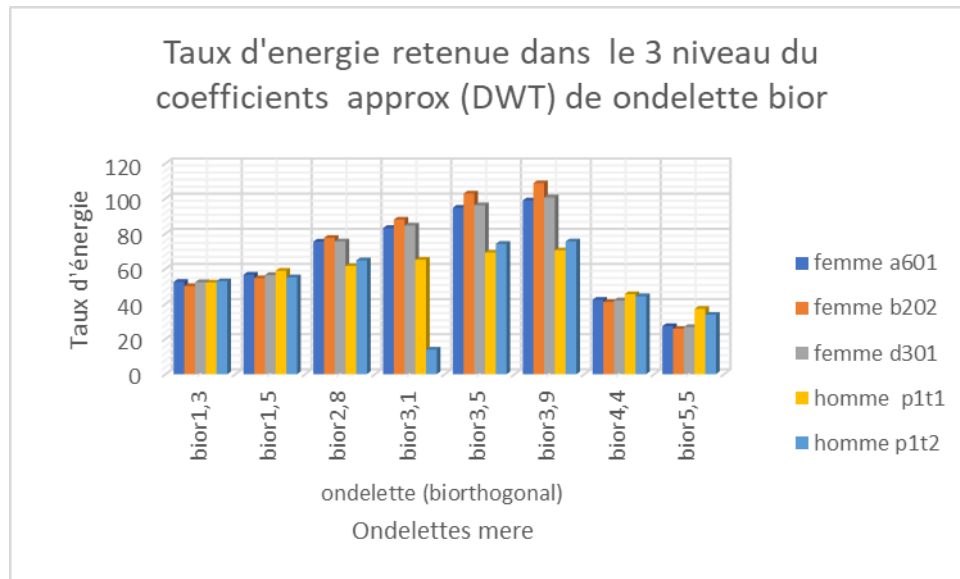


figure (4-3) : Comparaison entre les 10 branches de l'ondelette mère (biorthogonal) application (DWT) au troisième niveau

Tableau n° (4-4) : représente une comparaison des valeurs du pourcentage d'énergie retenue après application de la transformée en ondelettes discrète sur le signal de parole de 3 femmes , et 2 hommes, et ceci pour la même phrase à travers l'ondelette mère (biorthogonal) avec ses 8 branches (bior 1.3, bior 1.5, bior 2.8, bior 3.1, bior 3.9,, bior 5.5).

Après la comparaison, il a été constaté que le pourcentage d'énergie le plus élevé se situe au niveau de la fonction d'ondelette mère ((bior 3.9), que ce soit chez les hommes ou chez les femmes, et cela est également illustré par la figure (4-4) .

	ondelette (daubouchies)							
	db2	db3	db5	db7	db9	db20	db30	db45
femme a601	46,0637	47,7728	48,9584	49,4272	49,6777	50,2068	50,4120	50,7355
femme b202	43,7146	45,5327	46,9854	47,6269	47,9907	48,5026	49,1776	49,6380
femme d301	45,7798	47,4704	48,6229	49,0617	49,336	50,0654	50,4242	50,8671
homme p1t1	48,6484	49,3688	49,8227	49,9911	50,1084	50,4934	50,7677	51,2891
homme p1t2	48,3537	49,3167	49,874	50,0815	50,1843	50,5643	50,8538	51,6301

Tableau (4-5) : Comparaison entre les 45 branches de l'ondelette mère (Daubouchies) application (DWT) au troisième niveau

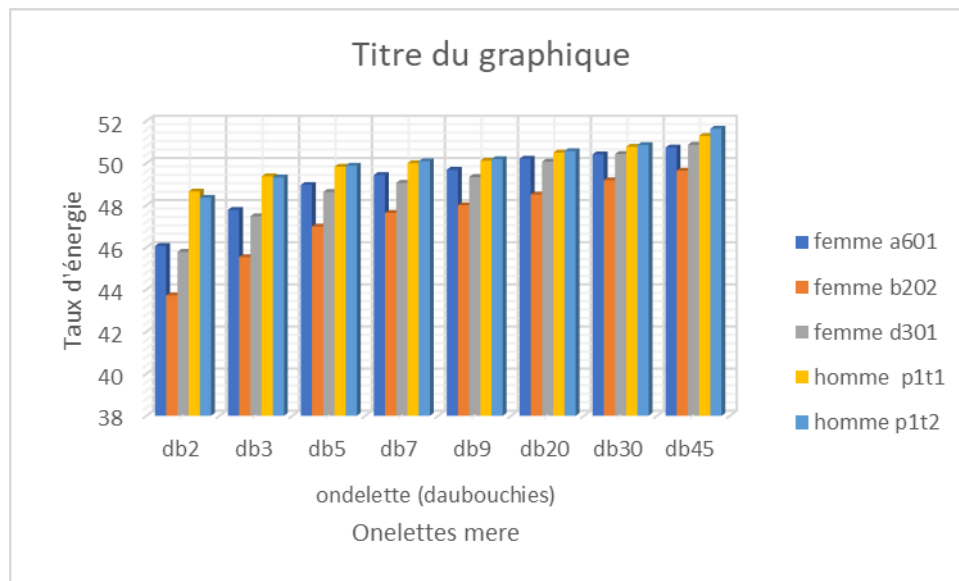


figure (4-5) : Comparaison entre les 45 branches de l'ondelette mère (Daubouchies) application (DWT) au troisième niveau

Tableau n° (4-5) : représente une comparaison des valeurs du pourcentage d'énergie retenue après application de la transformée en ondelettes discrète (DWT) sur le signal de parole de 3 femmes, et 2 hommes, et ceci pour la même phrase à travers l'ondelette mère (haar et mayer), et l'ondelette mère (Daubouchies) avec ses 45 branches (db1, db2, db3, db7, ,db45).

Après la comparaison, il a été constaté que le pourcentage d'énergie le plus élevé se situe au niveau de la fonction d'ondelette mère (db45), que ce soit chez les hommes ou chez les femmes, et cela est également illustré par la figure (4-5).

	comparaison des ondelettes					
	haar	meyer	biorthogo	coiflets	symlets	daubouchies
femme a601	40,4373	50,7579	97,2217	49,8978	49,5726	50,7355
femme b202	38,1499	49,5898	106,1706	48,2755	47,8377	49,638
femme d301	40,0709	50,8582	98,6206	49,596	49,2438	50,8611
homme p1t1	45,494	51,328	70,0121	50,2885	50,0542	51,2891
homme p1t2	44,2925	51,4776	75,5755	50,3898	50,1471	51,3602

Tableau (4-6) : Comparaison entre les 6 ondelettes mères et leurs 38 branches après application (DWT) au troisième niveau

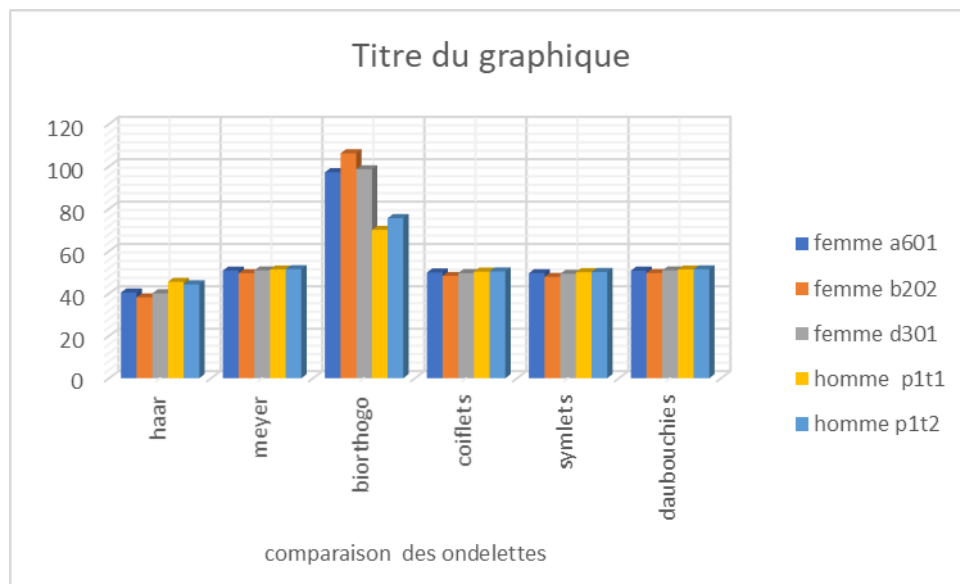


figure (4-6) : Comparaison entre les 6 ondelettes mères et leurs 38 branches après application (DWT) au troisième niveau

Tableau (4-6) : représente une comparaison des valeurs du pourcentage d'énergie retenue après application de la transformée en ondelettes discrète (DWT) sur le signal de parole de 3 femmes , et 2 hommes, et ceci pour la même phrase à travers les 6 ondelettes mères , et leur 38 branches

..

Après la comparaison, il a été constaté que le pourcentage d'énergie le plus élevé se situe au niveau de la fonction d'ondelette mère (biorthogonal), que ce soit chez les hommes ou chez les femmes, et cela est également illustré par la figure (4-6) , suivre par d'ondelette mère (Daubouchies).

	ECG	femme a601	femme b201	femme d301	homme p1t1	homme p1t2
coif5	26	49,898	48,276	49,596	50,288	50,389
sym8	57	49,5726	47,8377	49,2438	50,0542	50,1471
db2	57	46,0637	43,7146	45,7798	48,6484	48,3537
bior3,1	100	83,2203	87,9758	84,638	65,1766	14,0105
bior3,9		98,9292	100	100	70,5172	75,5755

Tableau (4-7) : Comparaison entre les 6 ondelettes mères et leurs 38 branches , et le signal ECG après application (DWT) au troisième niveau,

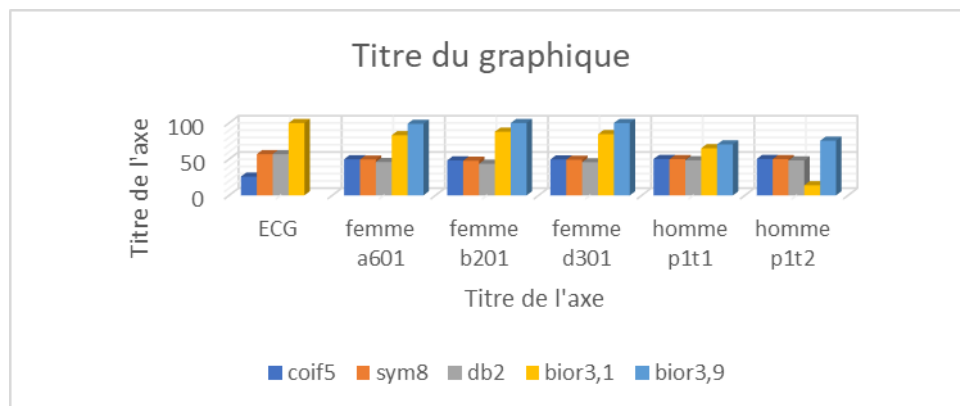


Tableau (4-7) : Comparaison entre les 6 ondelettes mères et leurs 38 branches, et le signal ,ECG après application (DWT) au troisième niveau

Tableau n° (4-7) : représente une comparaison des valeurs du pourcentage d'énergie retenue après application de la transformée en ondelettes discrète sur le signal de parole de 3 femmes , et 2 hommes, de la même phrase à travers les 6 l'ondelettes mères ,et ses 38 branches , et le signal ECG .

Après la comparaison, il a été constaté que le pourcentage d'énergie le plus élevé se situe au niveau de la fonction d'ondelette mère (bior3.7), que ce soit chez les hommes ou chez les femmes, ou de signal ECG , et cela est également illustré par la figure (4-7) .

Discussion

Après la calcul de l'énergie du signal d'origine $E_{X(n)}$, et l'énergie du signal des coefficients approximatifs dans le troisième niveau de la transformée en ondelettes discrète DWT E_{a3} pour tous les ondelettes mère , on calcule le taux d'énergie retenu E_{ra3} dans le domaine temporel selon le critère exprimé dans l'Equation (3-5) ,

Nous concluons de cette étude que la branche (bior3.9) de l'ondelette mère (biorthogonale) , est la meilleure pour extraire les caractéristiques du signal de parole , et retenir l'énergie dans le signal. Cette conclusion est basée sur l'analyse , et la comparaison des résultats obtenus à partir des 38 branches différentes, et des 6 ondelettes mères, le critère est considéré comme le meilleur , et le moins d'opérations de calcul pour extraire les caractéristiques du signal de parole , et le plus de rétention d'énergie dans le signal après analyse , selon les tableaux , et les figures . Le tableau de comparaison (4-6) , et la figure (4-6) montrent la différence entre les quantités d'énergies retenues à ce niveau d'analyse entre les ondelettes mères , et leurs branches ,En comparant les résultats de la recherche (Khadra, 2016) pour choisir l'ondelettes mère la plus appropriée dans son application au signal ECG , nous avons constaté que le critère que nous avons utilisé donnait de bons résultats , et des taux élevés pour tous les signaux , et pour les différentes branches . le tablau (4-7) , et la figure (4-7) montrant la comparaison de nos résultats avec les résultats de la recherche susmentionnée, nous disons donc que l'onde mère (biorthogonale) , est la meilleure pour analyser de signal pour tous les niveaux d'analyse , et la branche (bior3.9) , est la meilleure pour tous les signaux.

IV -3 Résultat de la deuxième partie

Comparaison le taux d'énergie d'un signal du parole de niveau troisième à l'aide de la transformée en ondelettes discrète (DWT) de l'ondelette mère Haar , meyer , biorthogonal , coiflets , symlets , daubechies

1- Introduction

Le signal de parole, est un signal non stationnaire, complexe et variable dans le temps [27]. De plus, la fréquence la plus élevée du signal de parole, est de 5 kHz. Les différences dans les caractéristiques linguistiques des êtres humains sont caractérisées par la variance des énergies pour les deux sexes . généralement le signal parole à trois caractéristiques Hauteur , Fréquence fondamentale , Energie .cette dernière , est sont petites chez les femmes par rapport les hommes .

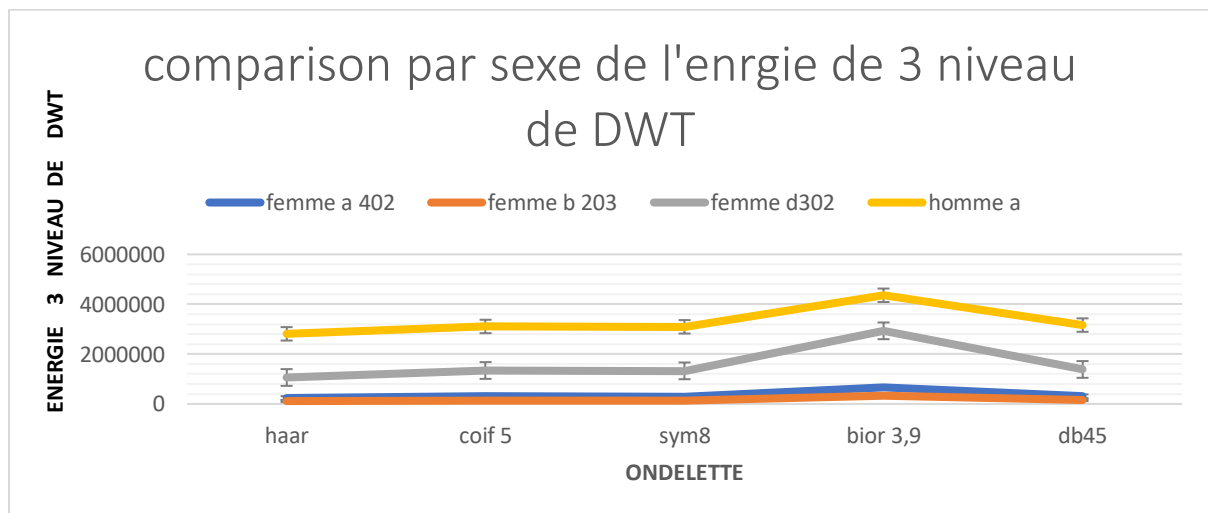


Figure (4-8) : comparaison globale d'une phrase « « do you speak english » » de sexe 3 femmes (a ,b ,d) on utilisant DWT

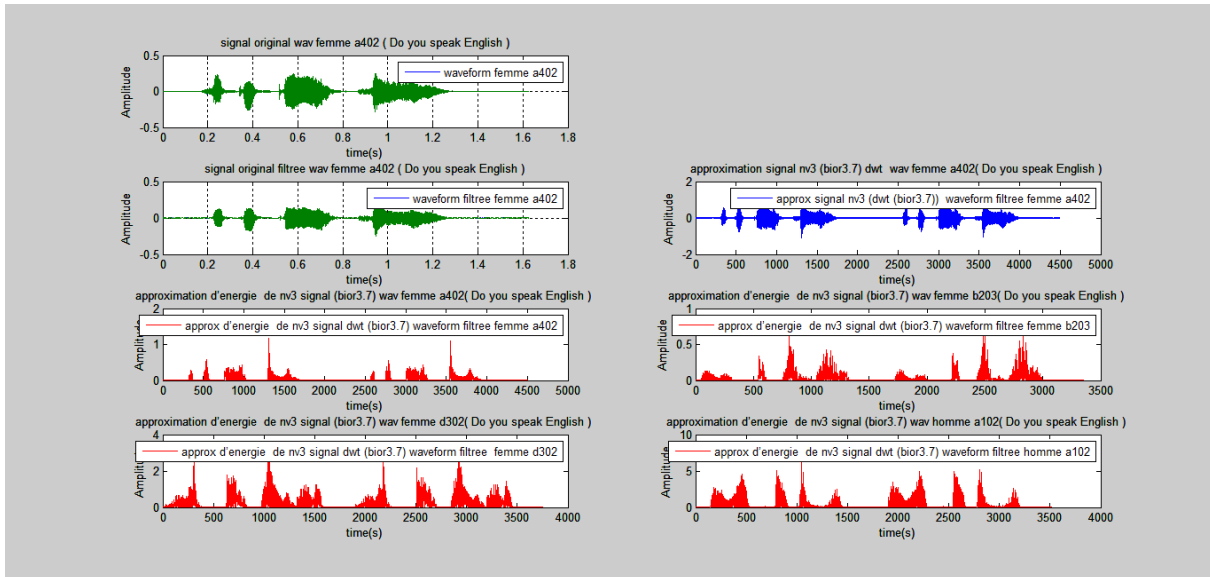


Figure (4-9) : comparaison globale d'une phrase « « do you speak english » » de sexe 3 femmes (a ,b ,d) on utilisant DWT

Discussion

A travers les expériences suivantes, nous remarquons que la plus grande énergie dans le troisième niveau de la transformée en ondelettes discret (DWT) pour les ondelettes mères suivantes, Haar, meyer, biorthogonal, coiflets, symlets, daubechies, varie d'une ondelette à l'autre, et de branche à l'autre dans les ondelettes, nous avons remarqué que la plus grande énergie dans la branche 8 chez les hommes dans la transformée en ondelettes discret (DWT) de l'ondelette mère symlet, Pour les deux sexes, par rapport tous les branches de deux sexes hommes, et femmes, de la même phrase, tandis que la plus élevée l'énergie dans les coiflets est dans la branche 5 chez les hommes par rapport aux autres branches des deux sexes, pour la même phrase, et la même DWT de troisième niveau. d'autre part l'énergie dans l'ondelette mère biorthogonale, est plus élevée chez les hommes dans la branche (bior3.9) par rapport au reste des branches Pour les deux sexes, et pour l'énergie la plus élevée dans le l'ondelette mère daubechies, est dans la branche 45 chez les hommes par rapport aux autres branches pour les deux sexes, de la même phrase, et le même niveau troisième de DWT. à travers le tableau de comparaison (4-7), nous concluons que l'énergie chez les hommes, est supérieure à celle chez les femmes dans le troisième niveau de la transformée en ondelettes discrète (DWT) de toutes les branches, et les ondelettes mères, et ceci Après avoir comparé la même phrase pour un groupe de femmes, et d'hommes.

Nous pouvons dire, que la branche bior3.9, est la plus élevée en énergie pour le signal parole, après le filtrage du signal par un filtre Butterworth degré $n=8$, pour la phrase, et nous mentionnons que le calcul d'énergie a été fait dans le domaine temporel du signal. et le signal parole est 'une phrase " Do you speak English" '.

IV -4 Résultat de la troisième partie

1- Discussion et résultats

À travers des tableaux , et les figures , qui sont les résultats expérimentaux de 200 expériences sur des signaux du parole , pour des phrases répétées , par différentes locuteurs femmes et hommes , après filtrage du signal de parole original à l'aide d'un filtre Butterworth degré $n = 8$, et de fréquence d'échantillonnage $f_e = 11025\text{hz}$, et après avoir calculé l'énergie de la branche (db7) pour le premier niveau de la transformée en ondelettes discrètes DWT de l'ondelette mère , et le calcul de nombre de passage par zéro du signal d'original , selon l'équation (3-8) , et l'équation (09), les résultats sont compilés dans le tableau (4-8) ci-dessous.

résultat de test d'une phrase répétée pour la même personne , et plusieurs personnes (4 femmes et 3 hommes) , la phrase est « « do you speak english » »							
	signal	Samples Number	Energy signal	Energy E_{a1}	ZCN signal	ZCN_{a1}	ZCR_{a1}
femmes	femme A 601	22243	879380	28077000	1268	4405,3	0,0570
	Femme A 602	21060	833610	25851000	1224	4282,3	0,0581
	Femme B 201	12487	271940	9406900	740	2623,8	0,0592
	Femme B 203	12820	275960	10579000	811	2874,1	0,0632
	Femme D 301	12449	1795300	9581200	768	2701,7	0,0617
	Femme D 302	12793	2313000	10444000	815	2885,4	0,0637
	Femme M 1000	15493	577710	11501000	723	2600,5	0,0467
	Femme M 1003	15988	583980	12943000	795	2833,3	0,0497
hommes	homme B 101	13292	5851300	6844900	526	1790,1	0,0396
	Homme B 102	13667	6016100	7346700	538	1864,2	0,0394
	Homme K 100	11341	317560	4442000	391	1371,3	0,0345
	Homme W 400	11685	773940	4254200	364	1285,6	0,0311
	Homme W 401	11638	425690	4075200	354	1245,3	0,0304

Tableau (4-8) : résultat de test d'une phrase répétée pour la même personne , et plusieurs personnes (4 femmes et 3 hommes) , la phrase est « « do you speak english » »

ZCN signal: Nombre de passage par zero du signal

E_{a1} : Energie de coefficient approximatif de première niveau (DWT)

ZCN_{a1} : Nombre de passage par zero de coefficient approximatif de première niveau (DWT)

ZCR_{a1} : Taux de passage par zero de coefficient approximatif de première niveau (DWT)

À travers le tableau (4-8) et le tableau (4-9), il a été constaté que l'énergie du signal d'original varie dans les deux sexes avec des valeurs différentes. L'énergie du signal d'original pour les hommes est beaucoup plus grande que l'énergie pour les femmes, et c'est pour la même phrase répétée, ou plusieurs phrases différentes pour la même femme, ou le même homme. Aussi, en utilisant la même phrase pour plusieurs femmes, et plusieurs hommes, comme le montre la Figure(4-10) :

4-1

$$E_{SH} \gg E_{SF}$$

E_{SH} : l'énergie du signal original pour les hommes

E_{SF} : l'énergie du signal original pour les femmes

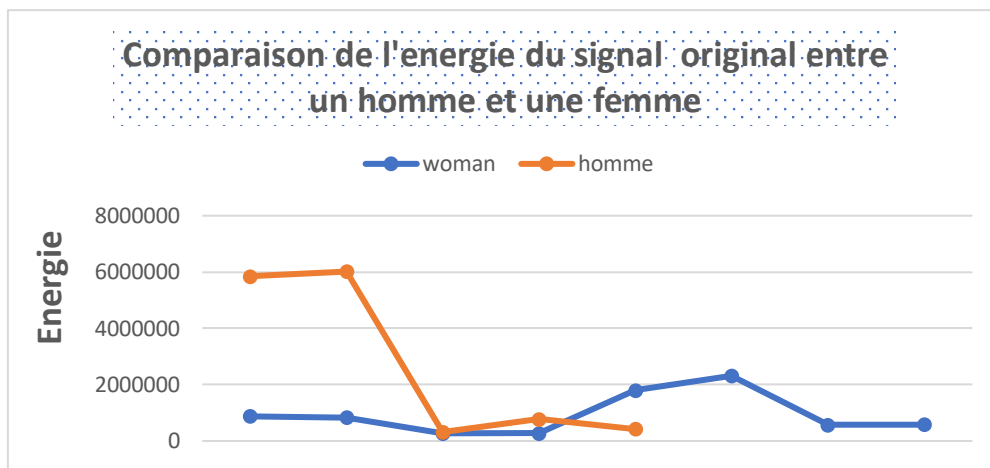


Figure (4-10) : Comparaison de l'énergie du signal original entre un homme et un femme

Mais dans le signal du coefficient approximatif a_1 , dans le premier niveau de la DWT, on trouve que l'énergie E_{a1} pour les femmes est supérieure à l'énergie pour les hommes, et ceci pour la même phrase répétée, ou plusieurs phrases différentes, pour la même femme, ou plusieurs femmes, ou le même homme, ou plusieurs hommes, comme indiqué dans le Tableau(4-8), Tableau (4-9), le Figure (4-11) présente cette comparaison.

4-2

$$E_{a1H} \gg E_{a1F}$$

E_{a1h} : l'énergie du coefficient approximatif a_1 pour les hommes

E_{a1f} : l'énergie du coefficient approximatif a_1 pour les femmes

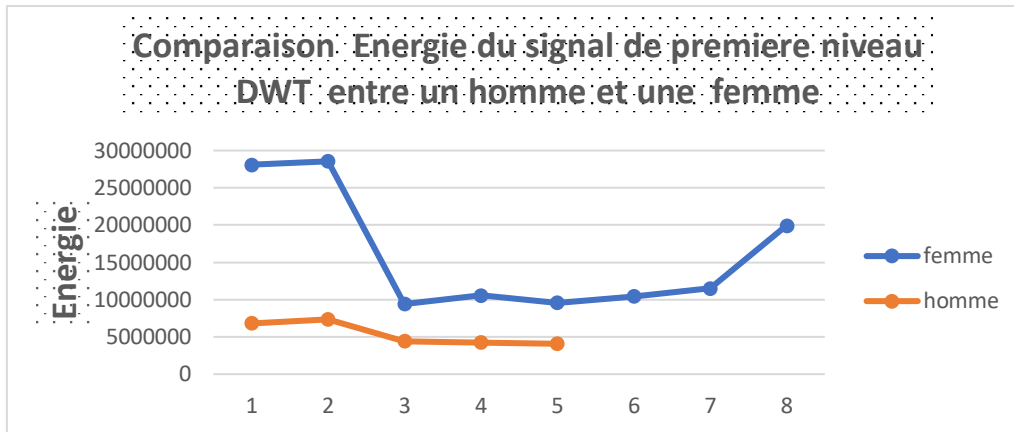


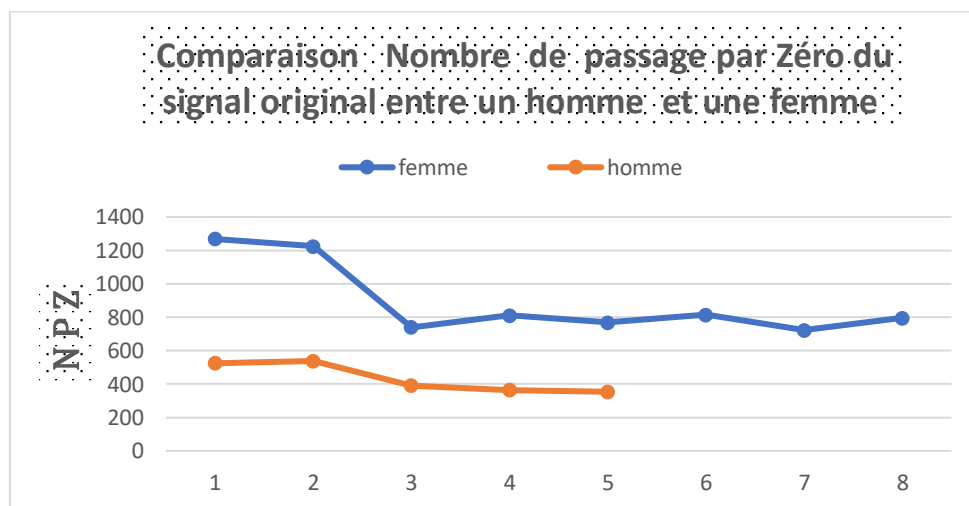
Figure (4-11) : Comparaison de l'énergie du signal de premier niveau DWT entre un homme et une femme

Il est à noter que le nombre de passage par zéro (NPZ) ou(ZCN) dans le signal d'origine diffère dans les deux sexes , et avec des valeurs diverses. Le passage par zéro du signal d'origine pour les hommes est beaucoup plus grande que le passage par zéro des femmes, et ceci pour la même phrase répétée , ou plusieurs phrases différentes , pour la même femme , ou plusieurs femmes , ou le même homme , ou plusieurs hommes , comme indiqué dans le tableau (4-8), tableau (4-9) , le Figure (4-12) présente cette comparaison .

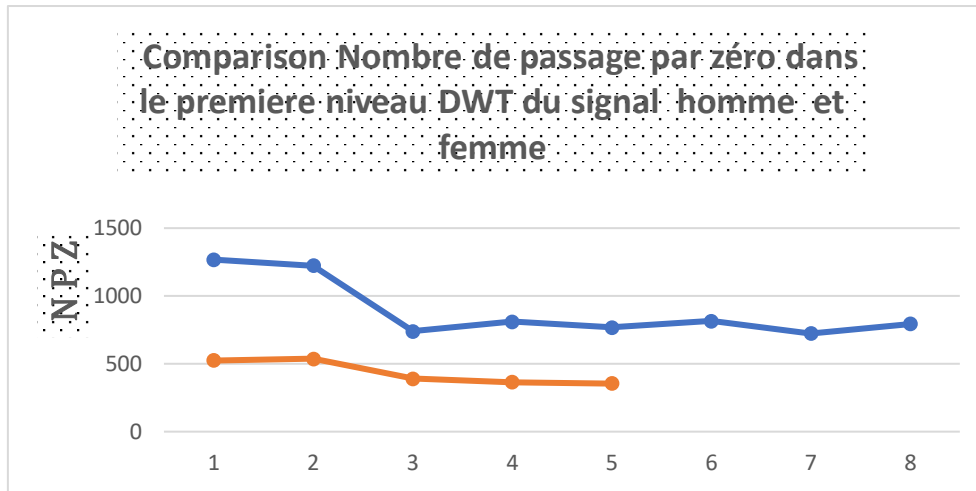
$$4-3 \quad npz_{SF} \gg npz_{SH}$$

Nnombre passage par zero de signal Homme : npz_{SH}

Nombre passage par zero de signal femme : npz_{SF}



Figure(4-12) : Comparaison du nombre de passages à zéro de signal original entre un 'homme et une femme



Figure(4-13) : Comparaison du nombre de passages à zéro dans le premier niveau DWT pour le signal de l'homme et de la femme

Pour le signal du coefficient approximatif a_1 dans le premier niveau de la DWT, on constate que le nombre de passage par NPZ_{a_1} , et le taux de passage par zéro (TPZ) ou (ZCR) pour les femmes est supérieur à celui pour les hommes, et ceci pour la même phrase répétée, ou plusieurs phrases différentes, pour la même femme, ou plusieurs femmes, ou le même homme, ou plusieurs hommes, comme indiqué dans le Tableau (4-8), Tableau (4-9) et Figure (4-13). Dans la Figure (4-14), nous concluons que le rapport passage par zéro pour un même niveau est plus important pour les femmes que pour les hommes.

4-4

$$npz_{a1F} \gg npz_{a1H}$$

npz_{a1H} : Nombre de passages à zéro dans le premier niveau DWT pour le signal de l'homme

npz_{a1F} : Nombre de passages à zéro dans le premier niveau DWT pour le signal de femme

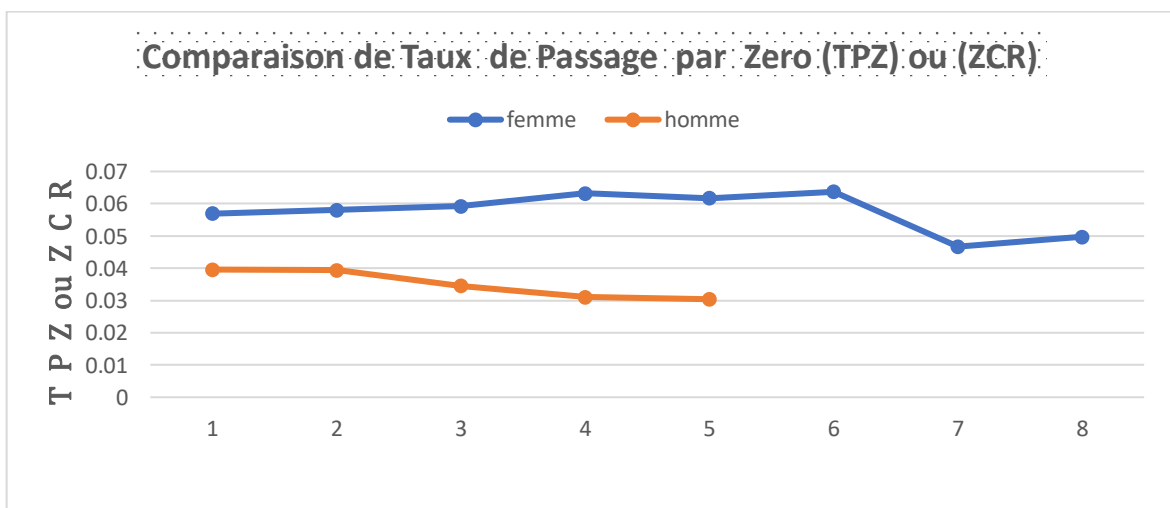


Figure (4-14) : Comparaison du taux de passage à zéro dans le premier niveau DWT pour le signal de l'homme et de la femme

Signal	Samples Number	Energy signal	Energy signal E_{a1}	ZCN_{a1}	ZCR_{a1}	WER
4 Women and 2 Men with the same sentence repeated 3 times: "do you speak english"						
femme A 601	22243	879380	28077000	4405.3	0.0570	363.3295
femme A 602	21060	833610	25851000	4282.3	0.0581	350.8538
femme B 201	12487	271940	9406900	2623.8	0.0592	212.4698
femme B 202	12645	274830	9855600	2715.3	0.0606	219.8762
femme D 301	12449	1795300	9581200	2701.7	0.0617	218.7851
femme D 302	12793	2313000	10444000	2885.4	0.0637	230.6052
femme m1000	15493	577710	11501000	2600.5	0.0467	206.3805
femme m1003	15988	583980	12943000	2833.3	0.0497	227.1477
Homme B 101	13292	5851300	6844900	1790.1	0.0396	151.3138
homme B 104	13666	6009400	7216300	1827.3	0.0391	154.3159
homme k 100	3411	317560	4442000	1373.3	0.0345	111.5123
homme w 400	11685	773940	4254200	1285.6	0.0311	103.0833
homme w 401	11638	425690	4075200	1245.3	0.0304	99.5563
4 Women with the same sentence: "see you later"						
femme A 2001	13595	420980	12289000	3129.3	0.0664	260.8490
femme B 2003	9503	311870	6561400	2401.4	0.0721	197.2349
femme D 402	5277	5571700	1748100	1171.4	0.0627	93.6040
femme m 3002	10365	322280	5379100	1826.2	0.0506	149.1963
homme C 201	12686	5501500	7000400	1923.6	0.0436	158.6357
Women and Men with different sentences						
femme A 401	20629	699960	24785000	4187.6	0.0583	345.1550
femme A 901	12242	306010	9660000	2726	0.0649	230.1245
femme A 803	9077	237420	5131100	1984.3	0.0622	160.9541
femme A 202	13809	374400	001	3059.5	0.0632	250.4241
femme A 101	12444	388450	10553000	2946.0	0.0676	242.3755
femme B 402	12880	486000	8577500	2319.3	0.0497	183.7694
femme B 501	8002	132600	5120800	2172.6	0.0813	191.7493
femme B 802	20612	662150	23670000	3982.7	0.0552	328.4199
femme B 901	11119	258280	7549900	2346.7	0.0620	199.6459
femme B 3003	13112	247580	11385000	3006.3	0.0665	251.8563
femme D 502	10428	2916500	7576900	2519.8	0.0696	209.3417
femme D 202	11089	1078200	7591400	2390.4	0.0617	195.8928
femme D 001	12320	1556200	9288400	2666.1	0.0613	213.5040
homme B 501	13160	2597500	8359000	2153.7	0.0487	189.0446
homme B 904	10492	2106500	3325500	1081.8	0.0299	92.0017
homme B 144	11025	2578600	5503800	1718.6	0.0455	145.8184
homme B 1004	10227	2481400	4480800	1520.2	0.0434	127.9606
homme B193	10998	2365600	5382100	1678.9	0.0446	143.1177
homme X102	8405	391600	5291400	2182.7	0.0748	181.4226
homme X 100	14152	1428800	13458000	3296.2	0.0668	272.9188
homme C100	11710	2518300	3975700	1196.7	0.0292	97.0313

Tableau(4-9) : Comparaison des valeurs (wer) de différents locuteurs hommes et femmes avec la même ,et différents phrases

ZCN_{a1} signal: Nombre de passage par zero du signal

E_{a1} : Energie de coefficient approximatif de première niveau (DWT)

ZCN_{a1} : Nombre de passage par zero de coefficient approximatif de première niveau (DWT)

ZCR_{a1} : Taux de passage par zero de coefficient approximatif de première niveau (DWT)

WER: Taux d'énergie d'ondelettes << Wavelets Energy Rate >>

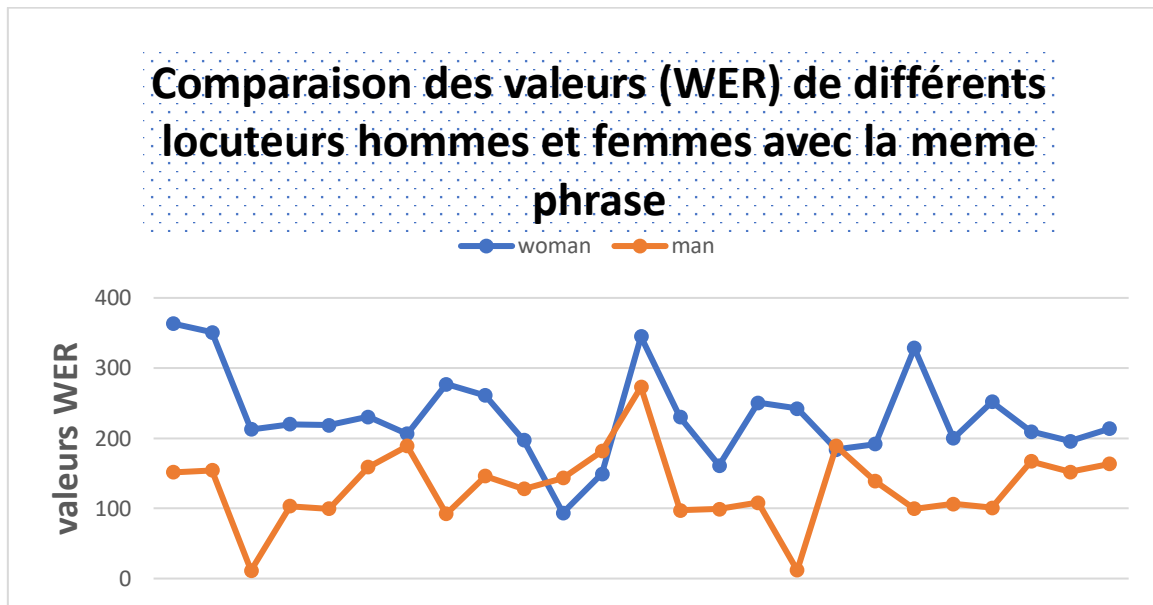


Figure (4-15): Comparaison des valeurs (wer) de différents locuteurs hommes et femmes avec la même phrase , et différents phrases .

IV 5- Résultats de la quatrième partie

Le processus découpage du signal nécessite qu'il soit la fréquence de signal est le double du signal à couper , et puisque la fréquence la plus élevée du signal de parole est de 5000 Hz , il est nécessaire que le signal de parole utilisé dans cette recherche soit à une fréquence de 11025 Hz .les deux figure(a ,b) représente le signal original .

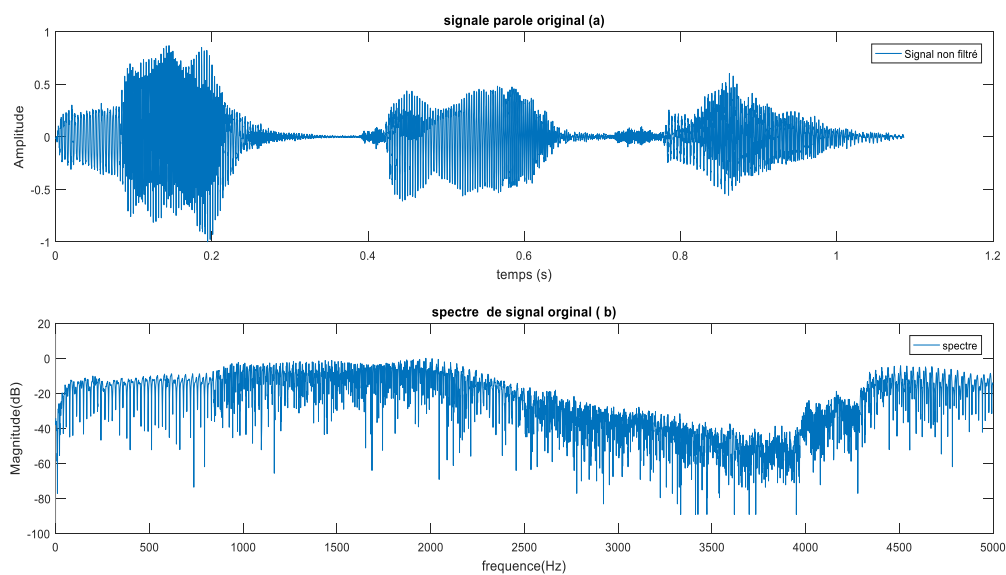


Figure (4-16) : a – signal original , b- spectre de signal original

1- Pré-traitement :

Le pré-traitement du signal parole nécessite de convertir le signal en signal mono, qui ne diffère pas beaucoup du signal d'original, et normaliser par la suite, comme représenté par la figure (4-17). Nous notons qu'après ce traitement, le signal normaliser ne change pas beaucoup par rapport au signal d'origine.

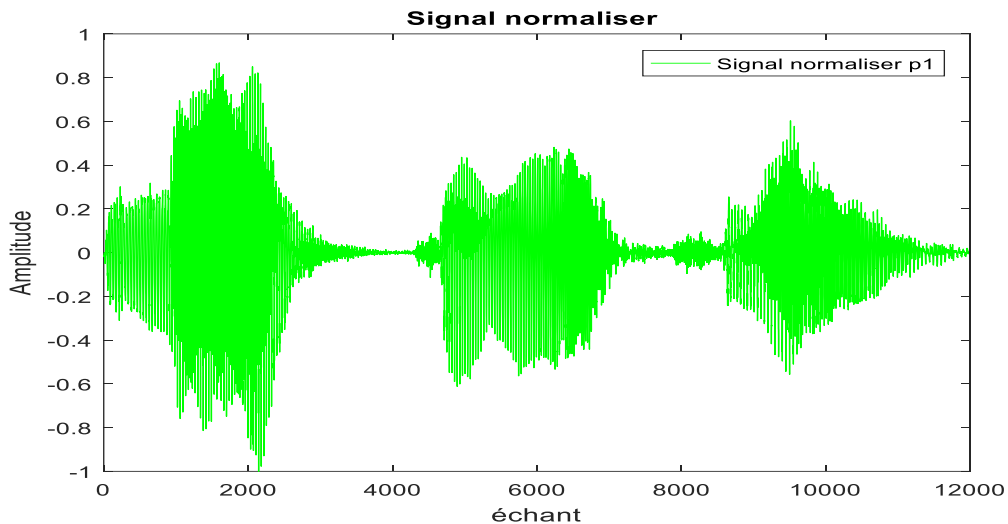


Figure (4-17) : signal original normaliser

2- filtrage du signal

Le processus de filtrage du signal wav est effectué en faisant passer le signal à travers un filtre linéaire passe-bas de Butterworth, généralement d'ordre $n=8$, et ceci pour réduire le bruit dans le signal qui peut être causé par des appareils d'enregistrement ou à partir de facteurs externes, et pour éliminer les fréquences supérieures à 600 Hz à l'aide de ce filtre avant de poursuivre l'analyse, comme présenté à la figure (4-18).

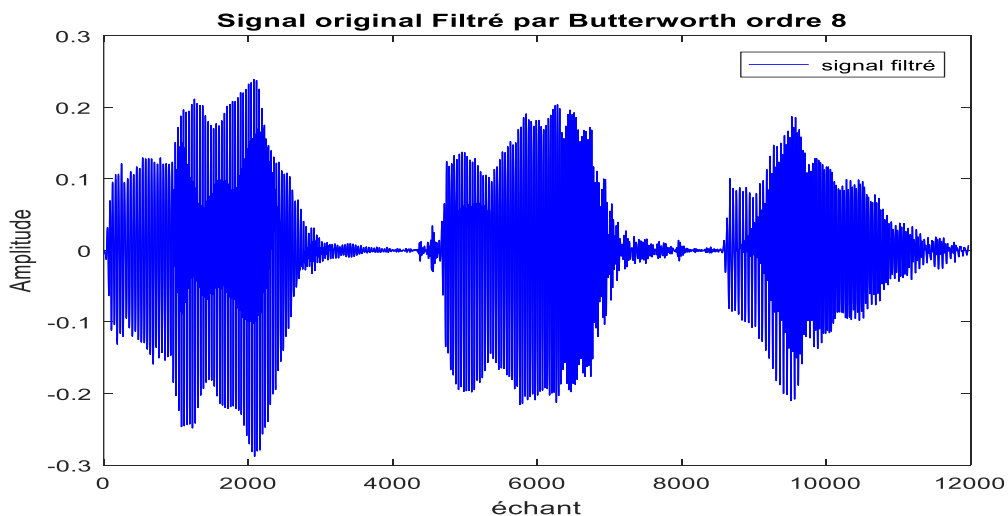


Figure (4-18) : signal original filtré

3- Transformée en ondelette discrète DWT.

Après avoir terminé le processus de Pré-traitement et le filtrage de signal wav , nous appliquons au signal résultant la transformée en ondelettes discrète DWT selon l'équation (3-2-a) ,(3-2-b) des coefficients approximatifs , et coefficients détaillés pour trois niveaux, comme représenté par la figure 4-5-a pour le premier niveau , et 4-5-b pour le deuxième niveau , et 4-5-c pour le niveau Le troisième , où à chaque niveau nous négligeons les coefficients détaillés et effectuons la transformée en ondelettes discrète DWT sur les coefficients approximatifs . Dans la transformée en ondelettes discrète , nous avons utilisé la branche bior 3.9 de l'ondelette mère biorthogonal qui est donner de bons résultats par rapport aux autres branches et ondelettes mère . Cette branche (bior 3.9) a donné la plus grande rétention des caractéristiques du signal dans son troisième niveau , comme le montrent les expériences que nous avons menées dans le mentionné dans le la première partie des troisième et quatrième chapitres est publiée dans la revue Science and Technologie .la figure (4-19) composé de 4 figure présente les trois niveaux de la transformée en ondelettes discrète DWT 4-19 a ,4-19-b , 4-19-c .

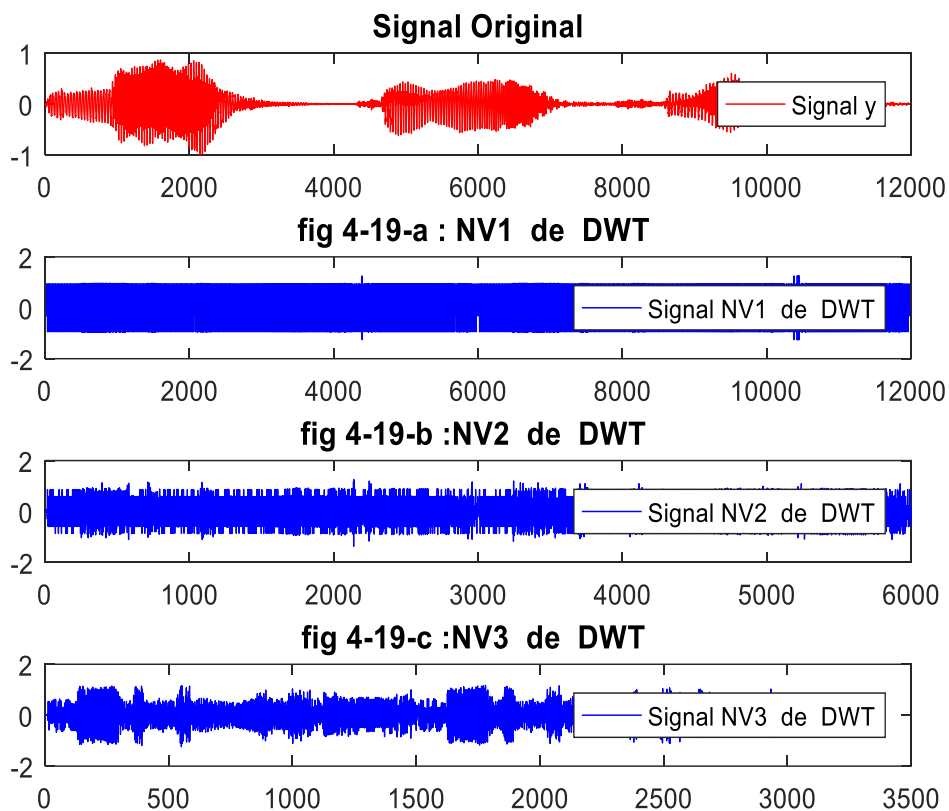


Figure (4-19) : les trois niveau de DWT(4-19 a ,4-19-b , 4-19-c)

4- Framming (Couper le signal en trames):

Au lieu de prendre tout le signal, nous divisons le signal parole en un groupe de trames, car chaque trame est constituée d'un ensemble d'échantillons qui expriment les caractéristiques de cette trame, et chaque trame de durée 30 millisecondes, ce qui est une période très courte pendant laquelle le signal est stable. Par conséquent, après le processus de transformation en ondelettes discrète DWT du signal parole, nous avons divisé le signal des coefficients approximatifs de premier, deuxième et troisième niveaux. Représente une trame d'une longueur de 330 échantillons, et représenté dans les figures (4-20-a),(4-20-b),(4-20-c) respectivement, et c'est parce que le signal parole est un processus aléatoire non stationnaire, or les outils de traitements du signal parole nécessite que le signal soit stable, et la stationnarité du signal, alors pour ce la le signal parole soit quasi stationnaire sur des courts segments de parole appelés "trame". Ces derniers sont des tranches temporelles de 10 à 30 ms, ou de 330 échantillons d'où l'appellation de l'analyse à court-terme. Comme présenter a la figure (4-20).

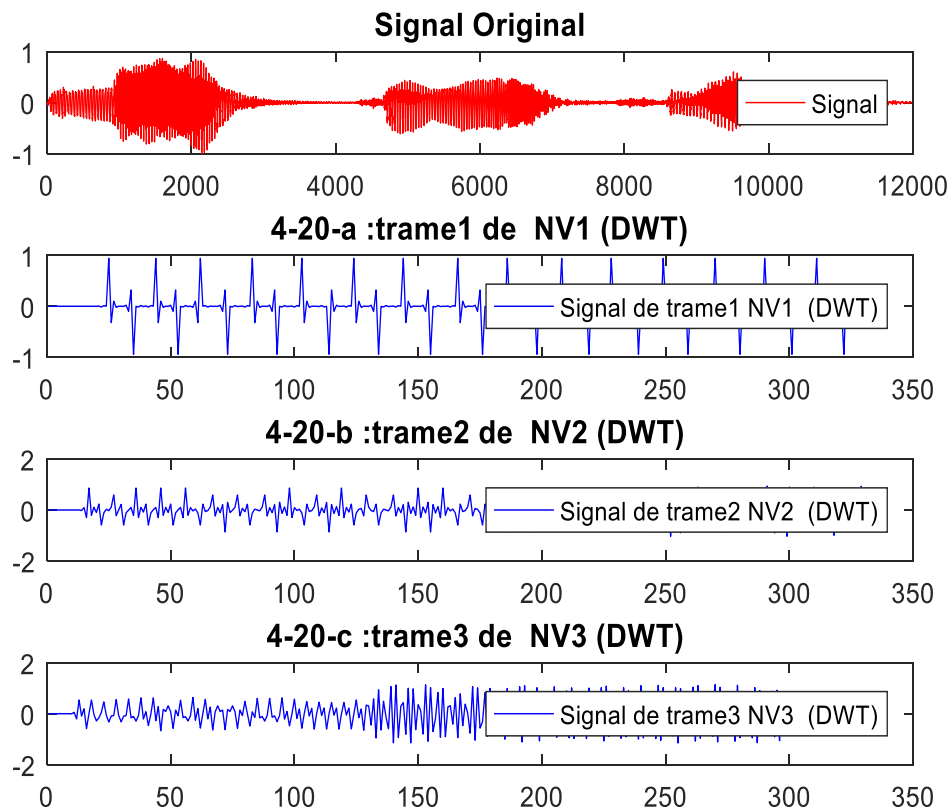
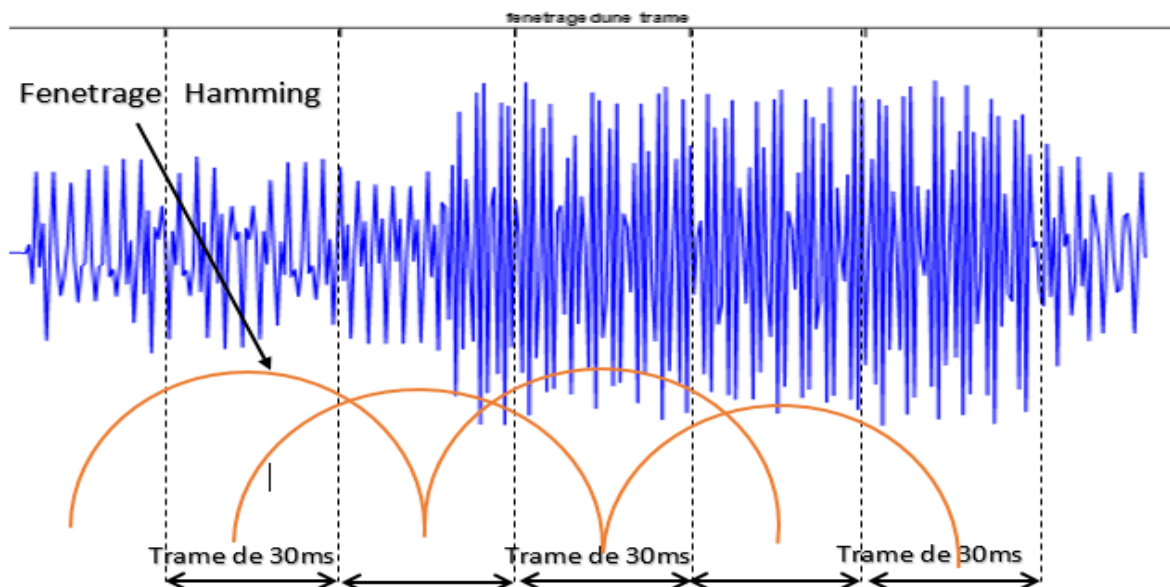


Figure (4-20) : les trois niveau de DWT(4-20 a ,4-20-b , 4-20-c)

5- Fenêtrage (Windowing)

Au cours de cette étape , le taux d'erreur est pouvant réduit à la suite de la division du signal de parole en trames, et des l'échantillonnes , et sa méthode consiste à créer une fenêtre sur la première trame, puis sur la deuxième trame, puis la troisième , et ainsi de suite . La recouvrement de première et la deuxième trame est de 50 à 75 % de , et la même pour la deuxième et la troisième etc.. La figure (4-6) représente le processus de fenêtrage sur les trames du signal.

En pratique la fenêtre de Hamming , est souvent la plus utilisée, puisque il a été démontré que toute variation rapide dans le domaine temporel correspond à des hautes fréquences dans le domaine fréquentiel qui se traduit par des ondulations sur le spectre. Alors on lui préfère d'autres fenêtres plus douces. la fenêtre de Hamming permet une grande atténuation en dehors de la bande passante comparativement à la fenêtre rectangulaire d'où son avantage. Les trames sont soumis au processus de fenêtrage à l'aide de la fenêtre de Hamming, afin d'annuler les discontinuités aux bords, et la fenêtre de Hamming est donnée par la relation ().



La figure (4-21) : présente le fenêtrage sur les trames du signal avec windows Hamming

6- Energie

l'Energie en absence de bruit est efficace pour séparer les sons non voisés de parole .

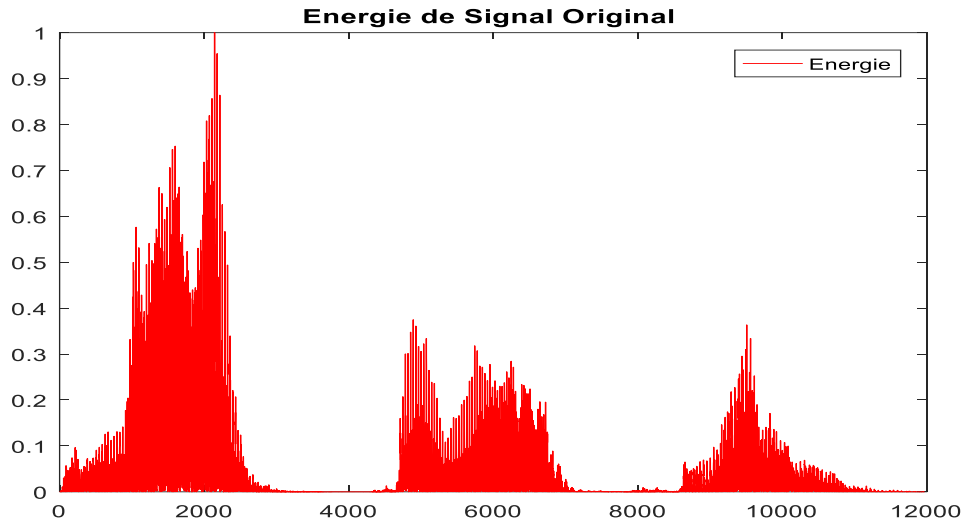


Figure (4-22) : signal original normaliser

7- Energie acourt terme

Pour distinguer les sons voissé et les sons non voisés , nous nous basé sur les algorithmes les plus courants , car ils dépendent de la calcul de l'énergie du signal , et ce la après avoir découpé le signal en trames de certaines longueurs et avec un certain recouvrement , comme le montre l'équation (4-5) En termes absolus.

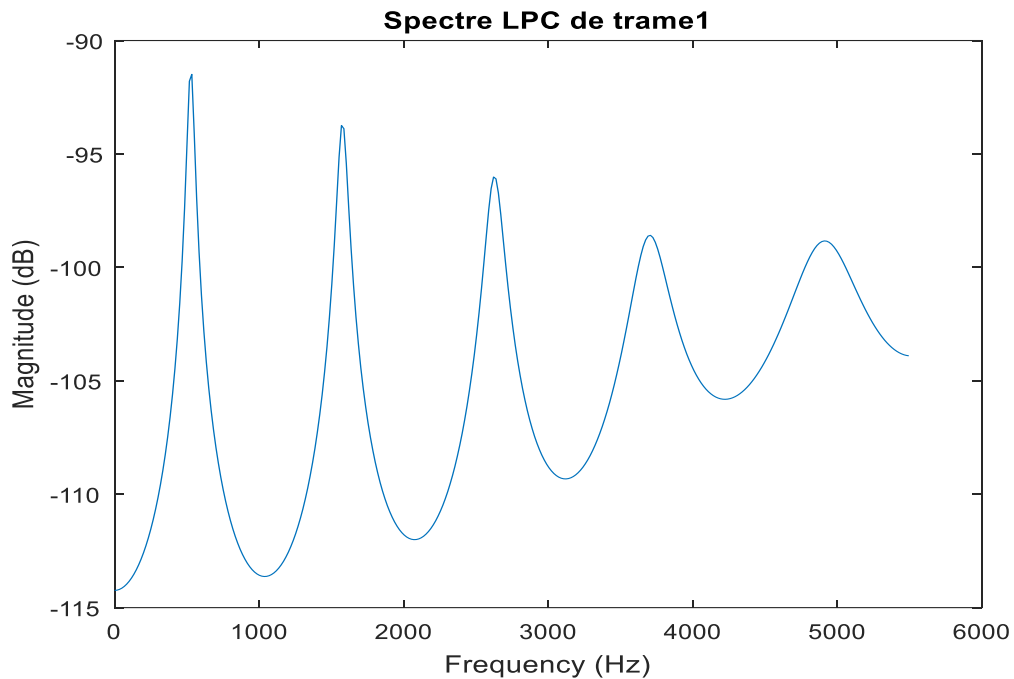
$$4-5 \quad E_n = \sum_{k=1}^n |y(k)|^2$$

8- Autocorrelation court -terme

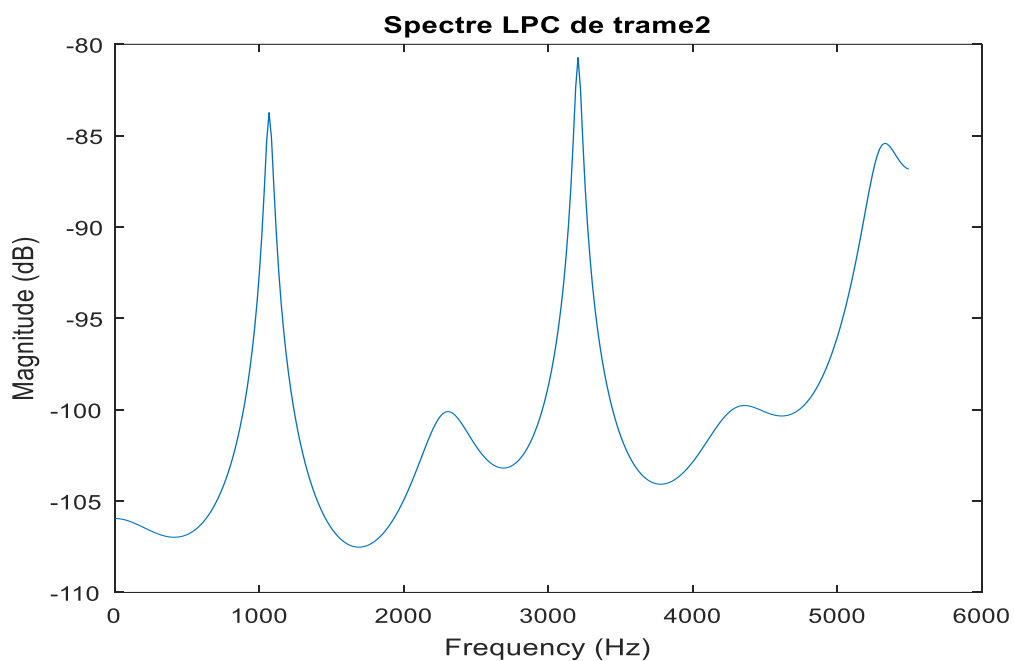
L'idée de l'utilisation de la fonction d'autocorrélation est de déterminer à quel point deux échantillons successifs d'un signal se ressemblent , La fonction d'autocorrélation d'un signal ergodique et stationnaire Sur un trame , c'est le cas d'un fenêtre de notre signal parole, et définie par L'équation (3-12) . comme présenter par les figures (4-23) ,(4-24) ,(4-25) pour les trames de trois niveau DWT successivement . Le résultat de l'auto -corrélacion est un vecteur symétrique de longueur 2N avec un maximum en son milieu. Alors on limite le calcul du maximum dans une seule partie du signal d'auto corrélacion soit partie droite ou gauche.

9- Analyse le LPC

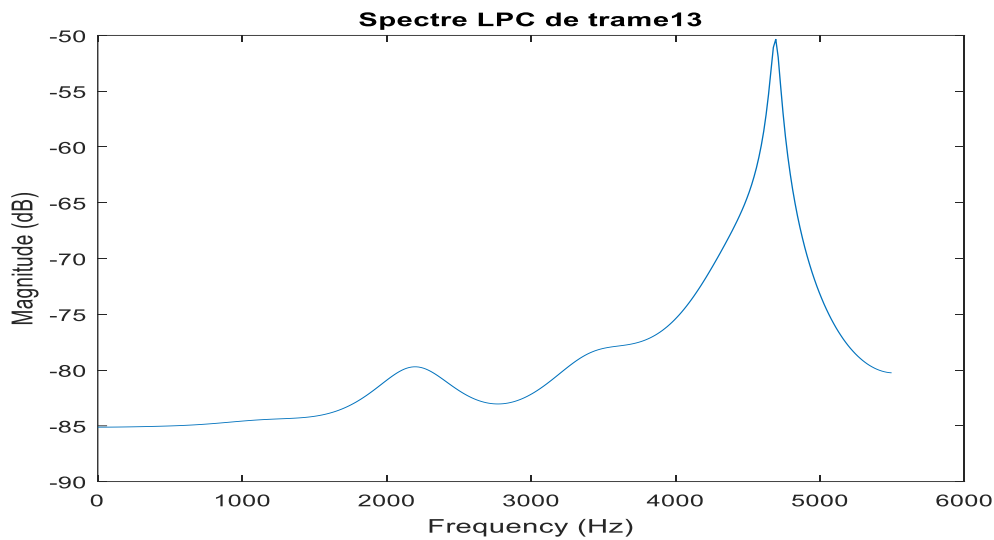
L'objectif de `lpc` est de fournir d'un nombre limité de coefficients de prédiction linéaire et la puissance de l'erreur de prédiction pour représenter directe du signal parole qui s'approche du système phonatoire. Les figures (4-26), (4-27), (4-28) est une représentation des spectres vocaux.



La figure (4-26) : spectre de la partie droite de trame 1 fenêtré par windows Hamming



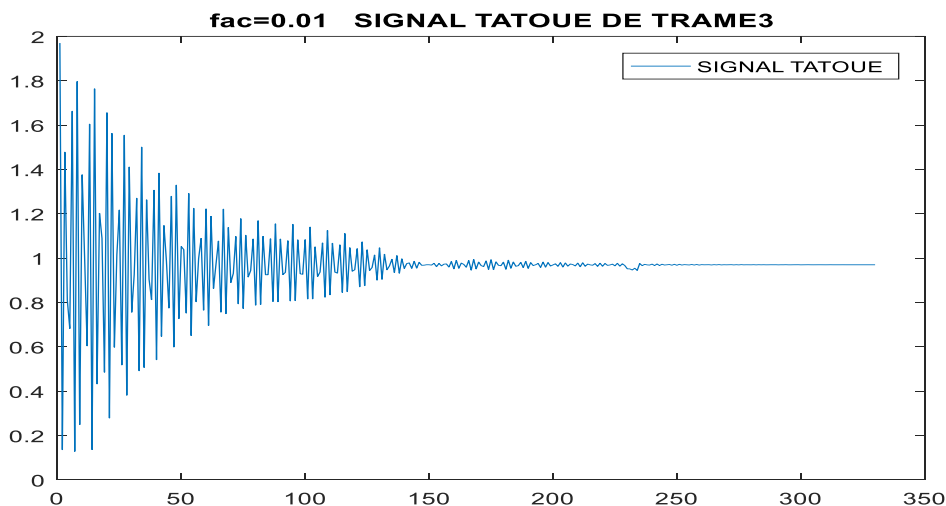
La figure (4-27) : spectre de la partie droite de trame 2 fenêtré par windows Hamming



La figure (4-28) : spectre de la partie droite de trame 3 fenêtré par windows Hamming

IV 6- Resultat et Discussion de la cinquième partie

Les figure (4-29) représente un Signal de sortie de parole tatoué avec un watermark est une trace de matrice 4×4 ces élément sont des paramètre de signal parole d'entre et des coefficients de prédiction linéaire LPC de les trois trames de trois niveau première de DWT .



La figure (4-29) : signal de sortie tatoué de troisième trame avec facteur 0.01

Les Tableaux (4-10), (4-11) présente une comparaison des résultats de lpc de trios niveaux DWT de deux femmes , et deux hommes pour identification et vérification , pour le tableau (4-11) présente le taux de tatouage de signal de sortie par rapport le signal d'entre et les autre taux soit pour l'identification soit vérification . cette comparaison et calculer après une 20 essai pour 38 branches d'ondelettes mères pour la même phrase répétée trois fois et pour des phrases différentes .

	femme 701	femme 702	femme 703	femme 301	femme 302	femme 303	
Sig Original	Nbrs échan	11978	11971	11972	9941	9947	9939
	NPZ	860	867	867	648	653	650
	TPZ	0,0718	0,724	0,0724	0,0652	0,0657	0,654
	Energie	536,2885	568,1532	529,0604	160,274	160,7886	179,6905
Trame 1 nv 1	1 pic max corr	0,038	0,0383	0,0381	0,0383	0,0366	0,0361
	Energie	2,409	2,1593	2,3405	1,9443	1,9214	1,9944
	Coef a2	0,4477	0,4466	0,397	0,3464	0,3535	0,3118
	Coef a3	0,0328	0,0188	0,0335	-0,0805	-0,102	-0,1271
	Coef a4	-0,0057	-0,0179	0,0074	-0,057	-0,0852	-0,076
	Coef a5	0,0089	0,0178	-0,0031	-0,0831	-0,0217	-0,0308
	gain	0,0000056	0,00000561	0,00000523	0,00000669	0,00000599	0,0000056
Trame 1 nv 2	1 pic max corr	0,0572	0,0539	0,543	0,093	0,0909	0,0887
	Energie	4,1793	4,6698	4,2604	2,2887	3,1061	2,8206
	Coef a2	0,5123	0,4901	0,482	0,5277	0,5364	0,5162
	Coef a3	-0,0172	-0,0369	-0,0133	0,1458	0,0988	0,1069
	Coef a4	0,1282	0,112	0,1424	0,3119	0,3927	0,3707
	coef a5	-0,115	-0,1655	-0,1466	0,1517	0,1335	0,1056
	gain	0,0000108	0,0000093	0,0000096	0,0000298	0,0000287	0,0000269
Trame 1 nv 3	1 pic max corr	0,1852	0,1803	0,1897	0,1278	0,1489	0,1487
	Energie	13,5287	14,7914	13,3618	3,1204	3,105	3,0029
	Coef a2	0,858	0,8511	0,8581	0,7926	0,8446	0,8439
	Coef a3	0,1865	0,1832	0,1871	0,2197	0,2216	0,2248
	Coef a4	-0,0821	-0,0696	-0,0855	0,2003	0,1287	0,1238
	Coef a5	0,2277	0,2393	0,2288	-0,0004	0,0202	0,0309
	gain	0,000107	0,0001025	0,0001293	0,0000544	0,0000749	0,0000746

Tableau(4-10) : Comparaison de deux femme avec même phrase

TPZ: Taux de passage par zero du signal

NPZ: Nombre de passage par zero du signal

1pic max corr: 1ere pic max du l'autocorrélation de premiere trame de niveau 3 (DWT)

a2 , a3 , a4 , a5 : Coefficient 2,3,4 du l'autocorrélation de premiere trame de niveau 3 (DWT)

WER : Taux d'énergie d'ondelettes << Wavelets Energy Rate >>

		Trame nv3	Sig Tatoué	Raport sig ta/tr3	Taux iden sexe	Taux de verif	taux de tatouage
Femme 701	Energie	13,5287	320,46	23,7			
	Coeff a2	0,858	-0,2303	-0,3			
	Coeff a3	0,1865	-0,7319	-3,9			
	coeff a4	-0,0821	-0,1645	2,0			
	Coeff a5	0,2277	0,2871	1,3			
	gain	0,000107	0,0171	159,8			
Femme 702	Energie	14,7914	320,4605	21,7			
	Coeff a2	0,8511	-0,2332	-0,3			
	Coeff a3	0,1832	-0,7279	-4,0			
	Coeff a4	-0,0696	-0,1589	2,3			
	Coeff a5	0,2393	0,2811	1,2			
	gain	0,0001025	0,0171	166,8	98,00 %	97,80 %	99,04 %
Femme 703	Energie	13,3618	320,46	24,0			
	Coeff a2	0,8581	-0,2297	-0,3			
	Coeff a3	0,1871	-0,7321	-3,9			
	Coeff a4	-0,0855	-0,1672	2,0			
	Coeff a5	0,2288	0,2911	1,3			
	gain	0,0001293	0,0171	132,3			
Homme A03	Energie	2,9235	320,46	109,6			
	Coeff a2	0,7738	-0,2475	-0,3			
	Coeff a3	0,2816	-0,6248	-2,2			
	Coeff a4	0,0236	-0,2259	-9,6			
	Coeff a5	0,1032	0,1266	1,2			
	gain	0,0000428	0,0181	422,9			
Homme B02	Energie	2,0872	320,46	153,5			
	Coeff a2	0,728	-0,2838	-0,4			
	Coeff a3	0,2989	-0,5611	-1,9			
	Coeff a4	0,0237	-0,2559	-10,8			
	Coeff a5	0,1733	0,1555	0,9			
	gain	0,0000278	0,0178	640,3			

Tableau(4-11) : Taux de tatouage

1pic max corr: 1ere pic max du l'autocorrélation de premiere trame de niveau 3 (DWT)
a2 , a3, a4, a5 : Coefficient 2,3,4 du l'autocorrélation de premiere trame de niveau 3 (DWT)
WER : Taux d'énergie d'ondelettes << Wavelets Energy Rate >>

IV -7 Conclusion et Discussion de quatrième chapitre

Le quatrième chapitre traite des résultats finaux des expériences appliquées présentées par des tableaux des mesures, et des formes de résultat du troisième chapitre, qui sont en cinq parties. Les résultats de la première partie du troisième chapitre se compose de 06 tableaux, et 06 figures, qui présentent une comparaison entre les ondelettes mères, leurs branches les plus appropriées, et une comparaison avec un signal de ECG, et des signaux différent de deux sexe de locuteur avec la même phrase. Les 06 tableaux de tableau (4-1) jusqu'à (4-6), et les 05 figures de 4-1 jusqu'à 4-6 représentent les résultats de la comparaison des énergies retenues dans la transformée discret du troisième niveau des ondelettes mères, et de leurs branches. a travers ces tableaux, et les figures. Nous avons conclu à identifier la branche la plus retenue d'énergie parmi les ondelettes mères.

A travers les valeurs indiquées dans le tableau (4-9) comparatif, qui nous donne les valeurs du critère (WER) << taux d'énergie d'ondelette >> pour le signal de la parole étudiée $x(n)$. Le signal est une phrase répétée pour la même femme, ou un même homme, ou bien la même phrase pour plusieurs hommes, ou plusieurs femmes comme le montre la Figure (11), est une comparaison des valeurs de ce critère par rapport à la plage des valeurs de fréquence fondamentale f_0 dans des recherches publiées récemment. Le critère WER fournit des valeurs approximativement proches des valeurs de f_0 pour des locuteurs de phrases de sexe différents. En conséquence, nous pouvons, à travers le WER, spécifier des valeurs dans une plage de 50 à 180 pour les locuteurs de sexe hommes, et des valeurs dans une plage de 180 à 500 pour les locuteurs de sexe femmes, qui sont presque similaires à la plage de valeurs de fréquence fondamentale f_0 pour les hommes, et les femmes, afin de déterminer le sexe du locuteur en fonction de la valeur de sa fréquence fondamentale selon différentes méthodes. La vitesse de calcul direct du WER reste plus rapide, et plus facile pour déterminer le sexe du locuteur. Le tableau (4-9) et la figure (4-15) montre la comparaison des valeurs WER et la figure (11) montre leur progression.

Après cette étape qui détermine le sexe du locuteur. L'étape suivante de vérification du locuteur est suivie par référence à notre base de données stockée, selon le processus de comparaison basé sur un algorithme permettant d'authentifier, de confirmer, et vérifier le locuteur comme le montre l'organigramme de la Figure (3-12) Et les figures de (4-16) jusqu'à (4-28), et les tableaux de (4-10) jusqu'à (4-11) ici présents. La quatrième partie traite la diversification, la modification de la reconnaissance, et de la vérification du locuteur en fonction de son signal vocal en utilisant, et en s'appuyant sur critère LPCC. Nous avons présenté une nouvelle approche utilisant des matrices dans le processus de comparaison entre les locuteurs basé sur données après extraction des coefficients spectraux de la prédiction linéaire, et de calcul d'une matrice de $4 * 4$ dimensions composée des propriétés du signal, et de calcul de la trace de cette matrice. Ceci selon des tableaux, et des figures, qui nous donne un pourcentage de 98 % de la reconnaissance de locuteur.

Enfin l'organigramme figure (3-19) , dans lequel nous montrons la méthode de sécurisation du signal vocal au moyen du watermark, de la matrice et la trace de cette matrice avec un facteur fixe. La figure (4-29) et les chiffres de tableau (4-11) montrent les résultats obtenus après application de cet organigramme à un signal wav, et c'est ce que nous avons vu à la cinquième phase du troisième chapitre pour améliorer et diversifier le processus de sécurisation des informations extraites à l'aide d'un Watermark pour assurer plus de protection contre les attaques externes et de tous types de piratage. Le tableau montre une comparaison entre le signal d'origine et le signal de sécurité.

Conclusion Générale

Dans notre monde, avec le développement de la technologie et de l'intelligence artificielle, le piratage des informations circulants dans ce monde virtuel a évolué avec lui, ce qui a obligé les chercheurs à trouver des nouvelles façons de protéger les informations contre les attaques électroniques. Pour cela nous utilisons des nombreuses méthodes de protection, parmi ces méthodes figure l'utilisation de watermark pour protéger les informations, qui est une empreinte digitale indiquant le propriétaire de ces informations.

La première partie se limite à sélectionner la meilleure ondelette mère et la branche la plus appropriée pour retenir les caractéristiques du signal dans le troisième niveau de la transformée en ondelettes discrète DWT. Après avoir comparé toutes les ondelettes mères, qui sont (06) avec leurs branches (38), et en appliquant cette transformation à d'autres signaux différents comme le signal cardiaque ECG, et selon le tableau de comparaison (4-7) et (4-6). Dans la première partie du quatrième chapitre, nous avons trouvé que la meilleure branche pour conserver les propriétés dans le troisième niveau de l'ondelette discrète est la branche (bior 3.9) de l'ondelette mère biorthogonal avec un pourcentage jusqu'à 98%.

Dans cette recherche, nous avons présenté une méthode différente et variée pour identifier, vérifier le locuteur et protéger le signal de parole à l'aide d'un watermark basé sur une matrice et la trace de cette matrice. Cette dernière, est un facteur fixe dans les calculs mathématiques. Nous avons effectué le processus de multiplication d'un coefficient fixe avec une matrice de dimension (4*4) dont les éléments de la première ligne sont constitués des caractéristiques de signal d'original, y compris l'énergie, le nombre et le taux de passage par zéro, et le nombre des échantillons de ce signal, respectivement. Pour les éléments des lignes de (2) à (4), sont respectivement les quatre premiers coefficients pour les coefficients spectraux de prédiction linéaire du segment de signal du locuteur après l'application des fenêtres hamming et une transformée en ondelettes discrètes DWT pour les trois niveaux de ce signal qui est une phrase parlée des deux sexes, qui a été utilisée à toutes les étapes de la recherche. Lors de l'utilisation des matrices, nous avons trouvé un nouveau critère pour identifier et vérifier le locuteur basée sur la comparaison de la matrice résultant des opérations mathématiques avec la matrice de la banque des données précédemment préparée, que nous avons discuté dans le troisième chapitre.

Afin de réduire la durée qu'il faut pour reconnaître, et identifier le locuteur, nous avons introduit un nouveau critère que nous avons appelé (WER). C'est un critère qui peut déterminer le sexe du locuteur. Il est similaire à la fréquence fondamentale. Ce dernier dépend d'une fourchette des valeurs à partir de laquelle on peut déterminer le sexe du locuteur. Le but de ce critère est de restreindre la comparaison du signal à un nombre limité de locuteurs selon le sexe d'un homme ou d'une femme afin de réduire les opérations mathématiques qu'au lieu de comparer le signal avec tous les signaux des deux sexes.

Ce critère fait l'objet d'un article que nous avons publié dans une revue internationale classe " B " sous le titre "A Novel Approach for Speaker Gender Identification and Verification using DWT First Level Energy and Zero Crossing " en 2022.

Nous pouvons dire que nous avons pu diversifier la reconnaissance de la parole, la vérification du locuteur et la protection de signal du locuteur en utilisant le watermark avec une nouvelle méthode d'utilisation contre les attaques et les pénétrations qui entraînent des dommages au signal.

Enfin, on ne peut pas dire que ces méthodes et critères suffisent à protéger l'information, mais ce travail reste acceptable pour un taux de reconnaissance vocale de 98%. Il est sujet à enrichir, et à discuter dans toutes ses étapes.

Les perspectives futures de développement du watermark dans le domaine du signal de parole s'annoncent prometteuses. Avec l'évolution des technologies de traitement du signal et de l'apprentissage automatique, on peut s'attendre à des watermarking plus robustes et invisibles, capables de résister aux attaques malveillantes tout en préservant la qualité audio. De plus, l'utilisation de ces techniques pourrait s'étendre à de nouveaux domaines tels que la vérification de l'authenticité des enregistrements vocaux et la protection des droits d'auteur dans l'audio numérique.

Bibliographie

- [1] : Max Chassé. Juillet 2002. Document d'analyse . Commission d'accès à l'information . Juillet 2002
- [2] : Zhitao, L., Dong Youn, K., & William A, P. (2000, July). Wavelet compression of ECG signals by the set partitioning in hierarchical trees (SPIHT) algorithm. IEEE Transactions on Biomedical Engineering, 47(7), 849 - 856.
- [3]: KapilTajane*, Rahul Pitale*, Dr. JayantUmale**-- January 2014 --ReviewPaper :Comparative Analysis Of MotherWaveletFunctionsWith The ECG Signals -- KapilTajane et al Int. Journal of Engineering Research and Applications--ISSN : 2248-9622, Vol. 4, Issue 1(Version 4) .
- [4] ياسر سعيد خضرا. (2016) Wavlet Wavelet for Mother Suitable Selection of the Transform - Application on the ECG Processing .مجلة جامعة دمشق للعلوم الهندسية، 32(2)، 71-82.
- [5] : Shweta, J., & M.P, P. (2013, october). Comparative Analysis of Various Wavelet Families used for R-Wave Detection of ECG Waveforms. International Journal of Innovative Research in Science ,Engineering and Technology (IJIRSET), 2(10), 5177-5181.
- [6]: Fadoua Bahja. 14 Jan 2014.Détection du fondamental de la parole en temps réel : application aux voix pathologiques. Université Mohammed V-Agdal UFR Informatique et Télécommunications .
- [7]: M. A. F. M. Rashidul hasan . Tetsuya Shimamura. May 2012Windowless-Autocorrelation-Based Cepstrum Method for Pitch Extraction of Noisy Speech .
- [8]:Galeyev, B. M. (2010). A brief history of sound in art. Leonardo Music Journal, 20, 33-35.
- [9] : https://fr.wikipedia.org/wiki/Histoire_des_supports_d%27enregistrement_sonore
- [10] : Ballou, G. (Ed.), 2008. Handbook for Sound Engineers, fourth edition. Focal Press.
- [11] : https://cmtext.indiana.edu/acoustics/chapter1_sound2.php
- [12] : https://leselectroniciens.com/sites/default/files/cours/introduction_au_traitement_du_signal.pdf.
- [13] : “ Types of Waves: Transverse and Longitudinal Waves “ by ThoughtCo .
- [14] :M. S. Longair, Theoretical Concepts in Physics, 2nd ed. Cambridge University Press, 2003.
- [15]: R. E. Moore, "Auditory perception," in The Psychology of Human Perception, 2nd ed., e . by D. Regan (Academic Press, 1981), pp. 29-75.
- [16] : Sound Waves and Music" by HyperPhysics

- [17] : Proakis, J. G., & Manolakis, D. G. (2006). Digital signal processing: principles, algorithms, and applications. Pearson Education.
- [18]: Ladefoged, P., & Johnson, K. (2011). A course in phonetics (6th ed.). Boston: Wadsworth Cengage Learning.
- [19]: Roach, P. (2009). English phonetics and phonology: A practical course (4th ed.). Cambridge: Cambridge University Press.
- [20]: "Introduction to Phonetics and Phonology" by John J. Ohala and Manuel J. Sevilla-Garcia
- [21]: R. Boite et al. 2000. Traitement de la parole.
- [23] : D. O'shaughnessy, "Speech communication human and machine," Addison-Wesley, 1987.
- [24]: Annie Leclerc .fevrier 2007. Le role de la vision dans la production de la parole etude articulatoire et acoustique . université du quebec a montreal.
- [25] : <https://www.msmanuals.com/fr/accueil/troubles-du-nez,-de-la-gorge-et-de-l%20%99-oreille/troubles-laryng%C3%A9s/paralysie-des-cordes-vocales>
- [26] : H. Gray, "Anatomy of the human body," Philadelphia: Lea & Febiger; Bartleby.com, 2000.
- [29]: Christian jutten . Juillet 2018 . Théorie du signal , Département Informatique et Electronique des Systèmes Embarqués, IESE4 ,Univ. Grenoble Alpes - Polytech' Grenoble
- [30] : Marie Tahon , 2014-2015 , Traitement du signal , Laboratoire d'Acoustique,
- [31] : https://fr.wikipedia.org/wiki/th%C3%A9orie_du_signal
- [32] : Francis Cottet . Aide mémoire traitement du signal .
- [33] : https://fr.wikipedia.org/wiki/Transformation_de_Fourier
- [34] : Christophe doignon . (2008-2009) - FIP 2A . Traitement Numérique du Signal D'déterministe . Université Louis Pasteur de Strasbourg
- [35] : Traitement du signal et la reconnaissance de la parole ,Présenté par : Maripo Tsivery Tanjona .Soutenu le 23 Avril 2010)
- [36].(<https://www.rapport-gratuit.com/les-caracteristiques-du-signal-de-la-parole/>).
- [37] : Cyril_Plapous_Thesis .Traitements pour la réduction de bruit .Application à la communication parlée. 12 Décembre 2005 . université de rennes .
- [39] : <https://wiki.aalto.fi/pages/viewpage.action?pageId=149890776>
- [40] : [<https://speechprocessingbook.aalto.fi/>],
- [41] : Hassani Ezzaidi .décembre 97. Détection de la double parole dans le contexte de radiotéléphone main – libre en véhicule .

- [42] :Jont B. Allen , Lawrence R , Rabiner , Fellow . Nov 1977. A Unified Approach to Short-Time Fourier Analysis and Synthesis . proceeding of the IEEE . vol .65 , N11.
- [43]: Traitement du signal et reconnaissance de la parole.Présenté par : MARIPO Tsivery Tanjona .Soutenu le 23 Avril 2010
- [44] : Lawrence R. Rabiner and Ronald W. Schafer.2007. Introduction to Digital Speech Processing . Rutgers University and University of California Santa Barbara .USA
- [45] : Cyril , Plapous . 12 Décembre 2005 .Traitements pour la réduction de bruit.Application à la communication parlée. Université de rennes 1- IRISA/ENSSAT
- [46] : M. B. Sachs, C. C. Blackburn, and E. D. Young, “Rate-place and temporal-place representations of vowels in the auditory nerve and anteroventral cochlear nucleus,” *Journal of Phonetics*, vol. 16, pp. 37–53, 1988.
- [47] :<http://www.cochlea.eu/cochlee/fonctionnement>.
- [48] : H. Fletcher and W. J. Munson. October 1933. “Loudness, its definition, measurement and calculation,” *Journal of Acoustical Society of America*, vol. 5, no. 2, pp. 82–108 .
- [49] : Published in: *IEEE Spectrum* (Volume: 7, Issue: 10, October 1970) .
- [50] : modélisation spectrale et compression de la parole a base débit . NQ67110
- [51] :Khaled Zaabi . 10 juin 2004 . Implémentation d’une méthode de reconnaissance de la parole sur le processeur de traitement numérique du signal TMS320C6711.
- [52] : CHARPENTIER , Christophe
- [53] :https://fr.wikipedia.org/wiki/ARMA#Mod%C3%A8le_autor%C3%A9gressif
- [54] : <https://www.thalesgroup.com/fr/europe/france/dis/gouvernement/biometrie/empreintes-digitales-et-identification>
- [55] : A.K. Jain, S. Prabhakar, and S. Chen. 1999. “Combining multiple matchers for a high security fingerprint verification system”, *Pattern Recognition Letters*.
- [56] : M. El-Abed, R.Giot, B.Hemery and C.Rosenberger, “A study of users’ acceptance and satisfaction of biometric systems”, In 44th IEEE International Carnahan Conference on Security .Technology, San Jose, California, USA, 2010.
- [57] : Nicolas MORIZET, Reconnaissance Biométrique par Fusion Multimodale du Visage et de l’Iris Thèse de Doctorat, École Doctorale d’Informatique, Télécommunications et Électronique de Paris,2009.
- [58] : F. Cherif, B. Hemery, R. Giot, M. Pasquet, C. Rosenberger .2009 “Performance evaluation of behavioral biometric systems”, In *Behavioral Biometrics for Human Identification: Intelligent Applications*, 2009.
- [59] : L. Allano .2009 . “ La biométrie mulimodale: Stratégie de fusion de scores et mesures de dépendance. Appliquées aux bases de personnes virtuelles”, Thèse de doctorat, Institut national des télécommunication .

[61] : J. Wayman, "Biometric Systems Technology Design and Performance Evaluation", London, 2005.

[62] : https://fr.wikipedia.org/wiki/Empreinte_digitale

[64] : Pierre Bonazza. 21 Juin 2019. Système de sécurité biométrique multimodal par imagerie, dédié au contrôle d'accès. Université Bourgogne Franche-Comté .

[65] : AfB and ICISA: 1998 Glossary of Biometric Terms, Association for Biometrics and International Computer Security Association, to be referred at URL: <http://www.afb.org.uk/> (1998).

[66] : van Renesse, R. L.1998. An Introduction to Biometrics, in Optical Document Security, Second Edition, Artech House, Rudolf L. van Renesse ed. Chapter 19 .

[67]:[https://www.bcvision.es/blog-vision-artificial/algorithmes-de-reconnaissance-faciale-cas - dutilisation-et-controverse/](https://www.bcvision.es/blog-vision-artificial/algorithmes-de-reconnaissance-faciale-cas-dutilisation-et-controverse/)

[68]: Frédéric Massicotte .Nov 2007.la biométrie ,sa fiabilité ses impacts sur la pratique de la démocratie libérale .université du Québec a Montréal .

[69]: <https://www.wedemain.fr/inventer/9-dates-qui-ont-marque-l-essor-de-la-reconnaissance-faciale-a4568-html/>

[70] : Peter Gregory And Michael A. Simon.2008 " Biometrics For Dummies", Cisa, Cissp.

[71] : café techno - L'histoire de la biométrie : d'Alphonse Bertillon à l'iPhone x auteurs : Antitza dantcheva (Inria) & Jean-Luc dugelay (Eurecom), Sophia Antipolis.

[72] : Takwa chihaoui . 06 Décembre 2018. Système d'Identification de personnes basé sur la rétine. Université Paris - Est – ESIEE-Paris Ecole Doctorale Mathématiques et Sciences et Technologies de l'Information et de la Communication (Mstic)

[73] :John D. Woodward, Jr, Christopher Horn, Julius Gatune, Aryn Thomas. Published 2003 by rand . Biometrics A Look at Facial Recognition.

[74] : Nicolas morizet .2009. Reconnaissance Biométrique par Fusion Multimodale du Visage et de l'Iris Thèse de Doctorat, École Doctorale d'Informatique, Télécommunications et Électronique de Paris.

[75] : Eugen Harinda, Etienne Ntagwirumugara.2015 .Security & Privacy Implications in the Placement of Biometric-Based ID Card for Rwanda Universities. Journal of Information Security, 6, 93-100.

[77] : J. Bhatnagar, A. Kumar, "On estimating performance indices for biometric identification", Pattern Recognition, 2009.

[78] : [Biometric systems for access granting].

[79] :https://atelieremploi.fr/wiki/iris_recognition

[80] :<https://labiometrie.wordpress.com/2017/02/11/reconnaissance-de-liris/>

[81] : John Daugman . 2004 . How Iris Recognition Works . EEE . transactions on circuits and systems for video technology .

[82]:<https://sante.journaldesfemmes.fr/fiches-anatomie-et-examens/2782213-iris-definition-schema-oeil-marron-bleu-jaune/>

[83] :Haroun djaafar. 29 / 06 / 2010. Identification basée sur le code d'iris . université Sétif., Algérie .

[84] :<https://www.synergielyoncancer.fr/glossaire/adn>

[85] : <https://perspective.usherbrooke.ca/bilan/servlet/BMEve/1517>

[86] : https://en.wikipedia.org/wiki/Photo_51#cite_note-due_credit-2

[87] : <https://www.google.com/search?client=firefox-b-d&q=definition+ADN>

[88] : Masaki Hashiyada , tohoku J EXP.Med .2004.Development of Biometric DNA Ink for Authentication Security.

[89] : Etude-comparative-entre-la-cryptographie .

[90] : <https://fr.wikipedia.org/wiki/Cryptographie>

[91]:<https://www.journaldunet.fr/web-tech/dictionnaire-du-webmastering/1203491-cryptographie-definition-traduction/>

[92] : Tirkel A. Z. , Osborne C.F ., Mee N. R.A., Rankin G.A., McAndrew A. Maximal Connected Set s Application to Microcell CDMA . Submitted to the International Journal of Digital and Analog Communications 1993.

[93] : <https://fr.wikipedia.org/wiki/St%C3%A9ganographie>

[94] :Tatouage_Robuste_Et_Aveugle_Dans_Le_Doma

[95] : <https://apcpedagogie.com/quest-ce-que-le-watermarking/>

[96]: Hai Tao *1 , Li Chongmin* 2 ,Jasni Mohamad Zain 1 , Ahmed N. Abdalla 3A Survey of Digital Watermarking Techniques and it Robust Image Watermarking Theories and T Robust Image Watermarking Theories and Techniques .

[97] : August 2012. "A comparative study of digital watermarking techniques" de Bhatia et Singh (2018).International Journal of Computer Applications (0975 – 8887) .Volume 52– No.20.

[99] : Prabhishkek Singh, R S Chadha .A Survey of Digital Watermarking Techniques. , March 2013.Applications and Attacks International Journal of Engineering and Innovative Technology (IJEIT) .Volume 2, Issue 9 .

[100] :Gwena ël Do ërr , Jean-Luc Dugelay .Problématique de la Collusion en Tatouage Vidéo Collusion Issue in Video Watermarking

[101] : Le tatouage de données audiovisuelles .. Philippe Nguyen – Séverine Baudry

[102] :Prabhishkek Singh, R S Chadha . A Survey of Digital Watermarking Techniques, Applications and Attacks

[103] : Frank Hartung , Martin Kutter Multimedia Watermarking Techniques .

- [104]: International Journal of Innovative Technology and Exploring Engineering (IJITEE),ISSN: 2278-3075, Volume-8 Issue-7S2, May 2019 Spread Spectrum Based Digital Audio Watermarking. Bodhvi Gaur, Chandrakala.C.B
- [105] : wikipedia. (2022, juillet 30). ondelette. Récupéré sur wikipedia: <https://fr.wikipedia.org/wiki/Ondelette>
- [106] : polikar, R. (2006). The Engineer's Ultimate Guide to Wavlet Analysis - The Wavlet Tutorial. College of Engineering Web Servers: Rowan University.
- [107] : Karanpreet, K., & Ankush, K. (2014). Performance Analysis of Convolutional Interleaved DWT based OFDM system. International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), 3(5), 6601-6604.
- [108] :Sumalatha, S., & Udaykiran Bhargava, Gollamandala, G. (2018, july 20). BER Analysis of Translation Invariant Wavelet Based Orthogonal Frequency Division Multiplexing System for 5G Wireless Communication Networ. international Journal of Intelligent Engineering and Systems.
- [109] :Sumithra M G, M. I., & Thanushkodi , k. (2009, December). Performance Evaluation of Different Thresholding Methods in Time Adaptive Wavelet Based Speech Enhancement. IACSIT International Journal of Engineering and Technology, 1(5), 439-447.
- [110] : Gao, r., & Ruqiang, Y. (2011). Wavelets Theory and applications for Manufacturing. Springer. doi:10.1007/978-1-4419-1545-0
- [111] :Merry, R. J. (2005, June 7). Wavelet theory and applications A literature study. 2005.053. Department of Mechanical Engineering: Eindhoven University of Technology.
- [112] : Misiti, M., Misiti, Y., Oppenheim, G., & Poggi, J.-M. (2009). Wavelet Toolbox™ 4 User's Guide MathWorks Inc.
- [113] : Speaker Identification Using a Butterworth Filter and Artificial Neural Networks" by Ahmed A. Alshammari, published in the International Journal of Computer Science Issues (IJCSI), Vol. 14, Issue 5, September 2017.
- [114] :Craciun Alexandru . Speaker Identification using Neural Networks . R.V Pawar, P.P.Kajave, and S.N.Mali . proceeding of world academy of science . engineering and technology volume 7 august 2005 ISSN 1307-6884.
- [115]: Noor Almaadeed 1 & Amar Aggoun 2 & Abbes Amira 1,3 . J Sign Process Syst (2016) 82:345–356.Text-Independent Speaker Identification Using Vowel Formants .
- [116] : How to cite this paper: Këpuska, V.Z. and Elharati, H.A. (2015). Robust Speech Recognition System Using Conventional and Hybrid Features of MFCC, LPCC, PLP, RASTA-PLP and Hidden Markov Model Classifier in Noisy Conditions. Journal of Computer and Communications, 3, 1-9. Published Online June 2015 in Scires .
- [117] :Li,J, Deng ,L, Haeb-Umbach,R,& Gong ,Y, (2015). Robust Automatic Speech Recognition : A Bridge to Pratical Applications . Academic Press.
- [118] : <https://www.elettroamici.org/fr/teorema-di-nyquist-shannon/>.

[119]: Abdelkader AMRAOUI¹ , Slami SAADI² . A Novel Approach for Speaker Gender Identification and Verification using DWT First Level Energy and Zero Crossing . Engineering, Technology & Applied Science Research Vol. 12, No. 6, 2022, 9570-9578 . p 9570 - 9578.

[120] :Par Art Pini , Avec la contribution de Rédacteurs nord-américains de Digi-Key , 24-03-2020 .Notions de base sur les filtres passe-bas antirepliement.

[121] :https://fr.wikipedia.org/wiki/Fréquence_de_coupure .

[122] : D.S.Shete¹ , Prof. S.B. Patil² ,Prof. S.B. Patil³ . Volume 4, Issue 1, Ver. I (Jan. 2014), PP 01-05 . IOSR Journal of VLSI and Signal Processing (IOSR-JVSP).Zerocrossing rate and Energy of the Speech Signal of Devanagari Script .

[123]: Ranganadh Narayanam * Assistant Professor, Eece, Its, Ifhe – India . Narayanam * et al. September, 2017 0 international journal of engineering sciences et resherch technology . voiced and unvoiced separation in speech auditory brinstem responses of human subjects using zero crossing rate (zcr) and energy of the speech signal .

[124]: Lawrence R. Rabiner¹, Ronald W. Schafer² . 2007 .Editorial Board . Introduction to Digital Speech Processing . Foundations and Trends R© in Signal Processing Volume 1 Issue 1–2 .Editorial Board

[125]: Khedidja Nekab .Montreal .12 décembre 2013 .Etude comparative des différentes méthodes d'estimation de la fréquence fondamentale des cris des nouveau-nés. Ecole de technologie supérieure .université du Québec .

[126] : M. B. Messaoud, A. Bouzid, N. Ellouze .Published in JEP/TALN/RECITAL 1 June 2012 .Computer Science, Mathematics. Estimation du pitch et décision de voisement parcompression spectrale de l'autocorrélation du produitmulti-échelle .

[127] : F. Ykhlef , R. Amiar¹ , S. Hecini , W. Benzaba , L. Bendaouia .2009 . Published in CIIA 2009.Computer Science .Etude Comparative des Performances de Plusieurs Techniques de Détection de la Fréquence Fondamentale des Signaux Vocaux.

[128] : R ,ajou , S sbaa, S aouragh, A. taleb ahmed . Octobre 2011. Détection du pitch par les ondelettes continues en temps réel por un signal parole basé sur un seuil adaptatif pour une détermination V/NV.

[129] : Laetitia Jeancolas.4 décembre 2019 . l'Université Paris-Saclay .détection précoce de la maladie deParkinson par l'analyse de la voix etcorrélations avec la neuroimagerie .