



Université ABBES LAGHROUR Khenchela
Faculté des Sciences et de la Technologie
Département de Génie Industriel



كلية العلوم والتكنولوجيا
قسم الهندسة الصناعية

N° Série :.....

Mémoire de fin d'étude

Pour l'obtention du diplôme de Master

Filière : Télécommunications

Spécialité : Télécommunications Avancées

THEME

Rehaussement du signal de Parole pour les
Communications Téléphoniques dans les
Milieux Bruités

Réalisé par : - HEZIL Ahlam Farah
- GASMI Khaoula

Dirigé par : Mr. CHERGUI Laid

Membres de jury :

Mr: Chergui Laid ----- Rapporteur

xxxxxxxxxxxxx----- Examineur

xxxxxxxxxxxxxxxxx---- Examineur

Promotion 2015/2016



Remerciements

Tout la gratitude et le merci au dieu le tout puissant qui a fait que ce travail soit mené et terminé.

Nous adressons nos vifs remerciements à notre promoteur consultant M^r CHERGUI Laid pour nous avoir diligentées tout au long de ce travail, pour sa Compréhension, sa patience, sa compétence, et ces remarques qui nous ont été précieuses.

Nous tenons à remercier également les membres du jury pour l'honneur qu'ils nous font en acceptant de juger et d'évaluer notre travail.

Nos derniers remerciements, vont à tous ceux qui ont contribué de près ou de loin à la réalisation de ce travail.

HEZIL Ahlam Farah

GASMI Khaoula





Dédicace

Je dédie ma gratitude et mes vertus

A mes très chers parents Boubaker et Noua qui m'ont aidé et qui é t pour m'instruire par leur sacrifice qu'ils trouvent ici le témoignage de profond respect et ma reconnaissance infinie.

Ce que je suis, tous

Ce que je réalise, merci de m'avoir soutenu et aimé durant

Toutes ces années d'être si fiers de moi.

A mes frères Abd Nour, Hilal et Mohamed, à la personne qui compte beaucoup pour moi, à mes tentes et mes oncles, mes chers cousins et cousines, et à toute ma famille HEZIL.

A tout mes amis et collègues de cette promotion surtout mon binôme Khaoula.

A tout ceux qui ont participés de prés ou de loi pour la réalisation de ce travail et ceux qui m'espèrent la réussite.

HEZIL, Ablam Farah.





Dédicace

Je dédie ce mémoire à :

Mes parents : Ma mère Nassima qui a œuvré pour ma réussite, de par son amour, son soutien, tous les sacrifices consentis et ses précieux conseils, pour toute son assistance et sa présence dans ma vie, reçois à travers ce travail aussi modeste soit-il, l'expression de mes sentiments et de mon éternelle gratitude.

Mon père Abd Elmalk qui peut être fier et trouver ici le résultat de longues années de sacrifices et de privations pour m'aider à avancer dans la vie. Puisse Dieu faire en sorte que ce travail porte son fruit ; Merci pour les valeurs nobles, l'éducation et le soutien permanent venu de toi.

Ma sœur Ikhelas mes frères Mohamed et Mahboubi et qui n'ont cessé d'être pour moi des exemples de persévérance, de courage et de générosité.

Mes tentes et mes oncles, mes chers cousins et cousines, et à toute ma famille.

Tous mes amis et mon amie d'enfance Majeda

Mes collègues de cette promotion surtout mon binôme Farah

Mes professeurs qui doivent voir dans ce travail la fierté d'un savoir bien acquis.

GASMI Khaoula.



Résumé :

Ce mémoire présente l'étude de quelques techniques de filtrage qui peuvent être employées dans le débruitage et le rehaussement du signal de parole contaminé par le bruit, un problème rencontré fréquemment dans les communications téléphoniques (Mobile ou GSM) qui sont faites généralement dans des milieux bruités.

Plusieurs techniques sont utilisées pour le filtrage et le débruitage du signal de parole. Dans ce mémoire on a fait l'étude des techniques de débruitage basées sur la soustraction spectrale, le filtrage de Wiener et le filtrage adaptatif au sens de Wiener basé sur les algorithmes récursifs tels que l'algorithme des moindres moyennes carrées (LMS) et l'algorithme de moindres moyennes carrées normalisées (NLMS).

Les performances des différentes techniques sont testées pour un signal de parole en présence de bruit pour différentes conditions.

Abstract :

This work present the study of some filtering technique that can be used in the enhancement of the speech signal affected by noises, a problem frequently encountered in telephonic communications (Mobile or GSM) which are generally made in noisy environnements.

Several techniques are used for speech signal denoising. In this work, we made the study of denoising techniques based on spectral subtraction, Wiener filtering and adaptive filtering. The later, is based on recursive algorithms such as the least mean square algorithm (LMS) and normalised least mean square algorithm (NLMS).

The performances of the different techniques are tested for speech signal in presence of noise for different conditions.

:

هذه المذكرة تعرض دراسة لبعض التقنيات المستعملة في الترشيح والتي يمكن استعمالها في تنقية وتحسين الصوتية الملوثة بالضوضاء، وهي إشكالية نصادفها في المكالمات الهاتفية (GSM)

عدة تقنيات تستخدم من اجل ترشيح و تنقية الصوتية. في هذه المذكرة قمنا بدراسة تقنيات التصفية التي تعتمد على الطرح الطيفي، ترشيح Wiener و الترشيح المكيف الذي يستند على الخوارزمية العودية كخوارزميات اقل متوسط مربع (LMS) خوارزميات اقل متوسط مربع المنتظمة (NLMS).

مختلف التقنيات تم تجربتها من اجل

Table des Matière

Chapitre 1 : Généralités sur le signal de parole.

1.1. Introduction.	1
1.2. Notion du signal :	1
1.3. Production et Perception du signal de parol.	1
1.3.1. Production du signal de parole.	1
1.3.2. Perception de la parole.	2
1.4. Analyse et paramétrisation du signal de parole.	2
1.4.1. Analyse LPC.	3
1.4.2. Cepstre.	4
1.4.3. Application.	6
1.5. Le signal du bruit.	7
1.5.1. Définition de bruit.	7
1.5.2. Nature et caractéristiques du bruit.	7
1.5.3. Les difficultés rencontrées lors de l'utilisation de la parole.	8
1.5.4. Qu' ce qu'un bruit blanc ?	8
1.5.5. Qu'est ce qu'un bruit musical.	9
1.5.6. Les différents types de bruit.	9
1.6. Rapport signal sur bruit (RSB) ou (SNR).	9
1.7. Conclusion.	10

Chapitre 2 : Les Transformées Discrètes Orthogonales.

2.1. Introduction.	12
2.2. La transformation de Fourier discrète (TFD).	12
2.2.1. Définition.	12
2.2.2. Réalisation pratique.	13
2.2.3 TFD de signaux classiques.	14
2.2.4. Analyse spectrale.	14
2.2.4.a. L'analyseur de spectre à balayage (analogique).	14
2.2.4.b. L'analyseur de spectre à FFT (numérique).	15
2.2.4.c. L'analyseur de signaux vectoriel (analogique/ numérique).	15
2.2.4. Analyse de séries temporelles.	15
2.2.5. Application de la TFD.	15
2.2.6. Quelques propriétés de la TFD.	16
2.3. Transformée Hartley discrète (DHT).	16
2.3.1. Définition.	16
2.3.2. Relation avec la transformation de Fourier discrète.	17
2.4. La transformée en cosinus discrète (DCT).	18
2.4.1. définition.	18
2.4.2. Les propriétés de la DCT.	19
2.4.3. Application.	20
5. Conclusion.	21

Chapitre 3 : Les techniques de rehaussement du signal de parole.

3.1. Introduction.	23
3.2. La soustraction spectrale.	23
3.3. Filtrage au sens de Wiener.	26
3.3.1 Filtrage de Wiener.	26
3.3.2. Notations.	29
3.3.3. Calcul de l'erreur quadratique minimale.	29
3.3.3. Equation de Wiener-Hopf.	30
3.3.3.1. Expression générale.	30
3.3.3.2. Solution pour un filtre FIR.	31
3.4. Algorithme du gradient stochastique (least-men-square –LMS).	32
3.4.1. Rappels sur l'algorithme du gradient déterministe.	32
3.4.2. L'algorithme LMS.	33
3.4.3. Résumé de l'algorithme LMS :	34
3.4.4.1. Convergence en moyenne.	35
3.4.4.2. Convergence en moyenne quadratique.	36
3.4.4.3. L'algorithme LMS pour des données complexes	41
3.5. Conclusion.	43

Chapitre 4 : Tests et Expérimentations.

4.1. Introduction :	45
4.2. Débruitage de la parole par la technique de l'annulation de bruit :	44
4.3. Débruitage de la parole par filtrage de WIENER .	46
4.4. Annulation de bruit par filtrage adaptatif :	47
4.5. Annulation du bruit du signal parole :	49
4.6. Mesure de l'amélioration du SNR après filtrage :	52
4.7. Conclusion :	52

Introduction

Général

Introduction Générale

Introduction Générale :

La téléphonie est un moyen de communication moderne basée sur la transmission du signal de parole, qui a constitué un moyen de communication avantageux entre les humains. La présence d'un bruit ambiant lors d'une communication téléphonique influe sur la qualité du signal de parole transmis et dégrade la qualité de la conversation.

Le débruitage du signal de parole constitue un champ d'études vastes dont l'objectif est d'améliorer la qualité du signal parole en le restaurant à partir de la conversation bruitée par des bruits ambiants qui sont généralement de nature acoustique. Dans ce mémoire nous ne considérons que le cas des bruits additifs qui s'ajoute au signal de parole.

Plusieurs techniques de filtrage ont été appliquées pour le débruitage du signal de la parole. Le débruitage de la parole basé sur le filtre de Wiener a été utilisé. Le calcul des coefficients optimaux du vecteur du filtre de Wiener qui permet de donner à la sortie du filtre la meilleure estimation du signal de parole à partir d'une observation bruitée nécessite le calcul de l'inverse de la matrice de l'autocorrelation du signal d'entrée, ce qui nécessite plusieurs opérations arithmétiques. Cette technique souffre de sa lourde complexité de calcul.

Les techniques de filtrage adaptatif qui permettent de réaliser un filtrage au sens de Wiener, tel que le filtrage adaptatif basé sur l'algorithme du gradient stochastique (LMS) permettent de se rapprocher de la solution optimale de Wiener de façon récursive en minimisant l'erreur quadratique moyenne. Cette technique est robuste, simple et présente une complexité de calcul réduite.

Les techniques de débruitage du signal de parole qui opèrent dans le domaine fréquentiel tel que la technique de débruitage basé sur la soustraction spectrale. L'objectif de ces techniques est la restauration instantanée du spectre du signal de parole par la soustraction d'une estimation de l'amplitude spectrale du bruit de celle du signal bruité. Le spectre d'amplitude restauré et combiné avec la phase du signal bruité pour former le signal de parole restauré.

Les trois techniques sont présentées dans ce mémoire. Nous avons limité les implémentations et les tests sous MATLAB aux techniques basées sur le filtrage de Wiener et le filtrage adaptatif au sens de Wiener basé sur l'algorithme LMS et l'algorithme LMS normalisé (NLMS). Les deux techniques sont implémentés dans le contexte de débruitage par annulation de bruit pour un signal parole, un bruit blanc et pour différentes conditions.

Ce mémoire est formé de 04 chapitres. Le premier chapitre donne des généralités sur le signal de parole et les signaux de bruit. Le deuxième chapitre présente une brève

Introduction Générale

description des quelques transformées utilisées en traitement du signal et en débruitage du signal de parole tel que la DFT, la DHT et la DCT. Le troisième chapitre est destiné à la description des trois techniques de filtrage utilisé pour le débruitage du signal de parole. Le quatrième chapitre est consacré aux différentes expérimentations et tests effectués pour évaluer les performances des techniques employées.

Une conclusion générale présente quelques remarques et conclusions.

Chapitre 1 :

Généralités sur le

Signal Parole.

Chapitre 1 : Généralités sur le Signal Parole.

1.1. Introduction :

Notre environnement est souvent bruité, les applications audio se trouvent ainsi confrontées au bruit ambiant. Ce chapitre résume les principales caractéristiques de la parole et du bruit : pour la parole, de sa production à sa Perception; pour le bruit, on citera les caractéristiques temporelles et spectrales de certains types de bruits liés aux applications de téléphonie mobile et de téléphonie mains libres à bord de véhicules.

1.2. Notion du signal :

Nous appelons signal tout variable ou source d'information évoluant en fonction du temps. Exemple : un signal $s(t)$, en électronique est une tension $v(t)$ ou un courant $i(t)$. Lorsque l'amplitude d'un signal est connue ou peut être déterminée à chaque instant, le signal est dit déterministe. Dans le cas où seule une information de nature statistique, telle la probabilité d'avoir une amplitude

donnée à un instant donné, le signal serait dit aléatoire. L'aspect aléatoire apparaît essentiellement au niveau des bruits (parasites).

Un signal $s(t)$ peut aussi se présenter sous différentes formes, selon que son amplitude est une variable continue ou discrète (quantifié) et sa variable indépendante 't' est continue ou discrète (signal échantillonné) [3].

1.3. Production et Perception du signal de parole :

1.3.1. Production du signal de parole :

Le signal de parole est le résultat de l'excitation du conduit vocal par un train d'impulsions ou un bruit donnant lieu respectivement aux sons voisés et non voisés figure 1.1. Dans le cas des sons voisés, l'excitation est une vibration périodique des cordes vocales suite à la pression exercée par l'air provenant de l'appareil respiratoire. Ce mouvement vibratoire correspond à une succession de cycles d'ouverture et de fermeture de la glotte. Le nombre de ces cycles par seconde correspond à la fréquence fondamentale F_0 . Quant aux signaux non-voisés, l'air passe librement à travers la glotte (du moins pas dans tout le conduit vocal) sans

Chapitre I : Généralités sur le signal de parole

provoquer de vibration des cordes vocales [1].

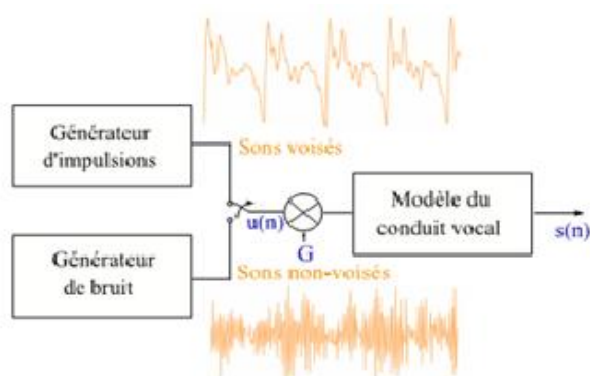


Figure 1.1 : Modèle simple de production de la parole

1.3.2. Perception de la parole :

Le signal de parole est un vecteur acoustique porteur d'informations d'une grande complexité, variabilité et redondance. Les caractéristiques de ce signal sont appelées traits acoustiques. Chaque trait acoustique a une signification sur le plan perceptuel.

Le premier trait est la fréquence fondamentale, fréquence de vibration des cordes vocales. Ses variations définissent le pitch qui constitue la perception de la hauteur (ou les sons s'ordonnent de grave à aigu). Seuls les sons quasi-périodiques (voisés) engendrent une sensation de hauteur tonale bien définie.

Le deuxième trait est le spectre fréquentiel dont dépend principalement le timbre de la voix. Le timbre est une caractéristique permettant d'identifier une personne à la simple écoute de sa voix. Le timbre dépend de la corrélation entre la fréquence fondamentale et les harmoniques qui sont les multiples de cette fréquence.

Le dernier trait acoustique est l'énergie correspondant à l'intensité sonore. Elle est habituellement plus forte pour les segments voisés de la parole que pour les segments non voisés [1].

1.4. Analyse et paramétrisation du signal parole :

Le signal de parole est un processus aléatoire non-stationnaire à long terme, mais il est considéré comme stationnaire dans des fenêtres temporelles d'analyse de l'ordre de 20 à 30ms. Cette propriété de stationnarité à court terme permet donc une analyse et modélisation progressive du signal de parole accompagnée, bien sûr, d'un chevauchement de

fenêtres pour permettre une continuité temporelle des caractéristiques de l'analyse et du modèle.

1.4.1. Analyse LPC :

Dans l'analyse par prédiction linéaire LPC, le conduit vocal est modélisé par une fonction de transfert qui suit un modèle autorégressif. Cette analyse est fort utilisée dans le codage de parole dans le but de réduire la redondance du signal vocal, ou pour extraire des paramètres pertinents pour la reconnaissance de parole.

L'estimation des coefficients de la fonction de transfert du conduit vocal est faite en supposant connaître le signal d'excitation. Pour les sons non voisés, le signal d'excitation est un bruit blanc de moyenne nulle et de variance unité. Pour les sons voisés, cette excitation est une suite d'impulsions d'amplitude unité. La fonction de transfert du conduit vocal dans le domaine Z est donnée par :

$$H(z) = \frac{S(z)}{U(z)} = \frac{G}{1 - A(z)} \quad (1.1)$$

$$\text{Ou } A(z) = \sum_{k=1}^p a_k z^{-k} \quad (1.2)$$

Est le prédicateur linéaire, a_k sont les coefficients de prédiction, $S(z)$ est le signal de parole produit en sortie, $U(z)$ est le signal d'excitation et G est un gain. Le signal de parole $s(n)$ à la sortie du modèle est donc représenté par la somme d'une combinaison linéaire des échantillons précédents et de la fonction d'excitation, tel que :

$$A(z) = \sum_{k=1}^p a_k z^{-k} S(z) + Gu(z) \quad (1.3)$$

Le modèle de prédiction exploite le fait que les échantillons successifs du signal de parole sont corrélés ; d'où l'intérêt de ce modèle dans le codage de la parole dans le sens où il permet de représenter la parole juste par ses paramètres pertinents, sans redondance. Signalons également que les coefficients sont

choisis de façon à minimiser l'erreur quadratique de prédiction sur chaque segment de la fenêtre d'analyse [1].

1.4.2. Cepstre :

Le Cepstre est basé sur une connaissance du modèle de production de la parole. Comme nous l'avons vu dans la section précédente, une modélisation du signal de parole consiste à définir ce signal comme le résultat de la convolution de la fonction de transfert du conduit vocal (filtre) par un signal d'excitation (source). Le but du cepstre est de séparer ces deux contributions (source et filtre) par application de la déconvolution à travers une transformée en cosinus discret. Le processus de calcul du cepstre est le suivant où s , u et h le signal de parole, le signal d'excitation (source) et la fonction de transfert du conduit vocal (filtre),

$$s = u * h \tag{1.4}$$

$$\begin{aligned} s &= u * h \\ TFD(s) &= UH \end{aligned} \tag{1.5}$$

Le logarithme de l'amplitude transforme le produit de la TFD en somme.

On obtient alors :

$$\begin{aligned} \text{Log } |S(\nu)| &= \log |U| + \log |H| \end{aligned} \tag{1.6}$$

Par transformation en cosinus discret (DCT), on obtient le cepstre.

L'expression du cepstre réel est donc :

$$c = DCT(\log(TFD(s))) \tag{1.7}$$

L'espace fréquentiel de représentation du cepstre est équivalent à un espace temporel. A partir du cepstre (Fig 1.2), il est possible de définir la fréquence fondamentale de la source u en détectant les pics périodiques (harmoniques) au-delà d'un certain nombre N de coefficients. En effet, les N premiers points du cepstre contiennent l'information la plus pertinente sur le spectre et permettent d'obtenir un spectre lisse, débarrassé des harmoniques dus à la contribution de la source. Cependant, déterminer la fréquence fondamentale d'un signal de parole reste encore un problème difficile. En effet, les algorithmes classiques manquent de robustesse quand le bruit est présent, quand la fréquence fondamentale change rapidement ou quand la valeur de celle-ci n'est pas assez élevée [1].

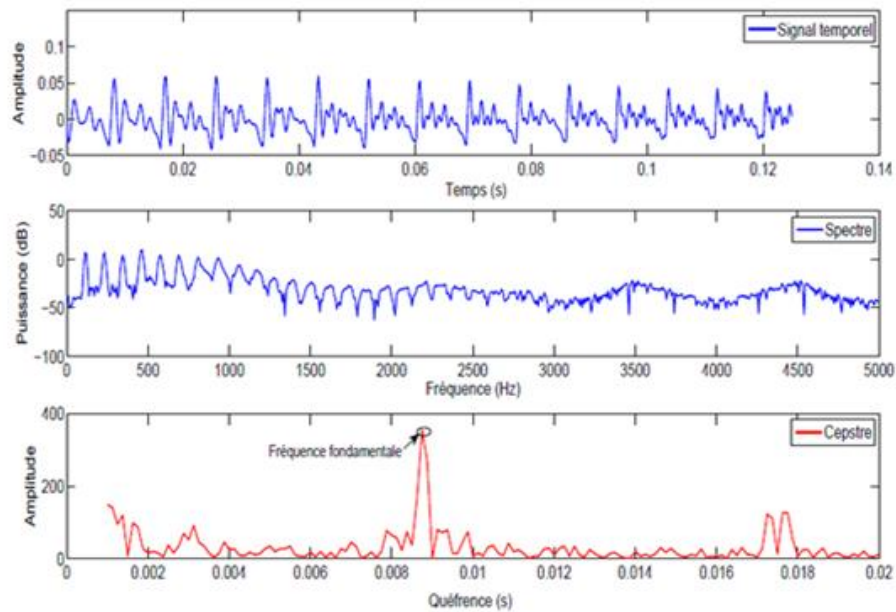


Figure 1.2 : Détermination de la fréquence fondamentale par Cepstre.

L'analyse cepstrale est basée sur un calcul de coefficients dits coefficients cepstraux de Mel, soit en abrégé MFCC (Mel Frequency Cepstral Coefficients). Le calcul est en effet basé sur une échelle de Mel. Cette échelle se rapproche de la perception fréquentielle de l'oreille. L'idée est de moyenniser le spectre dans des bandes de fréquence correspondant grossièrement au filtrage effectué par la membrane basilaire. L'échelle Mel est approchée par un banc de 15 à 24 filtres triangulaires espacés linéairement jusqu'à 1 KHz, puis espacés logarithmiquement jusqu'aux fréquences maximales. Elle a été conçue de telle façon que 1000 Hz correspondent à 1000mels. La formule de conversion :



Figure1.3 : Chaîne de calcul des coefficients MFCC

Chapitre I : Généralités sur le signal de parole

Du hertz en mel la plus utilisée est la suivante :

$$mel = 2595 \log_{10} \left(1 + \frac{f \text{ hz}}{700} \right) \quad (1.8)$$

La procédure de calcul pas à pas des MFCC est la suivante:

- Découpage en trame (stationnarité), chevauchement (éviter les transitions brusques de trame en trame).

- Pré-accentuation (pour donner plus d'énergie et renforcer la contribution des hautes fréquences) avec un filtre passe-haut de la forme $1-0.9z^{-1}$ puis fenêtrage de Hanning (pour la continuité aux bords). [1]

- Calcul de la TFD sur chaque trame.

- Filtrage par un banc de filtres triangulaires repartis le long de l'échelle de Mel.

- Calcul du logarithme du module de l'énergie en sortie du banc de filtres.

- Application de la Transformée en Cosinus Discrète inverse (joue le rôle d'une TFD inverse).

- Seuls les premiers coefficients sont conservés. Par exemple, dans la reconnaissance de la parole par HTK, on utilise les 12 premiers coefficients. Le premier coefficient c_0 représente l'énergie mais peut être remplacé par le log de l'énergie moyenne des échantillons. [1]

1.4.3. Application :

Le cepstre d'un signal est utilisé en traitement du son et de la parole, en reconnaissance vocale et en maintenance conditionnelle par la surveillance du comportement vibratoire des machines industrielles (ex: alternateurs de centrales électriques).

En traitement du son et plus spécifiquement le son de la parole, le signal est modélisé comme étant le résultat d'un filtre sur une excitation périodique. La représentation cepstrale permet de dissocier la source du filtre pour estimer la fréquence fondamentale ou les formants.[2]

1.5. Le signal du bruit :

1.5.1. Définition de bruit :

Le bruit est un signal perturbateur provenant du canal lui même ou de son environnement externe. Il est de comportement aléatoire et vient s'ajouter au signal véhiculant les informations et provoquer ainsi les erreurs de transmission (on va voir dans cette mémoire comment protéger l'information contre ce bruit). On distingue généralement un type de bruit : le bruit blanc. [3]

Ainsi, il apparaît évident qu'un problème fondamental en traitement du signal sera d'extraire le signal utile du bruit. La difficulté du problème dépend en particulier de la proportion entre signal et bruit. Ceci est mesuré par le rapport signal à bruit (RSB, ou SNR en anglais pour signal noise ratio). [3]

1.5.2. Nature et caractéristiques du bruit

On appelle bruit tout signal nuisible qui se superpose au signal utile en un point quelconque d'une chaîne de mesure ou d'un système de transmission. Il constitue donc une gêne dans la compréhension du signal utile, qui est dans notre cas, la parole. En physique, en acoustique et en traitement du signal, bien que le bruit soit, par nature, aléatoire, il possède certaines caractéristiques statistiques, spectrales ou spatiales. Le tableau (1.1), extrait de, représente les différentes classes auxquelles un bruit peut appartenir.

Tableau (1.1) : Différentes classes du bruit.

Propriétés	Types
Structure	Continu / Impulsif / Périodique
Type d'interaction	Additif/Multiplicatif / Convolutif
Comportement temporel	Stationnaire / Non stationnaire
Bande de fréquence	Étroite / Large

Chapitre I : Généralités sur le signal de parole

Dépendance	Corrélé / Décorrélé
Propriétés statiques	Dépendant / Indépendant
Propriétés spatiales	Cohérent / Incohérent

Comme notre but est essentiellement le débruitage et la reconnaissance de la parole pour des applications de téléphonie mobile et de téléphonie mains libres à bord de véhicules, on se limite dans notre étude aux bruits additifs, stationnaires ou non stationnaires, et décorrélés avec la parole (indépendance au sens statistique), tels que le bruit de conversation appelé Babble, le bruit de voiture appelé (car) et le bruit blanc gaussien (ce dernier est souvent utilisé mais peu réaliste) [1].

1.5.3. Les difficultés rencontrées lors de l'utilisation de la parole

La dégradation d'un signal vocal par le bruit ambiant engendre des problèmes perceptifs ennuyants pour l'oreille humaine et détériore les performances des systèmes de communication homme machine.

La réduction du bruit constitue un des problèmes majeurs, qui a donné naissance à un grand besoin de développer des algorithmes capables d'éliminer les signaux indésirables dégradant le signal vocal et dont l'objectif essentiel est d'améliorer l'aspect perceptuel et l'intelligibilité du signal vocal.

1.5.4. Qu'est-ce qu'un bruit blanc ?

Un bruit blanc est une réalisation d'un processus aléatoire dans lequel la densité spectrale de puissance est la même pour toutes les fréquences.

Le bruit blanc gaussien est un bruit blanc qui suit une loi normale de moyenne et variance données.

En synthèse et traitement du son, on ne considère que les fréquences comprises entre 20 Hz et 20 kHz puisque l'oreille humaine n'est sensible qu'à cette bande de fréquences (la sensibilité varie toutefois selon les personnes). L'impression obtenue est celle d'un souffle. Le son produit lors de l'effet de « neige » sur un téléviseur déréglé est un bon exemple de

bruit blanc [5].

1.5.5. Qu'est ce qu'un bruit musical :

Le bruit musical est un bruit résiduel perceptuellement gênant qui apparaît suite au débruitage de la parole par des algorithmes d'atténuation spectrale à court terme tels que la soustraction spectrale ou le filtrage de Wiener. Le spectre du bruit musical est particulièrement tonal, d'où le caractère musical. Son énergie moyenne est plus faible que celle du bruit initial, mais sa dispersion en fréquence est plus grande, ce qui le rend plus gênant que le bruit de départ du point de vue perception. Parmi les raisons d'apparition de ce type de bruit, on cite :

- Le traitement non linéaire des composantes négatives du signal débruité;
- l'estimation non précise de la densité spectrale de bruit;
- l'estimation basée sur des périodogrammes;
- la variabilité de la fonction de gain appliquée au signal bruité;
- la variance des estimateurs locaux de la densité spectrale des signaux... [1].

1.5.6. Les différents types de bruit

Les différents bruits pouvant influer sur un message peuvent être divisés en deux grandes catégories : les bruits additifs et les bruits convolutionnels. La distinction entre les deux peut être faite par le nombre d'agents agresseurs extérieurs à la transmission du message. Les bruits additifs sont causés par des agents extérieurs au trinôme source-voie-destinataire alors que les bruits convolutionnels sont causés par la moindre qualité de la voie de communication, celle-ci ayant alors un rôle ambigu, du point de vue du message, de médium et d'agresseur, l'effet est plus dévastateur si le bruit n'est pas stationnaire [1].

1.6. Rapport signal sur bruit (RSB) ou (SNR) :

La quantité de bruit présente sur un canal de transmission, est exprimé par le rapport de la puissance du signal transmis sur la puissance de bruit et prend le nom de rapport signal sur bruit, nous écrivons ce rapport S/B sous la forme :

$$\frac{S}{B} = \frac{P_S}{P_B} = \left(\frac{U_S}{U_B}\right)^2 \quad (1.9)$$

Chapitre I : Généralités sur le signal de parole

P_S : puissance du signal en W

P_B : puissance de bruit en W

U_S : tension efficace du signal en V

U_B : tension du bruit en V

Et on a coutume de l'exprimer sous la forme : $10.\log(S/B)$ en décibels (dB). Ce rapport varie dans le temps, puisque le bruit n'est pas uniforme, toutefois on peut en estimer une valeur moyenne sur un intervalle de temps. Le rapport signal sur bruit est aussi une caractéristique d'un canal de transmission [5] [3].

1.7.Conclusion :

Ce chapitre rappelle les notions de base pour les signaux paroles et de bruit, comme la Production & Perception du signal parole, suivie par des généralités sur le signal de bruit comme le bruit blanc et le rapport signal sur bruit (SNR), et quel que caractéristiques pour ce signal.

Chapitre 2 :
Les Transformées
Discrètes
Orthogonales.

Chapitre 2 :

Les Transformées Discrètes Orthogonales.

2.1. Introduction :

La transformée de Fourier discrète est la transformée de Fourier « exacte » d'un signal périodique et discret. Elle est très simple à calculer à partir de séries mathématiques limitées et ce calcul s'implante facilement sur calculateur ou circuit spécialisé (DSP) avec un algorithme FFT (Fast Fourier Transform) permettant d'en accélérer le temps de calcul de plusieurs centaines de fois.

Moyennant quelques précautions d'emploi, elle permet d'approximer en un temps record la transformée de Fourier d'un signal continu à partir de sa version échantillonnée d'où le grand intérêt de cette transformation pour les ingénieurs, scientifiques et traiteurs de signaux.

La transformation Hartley a été introduite pour la première fois par Hartley en 1942, sa discrétisation est plus récente, cette transformation permet d'associer une séquence réelle à une autre séquence réelle de même longueur, et semble mieux adaptée a priori, que la transformation de Fourier, à l'étude des signaux réels.

La transformée en cosinus discrète ou TCD (de l'anglais : DCT ou *Discrete Cosine Transform*) est une transformation proche de la transformée de Fourier discrète (DFT). Le noyau de projection est un cosinus et crée donc des coefficients réels, contrairement à la DFT, dont le noyau est une exponentielle complexe et qui crée donc des coefficients complexes. On peut cependant exprimer la DCT en fonction de la DFT, qui est alors appliquée sur le signal symétrisé.

La variante la plus courante de la transformée en cosinus discret est la DCT type-II, souvent simplement appelée « la DCT ». Son inverse, qui correspond au type-III est souvent simplement appelée « IDCT ».

2.2. La transformation de Fourier discrète (TFD) :

2.2.1. Définition :

La transformée de Fourier discrète (TFD) est un outil mathématique de traitement du signal numérique, qui est l'équivalent discret de la transformation de Fourier continue qui est utilisée pour le traitement du signal analogique.

Soit f une fonction de variable réelle à valeurs réelles ou complexes :

$$\begin{array}{l} \text{réelle à valeur} \\ f : \mathbb{Z} \rightarrow \mathbb{C} \\ x \mapsto f(x) \end{array} \quad (2.1)$$

On appelle transformée de Fourier (ou TF) de f l'intégrale :

$$F(v) = \int_{-\infty}^{+\infty} f(t) e^{-2i\pi vt} dt \quad (2.2)$$

Avec v réel. Cette intégrale existe si f est intégrable sur \mathbb{R} . On parle d'espace direct pour décrire (t) , et d'espace de Fourier pour décrire (v) .

Il ne faut pas confondre avec la transformation de Fourier rapide qui est un algorithme particulier de calcul de la transformation de Fourier discrète. On appelle transformée de Fourier discrète d'une suite de N termes $s(0), s(1), \dots, s(N-1)$, la suite de N termes $S(0), S(1), \dots, S(N-1)$. Sa définition mathématique pour un signal s de N échantillons est la suivante :

$$S(k) = \sum_{n=0}^{N-1} s(n) \cdot e^{-2i\pi k \frac{n}{N}} \quad \text{pour} \quad 0 \leq k \leq N-1 \quad (2.3)$$

La transformation inverse est donnée par :

$$s(n) = \frac{1}{N} \sum_{k=0}^{N-1} S(k) \cdot e^{2i\pi n \frac{k}{N}} \quad \text{pour} \quad 0 \leq n \leq N-1 \quad (2.4)$$

On obtient ainsi une représentation spectrale discrète du signal échantillonné $s(n)$. Il est important de comprendre que la TFD ne calcule pas le spectre continu d'un signal continu. La TFD permet seulement d'évaluer une représentation spectrale discrète (spectre échantillonné) d'un signal discret (signal échantillonné) sur une fenêtre de temps finie (échantillonnage borné dans le temps) [6].

2.2.2 Réalisation pratique :

Pour calculer la DFT un algorithme de calcul rapide appelé FFT (Fast Fourier Transform) a été développé. Cet algorithme est particulièrement performant dans le cas où $N = 2^M$ ou M est un nombre entier. En utilisant cet algorithme pour $N = 1024$, le temps de calcul est divisé par un facteur d'environ 1000 par rapport à l'utilisation directe de la formule (01). Cet algorithme est implanté sur les ordinateurs (processeurs actuels) et son exécution dure moins d'un micro seconde (μs).

Chapitre II : Les Transformées Discrètes Orthogonales.

2.2.3 TFD de signaux classiques :

Tableau (2.1) : Quelques signaux et leur TFD [6].		
$x^k = \frac{1}{N} \sum_{n=-N/2}^{(N/2)-1} X_n \cdot e^{2i\pi n \frac{k}{N}}$	$X^n = \sum_{k=0}^{N-1} x_k \cdot e^{-2i\pi k \frac{n}{N}}$	Note
$y^k = x^k \cdot e^{-2i\pi \frac{k}{N}}$	$y^n = x^{n-1}$	Propriété de translation
$y^k = x^{k-n}$	$y^n = x^n \cdot e^{-2i\pi n \frac{k}{N}}$	
$\begin{matrix} y \in \mathbb{R} \\ x^n \in \mathbb{R} \end{matrix}$	$\begin{matrix} Y_1 = X_1 \cdot e^{-2i\pi \frac{k}{N}} \\ X^n = X^{N-n} \end{matrix}$	TFD d'un signal réel
$\frac{e^{-i\pi k}}{a^n}$	$1 - \frac{e^{-2i\pi n/N}}{a}$	
$\frac{1}{N-1} k$	$\left(1 + \frac{e^{-2i\pi n/N}}{a}\right)^{N-1}$	

2.2.4. Analyse spectrale :

L'analyse spectrale des signaux est un élément essentiel en électronique pour de nombreuses raisons parmi lesquelles on peut citer :

- déterminer la largeur de bande de fréquence occupée par une transmission ;
- évaluer les distorsions harmoniques apportées par le traitement des signaux ;
- mesurer les filtres.
- L'électronicien qui a toujours besoin de vérifier expérimentalement, a besoin

d'un outil de mesure, l'analyseur de spectre. Il existe trois grandes familles d'analyseur de spectre, chacun ayant des caractéristiques intrinsèques :

2.2.4.a L'analyseur de spectre à balayage (analogique) :

Comme son nom l'indique, cet analyseur balaye une plage de fréquence en utilisant un filtre de largeur réglable. Il est capable de mesurer des plages de fréquence allant de l'audio à l'optique et ce pour des signaux d'amplitude très faible.

2.2.4.b L'analyseur de spectre à FFT (numérique) :

La FFT (Fast Fourier Transform ou transformation de Fourier rapide) est ici utilisée après échantillonnage du signal d'entrée basses fréquences (audio). Avantage : il est capable

Chapitre II : Les Transformées Discrètes Orthogonales.

de capturer les signaux en temps réel avec une résolution spectrale très fine qui dépend du nombre de points N et de la fenêtre de pondération utilisée.

L'augmentation de la rapidité et de la résolution des convertisseurs analogique numérique permettra d'analyser des signaux à des fréquences de plus en plus élevées.

2.2.4.c L'analyseur de signaux vectoriel (analogique/ numérique) :

Comme il combine les technologies des deux premiers (balayage et FFT), il permet d'analyser des signaux dont les fréquences ne sont séparées que de quelques MHz sur toute la gamme de fréquences radio. Très utilisé dans le domaine des transmissions numériques pour analyser des signaux complexes (QAM, QPSK,)

2.2.4. Analyse de séries temporelles :

La TFD est utilisée pour l'étude des séries temporelles (ou chronologiques) où le but est de trouver des corrélations entre deux séquences de données. Un exemple classique est l'analyse des cours de la bourse, afin de repérer des événements particuliers. La problématique est en général celle de la fouille de données, ou de la recherche par similarité. La TFD est utilisée ici comme un moyen de réduire la dimensionnalité du problème. La TFD permet en effet de décorréler les données de départ et de ne travailler que sur un petit nombre de coefficients significatifs. [10]

2.2.5. Application de la TFD :

Le principal usage de la TFD est lié à la mise en évidence des caractéristiques spectrales des signaux, le plus souvent analogiques. Nous allons voir qu'il est en effet possible d'estimer la transformée de Fourier d'un signal analogique quelconque (périodique ou pas) à l'aide de la TFD.

La TFD est utilisée dans un large spectre d'applications, seules les plus communes sont listées ici. Toutes ces applications nécessitent l'existence d'un algorithme rapide de calcul de la TFD et de son inverse.

On a montré que la TFD du signal échantillonné donnait presque exactement des échantillons de la TF continue du signal continu. Rappelons que l'échantillonnage nécessite que la bande spectrale du signal soit limitée, ce qui n'est pas rigoureusement possible pour un signal à durée limitée. Il y aura donc toujours un léger repliement, que l'on minimisera par une fréquence d'échantillonnage ν_e suffisamment grande [7].

2.2.6 Quelques propriétés de la TFD :

Les propriétés de la TFD sont les mêmes que celles de la TFTD, à condition de supposer que le signal d'entrée est périodique de période égale à N échantillons

Les propriétés de la TFD sont présentées au Tableau si dessus :

Chapitre II : Les Transformées Discrètes Orthogonales.

Tableau (2.2) : Propriétés principales de la transformée de Fourier à temps discret [7].

	X[n]	X(f)
Linéarité	$\alpha x(n) + \beta y(n)$	$\alpha X(f) + \beta Y(f)$
Translation	$x(n-k)$	$e^{-j2\pi k f} X(f)$
	$e^{j2\pi n f_0} x(n)$	$X(f - f_0)$
Convolution	$x(n) * y(n)$	$X(f) \cdot Y(f)$
	$x(n) \cdot y(n)$	$X(f) * Y(f)$

2.3. Transformée Hartley discrète (DHT):

2.3.1 Définition :

Une transformée de Hartley discrète (DHT) est une transformée de Fourier liée à une transformation de données discrètes périodiques semblables à la transformée de Fourier discrète (TFD), avec des applications analogiques dans le traitement de signal.

Soit $s(n)$, une séquence de N échantillons d'un signal réel, on peut lui associer une suite de N échantillons $S(k)$ par

$$S(k) = \sum_{n=0}^{N-1} s(n) \cdot \left(\cos\left(\frac{2\pi nk}{N}\right) + \sin\left(\frac{2\pi nk}{N}\right) \right) \quad \text{Avec } k = 0, 1, \dots, N-1. \quad (2.5)$$

De même la transformation inverse permet de retrouver la séquence $s(n)$ par :

$$s(n) = \frac{1}{N} \sum_{k=0}^{N-1} S(k) \cdot \left(\cos\left(\frac{2\pi nk}{N}\right) + \sin\left(\frac{2\pi nk}{N}\right) \right) \quad \text{Avec } n = 0, 1, \dots, N-1. \quad (2.6)$$

2.3.2. Relation avec la transformation de Fourier discrète :

Soit :

Soit :

$$F(k) = R(k) + jI(k) \tag{2.7}$$

la TFD de même séquence $s(n)$

la TFD de même séquence $s^*(n)$

$$F(k) = \sum_{n=0}^{N-1} S(n) \cdot \left(\cos\left(\frac{2\pi nk}{N}\right) + j \sin\left(\frac{2\pi nk}{N}\right) \right) \quad \text{Avec} \quad k = 0, 1, \dots, N-1. \tag{2.8}$$

On obtient alors assez facilement les relations suivantes avec les composantes de la THD (Transformée de Hartley discrète) :

$$R(k) = \frac{S(k) + S(N-k)}{2} \tag{2.9}$$

$$I(k) = \frac{S(k) - S(N-k)}{2} \tag{2.10}$$

$$R^2(k) + I^2(k) = \frac{X^2(k) + X^2(N-k)}{2} \tag{2.11}$$

Les transformées discrètes ont invariablement joué un rôle considérable dans divers sujets, particulièrement en Génie. Un exemple intéressante est la transformée discrète de Fourier, dite DFT. Une autre transformée assez riche et profondément connectée à la DFT est la transformée discrète de Hartley, la DHT, une version discrète de la transformée intégral symétrique du type Fourier, introduite par Ralph Hartley. D'ailleurs ses avantages côté numérique, la DHT s'est montrée un outil assez puissante depuis des années. Un fait décisif problèmes où la DFT s'avère utile est l'existence des algorithmes dite rapides (FT) envisagés pour son calcul efficace. La transformée rapide de Hartley existe elle aussi et est intimement liée aux applications de la DHT. Quelques situations pratiques récemment proposées pour les transformées discrètes concernent l'emploi de la transformée de Hartley de corps fini dans l'esquisse de systèmes de multiplexage numériques, systèmes à accès multiple et dans le projet de séquences multi état d'étalonnement spectral. Par-dessus d'être une transformée réel, la DHT est involutionnaire dans le sens qui les noyaux de la transformée inverse est le même de la transformée directe. Étant donné que la DHT est une version symétrique d'une transformée discrète, cette symétrie là peut être prise en avantage de façon à en déduire un nouveau algorithme rapide qui demande un nombre minimal de multiplications réelles à point flottant. Un algorithme rapide pour la DHT l'est aussi pour la DFT et vice versa Transformées

Chapitre II : Les Transformées Discrètes Orthogonales.

discrètes ayant un faible complexité multiplicative ont été un sujet d'intérêt de depuis longtemps. La complexité minimale multiplicative d'une transformée unidimensionnelle DFT pour une longueur donnée peut être calculée par la conversion de la DFT en un ensemble de convolutions cycliques multidimensionnelles. Dans cet article, un algorithme rapide est déduit lequel atteint la borne inférieure sur le nombre de multiplications d'une DFT ou DHT tout au moins pour des courtes longueurs de bloc. Il est fondée sur une décomposition multicouche de la DHT en faisant appel à la transformée de Walsh-Hadamard. Ces nouveaux schémas sont assez performant et peuvent être implémentés sur des multiprocesseurs numériques de signaux DSP et la régularité de son structure peut admettre le projet d'un circuit intégré de haut vitesse à bas coût [9].

2.3.3. Similarité :

Cette propriété permet d'exprimer simplement la THD d'une séquence $s(n)$ de longueur $2N$ obtenue en intercalant un zéro entre chaque échantillon de la séquence $s(n)$:

$$\begin{array}{ccc}
 \begin{array}{c} 2N \text{ obtenue en intercalant} \\ [s(0), s(1), \dots, s(N-1)] \end{array} & \xrightarrow{\text{mer simple et zéro entre chaque échantillon}} & \begin{array}{c} \text{THD} \\ [s(0), s(1), \dots, s(N-1)] \end{array} \\
 \end{array} \quad (2.12)$$

$$\begin{array}{ccc}
 \begin{array}{c} [s(0), s(1), \dots, s(N-1)] \\ [s(0), 0, s(1), 0, \dots, s(N-1), 0] \end{array} & \xrightarrow{\text{THD}} & \begin{array}{c} [s(0), \dots, s(N-1)] \\ [s(0), \dots, s(N-1)] \end{array} \\
 \end{array} \quad (2.13)$$

Si, comme pour la TFD, on utilise une longueur N qui soit une puissance de 2, on peut alors introduire une procédure de décomposition identique à celles de décimation en temps ou en fréquence [9].

2.4. La transformée en cosinus discrète (DCT) :

2.4.1. définition :

La DCT est une fonction linéaire inversible $\mathbb{R}^N \rightarrow \mathbb{R}^N$ ou de manière équivalente une matrice carrée $N \times N$ inversible. Il existe plusieurs légères variantes de la DCT. Voici les quatre types les plus connus.

2.4.2. Les propriétés de la DCT :

Le développement des algorithmes de calcul rapide des transformées DCT se basent sur la possibilité de décomposer la matrice de définition sous forme d'un produit de matrices dont le calcul est plus simple, et permet de réduire le nombre de multiplications scalaires, en profitant des identités remarquables de périodicité et symétries des fonctions sinusoïdales. Ainsi, on peut décomposer toute transformée DCT de \mathbb{R}^N en transformées plus simples en décomposant N en produit de facteurs premiers, et en composant des sous-transformées

Chapitre II : Les Transformées Discrètes Orthogonales.

dans \mathbb{R}^n où n est l'un de ces facteurs. En particulier, de nombreuses optimisations ont été développées quand N est une puissance de 2.

Cela revient à réécrire la matrice $N \times N$ sous forme de produit de sous-matrices identiques (disposées en pavage régulier et utilisant donc des coefficients réels communs ou différenciés uniquement par leur signe) et de matrices à coefficients unitaires ou nuls (-1, 0 ou 1), ces dernières ne nécessitant pas de multiplication.

La transformée en cosinus discrète dite en abrégé DCT (Discrète cosinus transform), est la transformée de Fourier d'une séquence réelle paire. Il s'agit donc d'une transformation réelle. Lorsque l'on considère le signal d'image comme un processus stationnaire au sens large, on peut montrer que la DCT est une bonne approximation de la KLT (transformée de Karhunen-Loève). De plus elle possède un ensemble de fonction base stable (les fonctions cosinus), ce qui n'est pas le cas de la TKL.

La transformation de Karhunen-Loève (KLT), est une méthode de la famille de l'analyse des données et plus généralement de la statistique multi variée, qui consiste à transformer des variables liées entre elles (dites "corrélées" en statistique) en nouvelles variables décorrélées les unes des autres. Ces nouvelles variables sont nommées "composantes principales", ou axes principaux. Elle permet au praticien de réduire le nombre de variables et de rendre l'information moins redondante.

Il s'agit d'une approche à la fois géométrique (les variables étant représentées dans un nouvel espace, selon des directions d'inertie maximale) et statistique (la recherche portant sur des axes indépendants expliquant au mieux la variabilité — la variance — des données). Lorsqu'on veut compresser un ensemble de N variables aléatoires, les n premiers axes de l'analyse en composantes principales sont un meilleur choix, du point de vue de l'inertie ou de la variance.

à pour but de transformer un tableau de données de i individus et de r réalisations pour pouvoir en donner une meilleure interprétation. L'idée de la transformation est de réaliser un changement de base pour obtenir de nouveaux axes où l'information contenue sur chaque axe est distribuée de façon optimale, c'est-à-dire de façon à éloigner au maximum le comportement des individus pour pouvoir mieux cerner leur comportement. Dans le cas d'une image décomposée en RGB [8].

2.4.3. Application :

La DCT, et en particulier la DCT-II est très utilisée en traitement du signal et de l'image, et spécialement en compression. La DCT possède en effet une excellente propriété de « regroupement » de l'énergie : l'information est essentiellement portée par les coefficients

Chapitre II : Les Transformées Discrètes Orthogonales.

basses fréquences. Pour les images naturelles, la DCT est la transformation qui se rapproche le plus de la transformée de Karhunen-Loève qui fournit une décorrélation optimale des coefficients pour un signal markovien. En pratique, les procédés de compression font donc l'hypothèse qu'une image naturelle peut être modélisée comme la réalisation d'un processus markovien et approximent la transformée de Karhunen-Loève, trop complexe en calcul et dépendante des données, par une DCT. L'intérêt d'une transformation se voit particulièrement bien sur une figure (voir ci-contre). Seuls un petit nombre de coefficients sont non-nuls, et peuvent être utilisés pour reconstruire l'image par transformée inverse (IDCT) lors de la décompression. Le gain en termes de compression vient de la suppression des coefficients nuls ou proches de zéro. Ce genre de mécanisme est utilisé dans les standards JPEG et MPEG, qui utilisent une DCT 2D sur des blocs de pixels de taille 8x8 (pour des raisons de complexité).

Les formats de compression de son avec perte AAC, Vorbis et MP3 utilisent une version modifiée de cette technique, la transformée en cosinus discrète modifiée, TCDM (MDCT en anglais).

La DCT est aussi employée pour la résolution de systèmes d'équations différentielles par des méthodes spectrales [8].

5. Conclusion :

La transformée de Fourier Discrète (TFD) est la manière rigoureuse de calculer la transformée de Fourier d'un signal discret. La TFD et sa transformation inverse permettent de relier les échantillons $\{s_k\}$ du motif du signal discret aux échantillons $\{S_n\}$ du motif de sa transformée de Fourier discrète.

Chapitre 3 :

Les techniques de rehaussement du signal de parole.

Chapitre 3 :

Les techniques de rehaussement du signal de parole.

3.1. Introduction :

Dans ce chapitre nous discutons les techniques de débruitage de la parole par annulation de bruit. Ces techniques sont basées sur des algorithmes de filtrage, parmi lesquels, nous discutons dans ce chapitre : le filtrage de Wiener, le filtrage adaptatif au sens de Wiener et le filtrage basé sur la soustraction spectrale.

3.2. La soustraction spectrale :

Dans cette soustraction spectrale une estimation de l'amplitude spectrale du signal est obtenue par la soustraction d'une estimation de l'amplitude spectrale du bruit de celle du signal bruité. La figure 3.1 donne une illustration de la configuration de la technique de la soustraction spectrale. Dans la soustraction spectrale le signal d'entrée $x(m)$, est divisé en trames de longueur de N échantillons. Chaque trame est fenêtrée, en utilisant une fenêtre (par exemple : hanning) et par la suite transformée par une transformée de Fourier discrète de N échantillons. Les fenêtres diminuent les effets des discontinuités au niveau des points limites de chaque segment. Le signal fenêtré est donnée par :

$$\begin{aligned}x_w(m) &= w(n)y(m) \\ &= w(n)[x(m) + n(m)] \\ &= [x_w(m) + n_w(m)]\end{aligned}\tag{3.1}$$

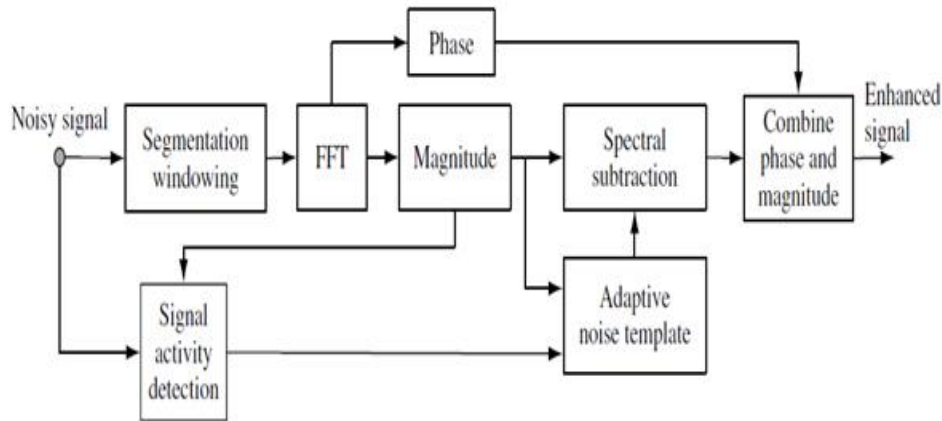


Figure 3.1. block diagramme d'un système de soustraction spectrale basé sur la FFT pour le débruitage de la parole.

L'opération du fenêtrage peut être exprimé dans le domaine fréquentielle comme :

$$\begin{aligned} \hat{Y}^w(f) &= W(f) * Y(f) \\ &= X^w(f) + N^w(f) \end{aligned} \quad (3.2)$$

Où * désigne l'opération de convolution. A travers ce chapitre, il est assuré que les signaux sont fenêtrés, pour la simplicité nous introduisant l'indice 'w' pour indiquer que les signaux correspondants est fenêtrés.

L'équation décrivant la soustraction spectrale doit être exprimée comme :

$$\hat{X}_k^b = \hat{Y}_k^b - \alpha(k) \hat{N}_k^b \quad (3.3)$$

Où \hat{X}_k^b est une estimation de l'amplitude spectrale du signal a la puissance de 'b' et \hat{N}_k^b est l'amplitude (time_ouveroyed magnitude) temporelle moyenne de spectre

Chapitre III : Les techniques de rehaussement du signal de parole

de bruit a la puissance de b . il est assuré que le bruit est un processus aléatoire stationnaire .pour la soustraction spectrale de l'amplitude, l'exposant $b=1$,et pour la soustraction spectrale de la puissance $b=2$. Le paramètre $\alpha(k)$ dans l'équation (3.3)control la quantité de bruit soustraite du signal bruité . Pour une soustraction du bruit complète, $\alpha(k)=1$ est pour une sur soustraction $\alpha(k)>1$.le spectre moyenne du bruit est obtenu a partir des périodes lorsque le signal est absent est seulement le bruit est présent comme :

$$\overline{N_{k,i}^b} = \frac{1}{M} \sum_{i=0}^{M-1} \overline{N_{k,i}^b} \quad (3.4)$$

Dans l'équation (3.4), $N_{k,i}$ est le spectre de l'équation i^{eme} trouve le bruit a la fréquence discrète k , et il est assuré qu'il ya M trame dans les périodes bruités seulement. En d'autre part, l'amplitude spectrale moyenne du bruit peut être obtenue comme la sortie d'un filtre numérique du 1^{er} ordre :

$$\frac{b}{N_{k,i}} = \frac{b}{aN_{k,i-1}} + (1 - \frac{b}{aN_{k,i}}) \quad (3.5)$$

Lorsque le coefficient 'a 'du filtre pas bas est typique ajusté entre 0.85 et 0.09 pour la restauration du signal dans le domaine temporel l'estimation de l'amplitude spectral \hat{X}_k est combinée avec la phase du signal bruité , et par la suite transformée au domaine temporel en utilisant la transformée de Fourier discret inverse comme :

$$\hat{x}(m) = \sum_{k=0}^{N-1} (\hat{X}_k e^{j\gamma k}) e^{j(\frac{2\pi}{N})km} \quad m = 0, \dots, N-1 \quad (3.6)$$

Ou $\theta_{\gamma k}$ est la phase de la fréquence du signal bruité , $\gamma(k)$ l'équation de la restauration du signal (3.6), est basé sur l'hypothèse que le bruit audible est principalement due la distorsion de l'amplitude spectrale ,et que la distorsion de la phase est largement inaudible les évaluation des effet perceptuels des distorsion de la phase simulé confirme cette hypothèse a cause des variation du spectre du bruit, la restauration spectrale peut résulté dans des estimation négative de l'énergie ou de l'amplitude spectrale. Le résultat est plus probable comme le rapport signal sur bruit diminue .pour éviter des estimation négative de l'amplitude

Chapitre III : Les techniques de rehaussement du signal de parole

de la sortie de la soustraction spectrale est post-traiter en utilisant fonction du Mapping T[.] de la forme :

$$T[X_k] = \begin{cases} \hat{X}_k & \text{si } \hat{X}_k > \beta \gamma_k \\ f_n[\gamma_k] & \text{autrement} \end{cases} \quad (3.7)$$

Par exemple, nous pouvons choisir une règle tel que si l'estimation $\hat{X}_k > 0.01\gamma_k$ (dans le spectre d'amplitude 0.01 est équivalente à 40db) puis \hat{X}_k doit être ajusté à une certaine fonction du signal traité $f_n[\gamma_k]$ dans ce plus simple forme, $f_n[\gamma_k]$ = noise floor ou noise floor est une constante positive

Une autre option de choix est $f_n[\gamma_k] = \beta \gamma_k$ dans ce cas ,

$$T[X_k] = \begin{cases} \hat{X}_k & \text{si } \hat{X}_k > \beta \gamma_k \\ \beta \gamma_k & \text{autrement} \end{cases} \quad (3.8)$$

La soustraction spectral peut être implémenter dans le domaine spectral de l'énergie ou de l'amplitude les deux méthodes sont similaires, en même temps il résulte théoriquement de peut de différence acceptable.

nous avons proposé une nouvelle technique de réduction du bruit qui consiste à intégrer une analyse du filtres au processus de suppression du bruit et on peut améliorer la qualité et l'intelligibilité du signal vocal rehaussé.

3.3. Filtrage au sens de Wiener :

3.3.1 Filtrage de Wiener :

Dans de nombreuses applications, les signaux temporels sont entachés d'une interférence ou d'un bruit non désirés. Il faut alors trouver une solution permettant de supprimer ou tout au moins.

Réduire ces composantes perturbatrices. Dans le cas où le spectre du signal désiré et celui du Signal perturbateur se superposent, il n'est pas possible de recourir au filtrage classique. Le filtre de Wiener apporte une solution à ce problème lorsque le processus est stationnaire. [11]

Le filtrage optimal de Wiener est schématisé par la figure 1.

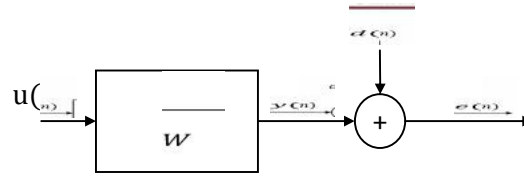


Figure 3.2. Schéma synoptique d'un filtre.

à travers ce chapitre nous supposons que tous les signaux sont de nature réels. On dispose d'une série d'observations :

$$u(n) = [u(n), u(n-1), \dots, u(n-N+1)]^T \quad (3.7)$$

et d'un filtre linéaire discret dont les coefficients sont :

$$w(n) = [w(0), w(1), \dots, w(N-1)]^T \quad (3.8)$$

Comme est présenté dans la figure 1, ce filtre fournit à sa sortie un signal de sortie $y(n)$ qui est le résultat de la convolution entre l'observation et le vecteur du filtre, il est donnée par :

$$y(n) = \sum_{k=0}^{N-1} u(n-k)w(k) \quad (3.9)$$

$$y(n) = u(n)w(0) + u(n-1)w(1) + \dots + u(n-N+1)w(N-1) \quad (3.10)$$

Notre objectif est d'estimer les coefficients du filtre w qui rendent $y(n)$ la meilleure estimation de $d(n)$ et minimise l'erreur $e(n)$ calculer par :

$$d(n) \text{ et minimise l'erreur} \quad (3.11)$$

$$e(n) = d(n) - y(n)$$

La solution est recherchée avec les contraintes suivantes :

1. Le filtre w est linéaire, à temps discret ;
2. Le critère est de minimiser l'erreur quadratique moyenne (MMSE : *Minimum Mean Square Error*), c'est-à-dire le minimum de $E[|e(n)|^2]$.

Chapitre III : Les techniques de rehaussement du signal de parole

Où $E(\cdot)$ désigne l'espérance mathématique. Ce problème peut être associé au schéma de la figure 1. L'erreur quadratique moyenne, J , est la fonction de coût que l'on cherche à minimiser :

$$J = E[e(n)e(n)^*] = E[|e(n)|^2] \quad (3.12)$$

L'opérateur gradient est définie par :

$$\nabla^k = \frac{\partial}{\partial a_k} \quad \text{avec } k = 0, 1, \dots, N-1 \quad (3.13)$$

Cet opérateur est appliqué à la fonction de coût J , d'où :

$$\nabla^k J = \frac{\partial J}{\partial a_k} \quad \text{avec } k = 0, 1, \dots, N-1 \quad (3.14)$$

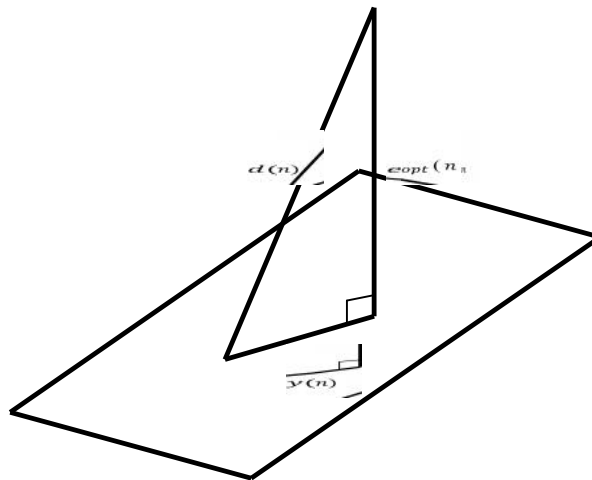


Figure3. 3. Représentation géométrique du théorème de la projection orthogonale.

La solution optimale correspond à une erreur quadratique minimale, notée $e_{opt}(n)$, telle que :

$$\nabla^k J = 0 \quad \text{avec } k = 0, 1, \dots, N-1 \quad (3.15)$$

Chapitre III : Les techniques de rehaussement du signal de parole

Ou encore :

$$E[u(n-k)e_{opt}(n)] = 0 \quad \text{avec } k = 0, 1, \dots, N-1 \quad (3.16)$$

On remarque que l'erreur optimale est orthogonale à toutes les entrées, $u(n), u(n-1), \dots, u(n-N+1)$. On retrouve ainsi le théorème de la projection orthogonale dont représentation géométrique est présenté dans la figure 2.[11]

En corolaire à cette remarque, l'approximation optimale notée $y_{opt}(n)$, qui s'écrit :

$$y_{opt}(n) = \sum_{k=0}^{+\infty} W_{opt}^* u(n-k) \quad (3.17)$$

Cette approximation optimale est donc une combinaison linéaire de termes $u(n-k)$ qui sont tous orthogonaux à $e_{opt}(n)$. Par conséquent, $y_{opt}(n)$ est elle-même orthogonale à $e_{opt}(n)$, c'est-à-dire que l'on peut écrire :[11]

$$E[y_{opt}(n)e_{opt}(n)] = 0 \quad (3.18)$$

3.3.2. Notations :

Soit U_n l'espace engendré par tous les $u(k), k \leq n$, on notera $\hat{d}(\frac{n}{U_n})$ l'estimation optimale au sens de l'erreur quadratique moyenne minimale (MMSE), obtenue à partir des données jusqu'à l'instant n inclus. D'après les résultats précédents, on peut donc écrire : [11]

$$y_{opt}(n) = \hat{d}(\frac{n}{U_n}) \quad (3.19)$$

3.3.3. Calcul de l'erreur quadratique minimale :

On rappelle que $e(n) = d(n) - y(n)$ et qu'à l'optimum, on a :

$$e_{opt}(n) = d(n) - \hat{d}(\frac{n}{U_n}) \quad (3.20)$$

La valeur minimale de l'erreur quadratique moyenne est alors :

Chapitre III : Les techniques de rehaussement du signal de parole

$$J_{min} = E[|e_{opt}(n)|^2] \quad (3.21)$$

D'après (3.21), on peut écrire, $d(n) = \hat{d}\left(\frac{n}{U_n}\right) + e_{opt}(n)$, de plus on sait que :

$\hat{d}\left(\frac{n}{U_n}\right) = y_{opt}(n) - e_{opt}(n)$. On en déduit donc par le théorème de Pythagore l'égalité :

$$E[|d(n)|^2] = E[|\hat{d}(n/U_n)|^2] + E[|e_{opt}(n)|^2] \quad (3.22)$$

c'est-à-dire :

$$\delta_d^2 = \hat{\delta}_d^2 + J_{min} \quad (3.23)$$

On en déduit finalement :

$$J_{min} = \frac{\hat{\delta}_d^2}{\delta_d^2} = 1 - \left(\frac{\hat{\delta}_d}{\delta_d}\right)^2 \quad (3.24)$$

En pratique, on normalise l'expression en divisant par δ_d^2 d'où l'erreur quadratique moyenne normalisée minimale :

$$\varepsilon = \frac{J_{min}}{\delta_d^2} = 1 - \left(\frac{\hat{\delta}_d}{\delta_d}\right)^2 \quad (3.25)$$

On remarque que cette erreur normalisée vérifie $0 \leq \varepsilon \leq 1$. Si $\varepsilon = 0$, on a alors $\delta_d^2 = \hat{\delta}_d^2$ ce qui correspond à une adéquation parfaite entre d et \hat{d} [11].

3.3.3 Equation de Wiener-Hopf :

3.3.3.1 Expression générale :

A partir du principe d'orthogonalité (3.17), de la définition de $e_{opt}(n)$ et $d(n)$, l'équation (3.17) devient :

Chapitre III : Les techniques de rehaussement du signal de parole

$$E[u(n-k)d(n) - \sum_{j=0}^{N-1} w_{opt,j} u(n-j)] = 0 \quad \text{avec} \quad k = 0, 1, \dots, N-1 \quad (3.26)$$

$$E[u(n-k)d(n)] = \sum_{j=0}^{N-1} w_{opt,j} E[u(n-k)u(n-j)] = 0 \quad \text{avec} \quad k = 0, 1, \dots, N-1 \quad (3.27)$$

Or, pour un signal $u(n)$ stationnaire, le terme $E[u(n-k)u(n-j)]$ n'est autre que le terme $R_{uu}(j, k)$ de la matrice d'autocorrélation. Le terme $E[u(n-k)d^*(n)]$ est l'inter-corrélation entre l'entrée du filtre et la sortie désirée, que l'on note $R_{ud}(k)$. L'équation (3.20) peut donc s'écrire :

$$R_{uu}(j, k) = \sum_{j=0}^{N-1} w_{opt,j} R_{ud}(k) \quad \text{avec} \quad k = 0, 1, \dots, N-1 \quad (3.28)$$

Cet ensemble infini d'équations porte le nom d'équations de Wiener-Hopf [11].

3.3.3.2 Solution pour un filtre FIR :

Si on suppose que le filtre est à réponse impulsionnelle finie (RIF) de longueur, N , on peut écrire les équations (3.32) sous forme d'un système de N équations :

$$R_{uu}(j, k) = \sum_{j=0}^{N-1} w_{opt,j} R_{uu}^{-1}(j, k) = 0 \quad \text{avec} \quad k = 0, 1, \dots, N-1 \quad (3.29)$$

La matrice d'auto-corrélation s'écrit :

$$R_{uu} = E[u(n)u^*(n)] = \begin{bmatrix} R_{uu}(0) & R_{uu}(1) & \dots & R_{uu}(N-1) \\ R_{uu}(1) & & & R_{uu}(N-2) \\ \vdots & & \ddots & \vdots \\ R_{uu}(N-1) & R_{uu}(N-2) & \dots & R_{uu}(0) \end{bmatrix} \quad (3.30)$$

Les équations de Wiener-Hopf s'écrivent sous la forme matricielle :

Chapitre III : Les techniques de rehaussement du signal de parole

$$R_{uu, w_{opt}} = R_{ud} \quad (3.31)$$

Avec $w_{opt} = [w_{opt,0}, w_{opt,1}, \dots, w_{opt,N-1}]^T$. Si la matrice d'autocorrélation R_{uu} est régulière, R_{uu}^{-1} existe et la solution optimale est donnée par :

$$w_{opt} = R_{uu}^{-1} R_{ud} \quad (3.32)$$

L'estimation de w_{opt} nécessite donc la connaissance de la matrice d'autocorrélation (c'est-à-dire des statistiques à l'ordre 2) des entrées du filtre et de l'inter-corrélation entre les entrées du filtre et la sortie désirée. Dans le cas d'un filtre d'ordre N (N coefficients), il faut donc estimer d'abord les $\frac{N(N+1)}{2}$ termes R_{uu} et les N coefficients de R_{ud} . [11]

3.4. Algorithme du gradient stochastique (least-men-square –LMS) :

L'algorithme du gradient stochastique est une approximation de l'algorithme du gradient déterministe. L'algorithme LMS est certainement l'algorithme adaptatif le plus populaire qui existe en raison de sa simplicité [12]

Plan :

- .Rappels sur l'algorithme du gradient déterministe
- .L'algorithme LMS
- . Convergence de l'algorithme LMS
- .Quelques règles
- .L'algorithme LMS pour des données complexes
- .Exemple : égalisation adaptative
- .Résumé [12]

3.4.1. Rappels sur l'algorithme du gradient déterministe :

L'algorithme du gradient déterministe est :

$$h(n+1) = h(n) - \frac{1}{2} \mu g(n) \quad (3.33)$$

Ou

$$\begin{aligned}
 \left(\frac{\partial J[h(n)]}{\partial h(n)} \right) &= -2E\{X(n)e(n)\} \\
 &= -2p + 2Rh(n)
 \end{aligned} \tag{3.34}$$

Est le gradient de la fonction coût $J[h(n)] = E\{e^2(n)\}$. Cet algorithme peut encore s'écrire en utilisant le signal d'erreur :

$$\begin{aligned}
 \text{coût } J[h(n)] &= E\{e^2(n)\}. \text{ Cet} \\
 e(n) &= d(n) - x^T(n)h(n)
 \end{aligned} \tag{3.35}$$

$$h(n+1) = h(n) + \mu E\{x(n)e(n)\}. \tag{3.36}$$

Problème : en pratique, $E\{x(n)e(n)\}$ ou de manière équivalente R et p ne sont pas connus. [12]

3.4.2. L'algorithme LMS :

Puisque $R = E\{x(n)x^T(n)$, et $p = E\{x(n)d(n)\}$ sont inconnus, on approchera ces grandeurs déterministes par des estimées $\hat{R}(n)$ et $\hat{P}(n)$ à l'instant n . dans le cas du LMS on choisit les estimées les plus simples possibles, à savoir :

$$\begin{aligned}
 \hat{R}(n) \text{ et } \hat{P}(n) \text{ à l'instant } n, \text{ à savoir :} \\
 \hat{R}(n) = x(n)x^T(n),
 \end{aligned} \tag{3.37}$$

$$\hat{P}(n) = x(n)d(n). \tag{3.38}$$

Ce sont simplement les estimées instantanées des corrélations.

En remplaçant $\hat{R}(n)$ et $\hat{P}(n)$ dans l'algorithme gradient déterministe [éq.(3.37)], on obtient :

$$\begin{aligned}
 & h(n) + \mu [\hat{P}(n) - \hat{P}(n-1)] \\
 & h(n) + \mu x(n)[d(n) - x^T(n)h(n)] \\
 & h(n) + \mu x(n)e(n),
 \end{aligned} \tag{3.39}$$

Qui est l'algorithme LMS. On remarquera que $h(n)$ est maintenant une variable aléatoire [puisque à chaque nouvelle itération n , $h(n)$ dépend des processus aléatoires $x(n)$ et $d(n)$] [12].

3.4.3. Résumé de l'algorithme LMS :

Calcul de la sortie du filtre :

$$y(n) = \sum_{k=0}^{L-1} h_k(n)x(n-k) \tag{3.40}$$

Calcul du signal d'erreur :

$$e(n) = d(n) - y(n) \tag{3.41}$$

Mise à jour du filtre :

$$h(n+1) = h(n) + \mu x(n)e(n) \tag{3.42}$$

μ est le pas d'adaptation de l'algorithme qui démarre avec une initialisation quelconque $h(0)$.

L'algorithme LMS est très simple : il nécessite seulement $2L+1$ multiplications et $2L$ additions par itération, ou L est le nombre de coefficients du filtre [12].

3.4.4. Convergence de l'algorithme LMS :

L'analyse de la convergence du LMS se fait en utilisant les deux critères suivants :

.convergence en moyenne du filtre $h(n)$, c'ad :

$$\lim_{n \rightarrow \infty} E\{h(n)\} = h_{opt} \tag{3.43}$$

Chapitre III : Les techniques de rehaussement du signal de parole

.Convergence du critère $J(n)$ (en moyenne quadratique),càd :[12]

$$\lim_{n \rightarrow \infty} E\{J(n)\} = J(\infty) \text{ constante.} \quad (3.44)$$

3.4.4.1. Convergence en moyenne :

L'équation du LMS est :

$$\begin{aligned} h(n+1) &= h(n) + \mu e(n)x(n), \\ &= h(n) + \mu x(n)[d(n) - x^T(n)h(n)] \\ &= [I - \mu x(n)x^T(n)]h(n) + \mu d(n)x(n), \end{aligned} \quad (3.45)$$

En prenant l'espérance mathématique et en supposant l'indépendance entre les données $x(n)$ et les coefficients du filtre $h_l(n)$, on a :

$$E\{h(n+1)\} = [I - \mu R]E\{h(n)\} + \mu p \quad (3.46)$$

Posons le vecteur maligned :

$$c(n) = h(n) - h_{opt} \quad (3.47)$$

$$E\{c(n+1)\} = [I - \mu R]E\{c(n)\}. \quad (3.48)$$

L'équation (14) devient après avoir additionné les deux cotés avec $-h_{opt}$:

$$E\{c(n+1)\} = [I - \mu R]E\{c(n)\} \quad (3.49)$$

Puisque $R = Q Q^T$, en prenant $v(n) = Q^T c(n)$, l'équation précédente est maintenant :

$$E\{v(n+1)\} = [I - \mu \Lambda]E\{v(n)\}, \quad (3.50)$$

Ou encore, comme on l'a fait pour le gradient déterministe :

$$E\{v^l(n)\} = \mu \lambda^l E\{v^l(0)\} \quad l = 0, 1, \dots, L-1. \quad (3.51)$$

On voit bien que la condition de stabilité est :

Chapitre III : Les techniques de rehaussement du signal de parole

$$0 < \frac{c}{\lambda_{max}} < \frac{2}{\lambda_{max}} \quad (3.64)$$

Où λ_{max} est la plus grande valeur propre de la matrice R dans ce cas :

$$\lim_{n \rightarrow \infty} \lambda_{max} = 0 \quad (3.65)$$

Et par conséquent : [12]

$$\lim_{n \rightarrow \infty} \lambda_{max} = h^2 \quad (3.52)$$

3.4.4.2. Convergence en moyenne quadratique :

Le signal d'erreur produit par l'algorithme LMS est :

$$\begin{aligned} e(n) &= d(n) - x^T(n)h(n) \\ &= d(n) - x^T(n)[h(n) - h_{opt}] \\ &= e_{min}(n) - x^T(n)c(n), \end{aligned} \quad (3.53)$$

Où $e_{min}(n)$ est le signal d'erreur obtenu avec le filtre optimal de Wiener.

Reprenons l'équation du LMS :

$$\begin{aligned} h(n+1) &= h(n) + \mu x(n)e(n) \\ &= h(n) + \mu x(n)e_{min}(n) - \mu x(n)x^T(n)c(n), \end{aligned} \quad (3.54)$$

Qui s'écrit aussi en fonction du vecteur $c(n) = h(n) - h_{opt}$ comme suit :

$$h(n+1) = [h(n) - h_{opt}] + \mu x(n)x^T(n)c(n) + \mu x(n)e_{min}(n) \quad (3.55)$$

La matrice d'autocorrélation du vecteur misalignement $c(n)$ est :

$$K(n) = E\{c(n)c^T(n)\} \quad (3.56)$$

En utilisant l'hypothèse d'indépendance et à partir de (24) on déduit une équation pour $K(n)$:

Chapitre III : Les techniques de rehaussement du signal de parole

$$K(n+1) = \frac{1}{1 - \mu R} K(n) (1 - \mu R) + \mu^2 j_{min} R, \quad (3.57)$$

Ou :

$$j_{min} = E\{e_{min}^2(n)\} \quad (3.58)$$

Est l'erreur quadratique moyenne minimal.

En utilisant la définition précédente et en invoquant l'hypothèse d'indépendance, l'erreur quadratique moyenne $J(n)$ du LMS peut être évaluée de la manière suivante :

$$\begin{aligned} J(n) &= E\{[e_{min}(n) - x^T(n)c(n)][e_{min}(n) - x^T(n)c(n)]\} \\ &= E\{e_{min}^2(n) - 2x^T(n)c(n)e_{min}(n) + x^T(n)c(n)x^T(n)c(n)\} \\ &= j_{min} + E\{c^T(n)x(n)c(n)\} \end{aligned} \quad (3.59)$$

D'autre part :

$$\begin{aligned} E\{c^T(n)x(n)c(n)\} &= E\{tr[c^T(n)x(n)c(n)]\} \\ &= E\{tr[x(n)x^T(n)c(n)c^T(n)]\} \\ &= tr\{E[x(n)x^T(n)]E[c(n)c^T(n)]\}. \end{aligned} \quad (3.60)$$

En invoquant l'hypothèse d'indépendance on a finalement :

$$\begin{aligned} E\{c^T(n)x(n)c(n)\} &= tr\{E[x(n)x^T(n)]E[c(n)c^T(n)]\} \\ &= tr\{RK(n)\}. \end{aligned} \quad (3.61)$$

On peut réécrire l'erreur quadratique moyenne :

$$J(n) = j_{min} + tr\{RK(n)\}. \quad (3.62)$$

On définit l'erreur quadratique moyenne excédentaire (excess mean-square-error) comme :

Chapitre III : Les techniques de rehaussement du signal de parole

$$\begin{aligned}
 J_{ex}(n) &= J(n) - J_{min} \\
 &= tr[RK(n)] > 0.
 \end{aligned}
 \tag{3.63}$$

En utilisant la factorisation $R = Q \Lambda Q^T$ dans l'expression précédente, on obtient :

$$\begin{aligned}
 J_{ex}(n) &= tr[Q \Lambda Q^T K(n)] \\
 &= tr[\Lambda Q^T K(n) Q] \\
 &= tr[\Lambda Z(n)],
 \end{aligned}
 \tag{3.64}$$

Où $Z(n) = Q^T K(n) Q$, puisque est une matrice diagonale, on peut encore écrire :

$$J_{ex}(n) = \sum_{l=0}^{L-1} \lambda_l z_l(n)
 \tag{3.65}$$

Où $z_l(n)$ sont les éléments diagonaux de la matrice $z(n)$.

On se souvient de l'équation récursive :

$$z_l(n+1) = (1 - \mu \lambda_l) z_l(n) + \mu^2 J_{min} \lambda_l
 \tag{3.66}$$

Soit en multipliant à gauche et à droite par Q^T et Q :

$$Z(n+1) = (1 - \mu \Lambda) Z(n) + \mu^2 J_{min} \Lambda
 \tag{3.67}$$

Et les éléments diagonaux de $Z(n+1)$ se calculent récursivement :

$$z_l(n+1) = (1 - \mu \lambda_l) z_l(n) + \mu^2 J_{min} \lambda_l
 \tag{3.68}$$

La récursion converge si $|1 - \mu \lambda_l| < 1$, l . L'algorithme LMS est donc convergent en moyenne quadratique si :

$$0 < \mu < \frac{2}{\lambda_{max}}
 \tag{3.69}$$

A l'infini et avec la condition de convergence précédente, on a :

Chapitre III : Les techniques de rehaussement du signal de parole

$$\lim_{z \rightarrow \infty} j_{ex}(z) = (1 - \frac{\sum_{l=0}^{L-1} \lambda_l z^{-l}}{\mu \lambda^l z^{-l}}) + \frac{\sum_{l=0}^{L-1} \lambda_l z^{-l}}{\mu^{-j_{min}} \lambda^l} \quad (3.70)$$

Soit

$$z^{-l} = \frac{\mu^{-j_{min}} \lambda^l}{\mu \lambda^l}, l = 0, 1, \dots, L-1. \quad (3.71)$$

On en déduit:

$$j_{ex}(\infty) = \sum_{l=0}^{L-1} \lambda_l z^{-l} = j_{min} \sum_{l=0}^{L-1} \frac{\mu \lambda_l}{\mu \lambda^l} \quad (3.72)$$

D'où :

$$\lim_{n \rightarrow \infty} j(n) = j(\infty) = j_{ex}(\infty) + j_{min} = j_{min} (1 + \sum_{l=0}^{L-1} \frac{\mu \lambda_l}{\mu \lambda^l}) = \text{constante} \quad (3.73)$$

Avec la condition de convergence. [12]

Quelques règles :

Comment choisir le pas d'adaptation : $0 < \mu < \frac{2}{\lambda_{max}}$?

En pratique λ_{max} n'est pas facile à déterminer.

On prend une estimée conservatrice :

$$\text{tr}[R] = \sum_{l=0}^{L-1} \lambda_l - \text{tr}[R] > \lambda_{max} \quad (3.74)$$

D'autre part :

$$\text{tr}[R] Lr(0) = L e^{\frac{\lambda_{max}}{2}} \quad (3.75)$$

D'où :

Chapitre III : Les techniques de rehaussement du signal de parole

$$0 < \mu < \frac{2}{L\sigma^2} \quad (3.76)$$

Qui est un choix sûr pour garantir la convergence de l'algorithme LMS.

Soit la quantité suivante :

$$\lambda_{moy} = \frac{1}{L} \sum_{l=0}^{L-1} \lambda_l \quad (3.77)$$

Qui représente une moyenne des valeurs propres de la matrice R. en utilisant certains résultats obtenus sur la constante de temps de l'algorithme du gradient déterministe, on définit la constante de temps moyenne de la courbe d'apprentissage de l'algorithme LMS :

$$\tau_{eqmoy} \approx \frac{1}{2\mu\lambda_{moy}} \quad (3.78)$$

On définit le « misadjustment » comme :

$$m = \frac{J_{min}}{J_{min}} = \sum_{l=0}^{L-1} \frac{\mu\lambda_l}{\mu\lambda_l} \quad (3.79)$$

Si $\mu\lambda_{max} \ll 1$ on a :

$$m \approx \frac{\mu}{2} \sum_{l=0}^{L-1} \lambda_l = \frac{\mu}{2} L\sigma^2 \quad (3.80)$$

En utilisant le pas normalisé :

$$\mu = \frac{\alpha}{L\sigma^2} \quad (3.81)$$

Où $0 < \frac{\alpha}{2} < 1$ on obtient :

Chapitre III : Les techniques de rehaussement du signal de parole

$$m \approx \frac{L}{4\tau_{eqmoy}} \quad (3.82)$$

D'où les observations suivantes :

.la valeur de τ augmente linéairement avec la longueur L du filtre h .

. m est inversement proportionnel au pas d'adaptation, d'où conflit entre vitesse de convergence et valeur finale de l'EQM [12].

3.4.4.3. L'algorithme LMS pour des données complexes :

Calcul de la sortie du filtre :

$$y(n) = \sum_{k=0}^{L-1} h_k x(n-k) \quad (3.83)$$

Mise à jour du filtre :

$$h_k(n+1) = h_k(n) + \mu x(n-k) e^*(n) \quad (3.84)$$

$\mu, 0 < \mu < \frac{2}{Lr_x^2}$, est le pas d'adaptation de l'algorithme qui démarre avec une initialisation quelconque $h(0)$ [12].

Exemple : égalisation adaptative

Dans cet exemple on étudie l'utilisation de l'algorithme LMS pour l'égalisation adaptative d'un canal linéaire adaptatif.

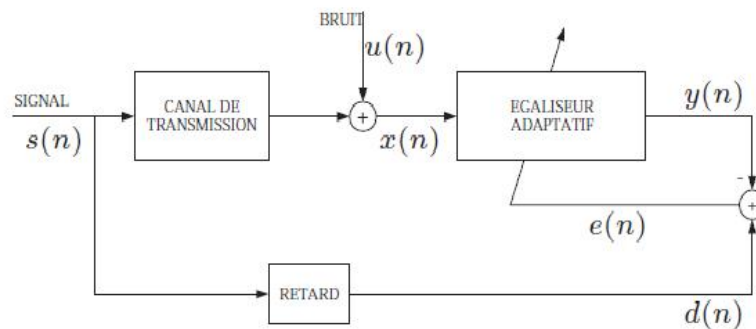


Figure 3.4: Egalisation adaptative.

Chapitre III : Les techniques de rehaussement du signal de parole

Le signal de transmission $s(n)$ est une séquence aléatoire ou $s(n) \in \{-1, 1\}$. le signal déformé par un canal dont la réponse impulsionnelle est :

$$w^k = \begin{cases} 1 \\ \frac{1}{2} \left[1 + \cos\left(\frac{2\pi}{\beta}(k-1)\right) \right] \end{cases}, k = 0, 1, 2 \quad (3.85)$$

Où β contrôle la quantité de distorsion produite par le canal (plus β est grand et plus la distorsion est importante). dans notre exemple, $\beta = 0.25$.

Un bruit blanc de moyenne nulle et de variance

$$\frac{\sigma_u^2}{2} = 0.001 \text{ est ajouté à la sortie du canal.}$$

Le signal d'entrée de ce filtre est :

$$x(n) = \sum_{k=0}^L w_k s(n-k) + u(n). \quad (3.86)$$

D'après la Fig., le signal d'erreur est :

$$e(n) = d(n) - y(n) = s(n - \delta) - h_{rc}(n) * x(n). \quad (3.87)$$

La moyenne d'ensemble est approximée en moyennant 200 réalisations indépendantes.

La Fig.2 montre les courbes d'apprentissage de l'algorithme LMS pour deux pas d'adaptation différents. on voit bien que pour un pas grand, l'algorithme converge plus rapidement qu'avec pas plus petite mais l'erreur résiduelle est plus importante qu'avec un pas plus petite mais l'erreur résiduelle est plus importante [12].

3.5. Conclusion :

L'algorithme LMS est simple et ces performances dépendent de trois facteurs :

1. le pas d'adaptation μ ,
2. les valeurs propres λ_l , de la matrice R
3. la longueur du filtre h.

Avec un pas d'adaptation petite, le LMS converge lentement mais l'EQM excédentaire est petite. Avec un pas d'adaptation grand, le LMS converge rapidement mais l'EQM excédentaire est grande.

Chapitre 4 :

Tests et

Expérimentations.

Chapitre 4 :

Tests et Expérimentations.

4.1 Introduction :

Dans ce chapitre on a fait un étude de débruitage de la parole basé sur la technique de l'annulation de bruit et par la technique de filtre WIENER .Aussi on a considéré les performances du filtre adaptatif sont testées en employant différents algorithmes du gradient stochastique LMS et NLMS.

Les résultats obtenus sont comparés pour voir lequel des algorithmes qui rendent le filtre plus performant.

Les différents algorithmes sont implémentés sous MATLAB.

4.2. Débruitage de la parole par la technique de l'annulation de bruit :

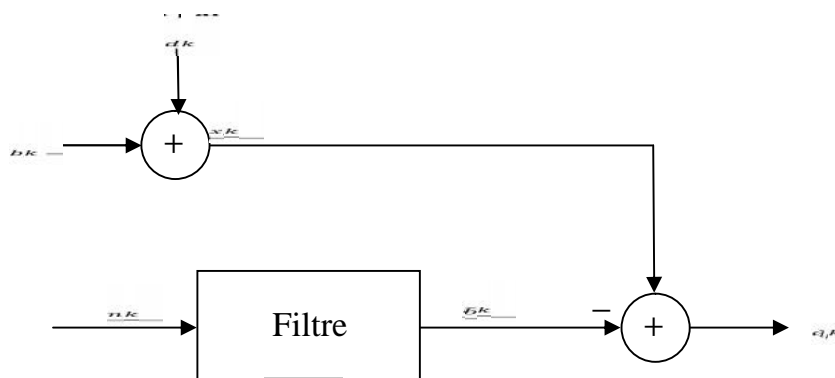


Figure 4.1 : Schéma synoptique d'un annulateur de bruit.

Le signal d'entrée est x_k qui est le signal de parole original d_k bruité par un bruit additif b_k .

$$x_k = d_k + b_k \quad (4.1)$$

Dans ce cas le filtre a pour but de donner une estimation \hat{b}_k du bruit b_k a partir d'un signal d'entrée n_k . Le signal estimer est utilisé pour réduire le bruit b_k du signal bruité x_k . La minimisation est faite par soustraction donnant comme résultat un signal \hat{d}_k qui est une estimation du signal original d_k .

Chapitre IV : Test et expérimentations

$$\hat{q}_k = x_k - \hat{b}_k$$

(4.2)

$$\begin{aligned} &= d_k + b_k - \hat{b}_k \\ \hat{d}_k &= d_k + (b_k - \hat{b}_k) \end{aligned}$$

(4.3)

Dans le cas optimale, $\hat{b}_k = b_k$ et par conséquence $\hat{d}_k = d_k$ et l'erreur d'estimation qui est mesuré comme erreur quadratique moyenne :

$$E\{|e_k|^2\} = E\{|(b_k - \hat{b}_k)|^2\} = E\{|(d_k - \hat{d}_k)|^2\} = 0 \quad (4.4)$$

4.3. Débruitage de la parole par filtrage de WIENER :

Dans cet exemple, nous appliquons le filtrage de WIENER dans le contexte de l'annulation de bruit présenté dans la figure 4.1.

Dans cet exemple, le signal original est considéré dans un premier temps comme une onde sinusoïdale. Le bruit additif b_k et d'entrée n_k sont considérés comme bruits blanc.

Le bruit est ajouté au signal original avec un rapport signal sur bruit (RSB) égale à 0 décibel (dB).

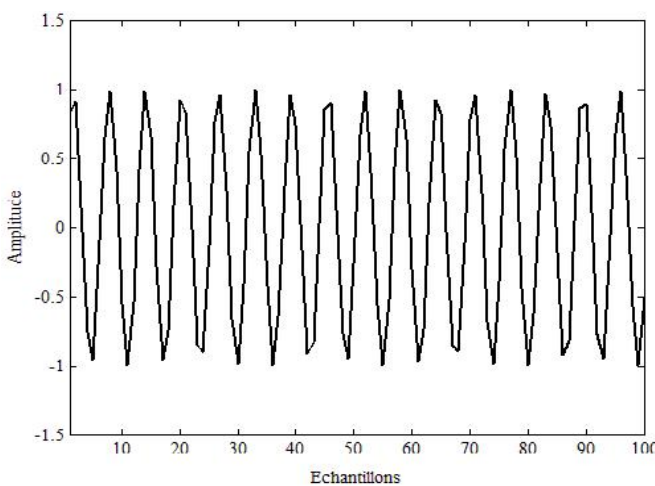


Figure 4.2 : Signal sinusoïdale original.

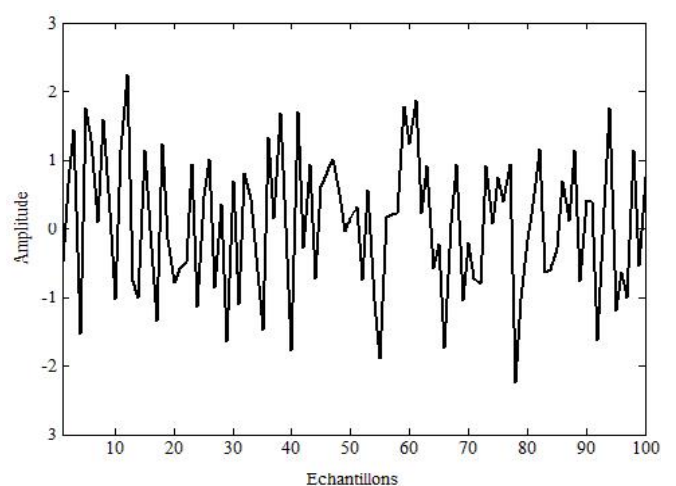


Figure 4.3 : Bruit.

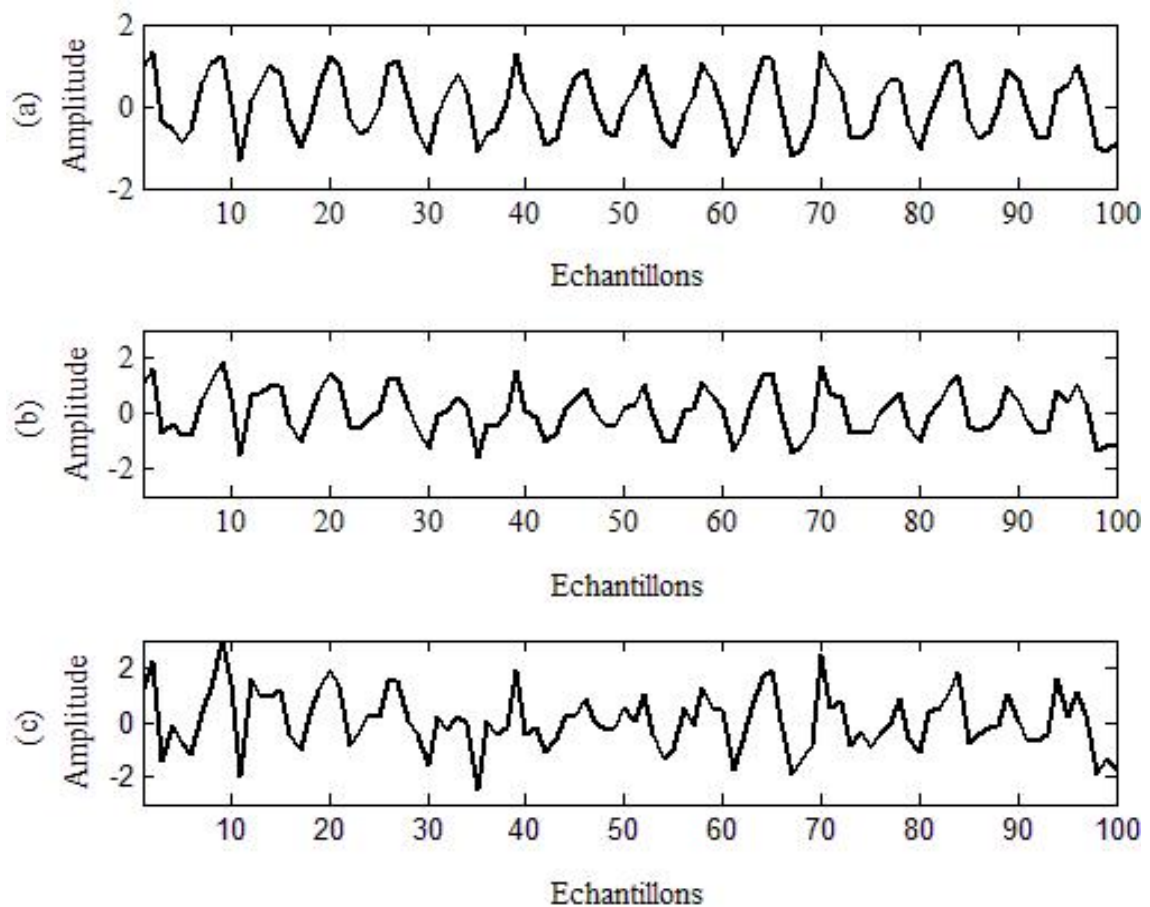


Figure 4.4 : Signal original bruité par différents niveau du RSB.

On a appliqué le filtrage de WIENER sur un signal sinusoïdal dans le but de réduire le bruit additif avec des différents niveaux du SNR.

4.4. Annulation de bruit par filtrage adaptatif :

Dans cette sous section nous allons utiliser le filtrage adaptatif dans l'annulation de bruit. Pour cela nous allons introduire l'algorithme du filtrage adaptatif LMS et NLMS dans la structure de la figure 4.1. La structure résultante est présentée dans la figure 4.2.

Chapitre IV : Test et expérimentations

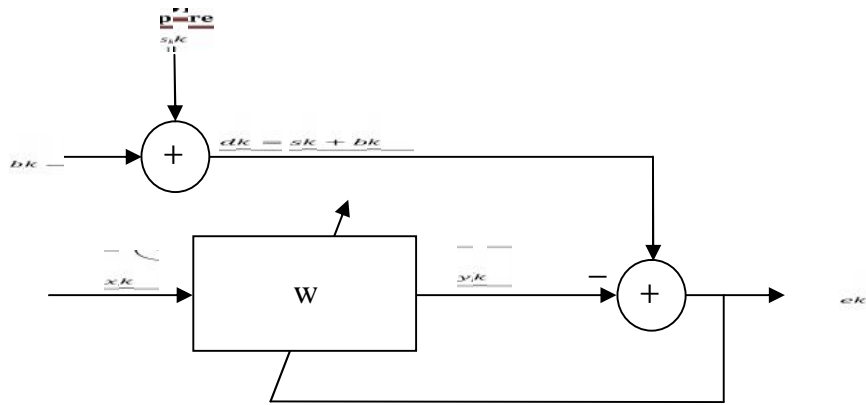


Figure 4.5 : Schéma synoptique d'un annulateur adaptatif de

Comme est discuté dans le chapitre 3, l'algorithme de moindres moyennes carrées LMS (least means squares) est un algorithme adaptatif utilisé dans le filtrage adaptatif.

$$y_k = W_k^T x_k \quad (4.5)$$

$$e_k = d_k - y_k \quad (4.6)$$

$$W_{k+1} = W_k + \mu e_k x_k \quad (4.7)$$

$$W_{k+1} = W_k + \frac{\mu e_k x_k}{(x_k^T x_k + \epsilon)} \quad (4.8)$$

Où y_k est l'échantillon du signal filtré.

e_k est l'échantillon du signal de l'erreur qui est la différence entre l'échantillon du signal désiré d_k et du bruit filtré y_k . w_k est le vecteur des coefficients du filtre à l'instant de l'échantillon k . L'équation (4.7) présente l'adaptation du vecteur w_k pour le cas de l'algorithme LMS et l'équation (4.8) présente l'adaptation du vecteur w_k pour le cas de l'algorithme NLMS où le terme $(x_k^T x_k + \epsilon)$ permet de réaliser une normalisation de l'énergie du signal d'entrée x_k et de le rendre l'algorithme indépendant de l'énergie du signal d'entrée. Cette normalisation permet de rendre le pas d'adaptation qui est constant dans le cas de l'algorithme LMS à temps variable dans le cas de l'algorithme NLMS ou l'équation (4.8) s'écrit :

$$W_{k+1} = W_k + \frac{\mu e_k x_k}{(x_k^T x_k + \epsilon)} \quad (4.9)$$

Où

$$\mu_k = \frac{\mu}{(x_k^T x_k + \epsilon)} \quad (4.10)$$

Chapitre IV : Test et expérimentations

La normalisation permet d'améliorer les performances de l'algorithme NLMS.

4.5. Annulation du bruit du signal parole :

Les paramètres de la simulation sont :

Le pas d'adaptation $\mu = 0.05$ et la longueur du filtre $N = 16$

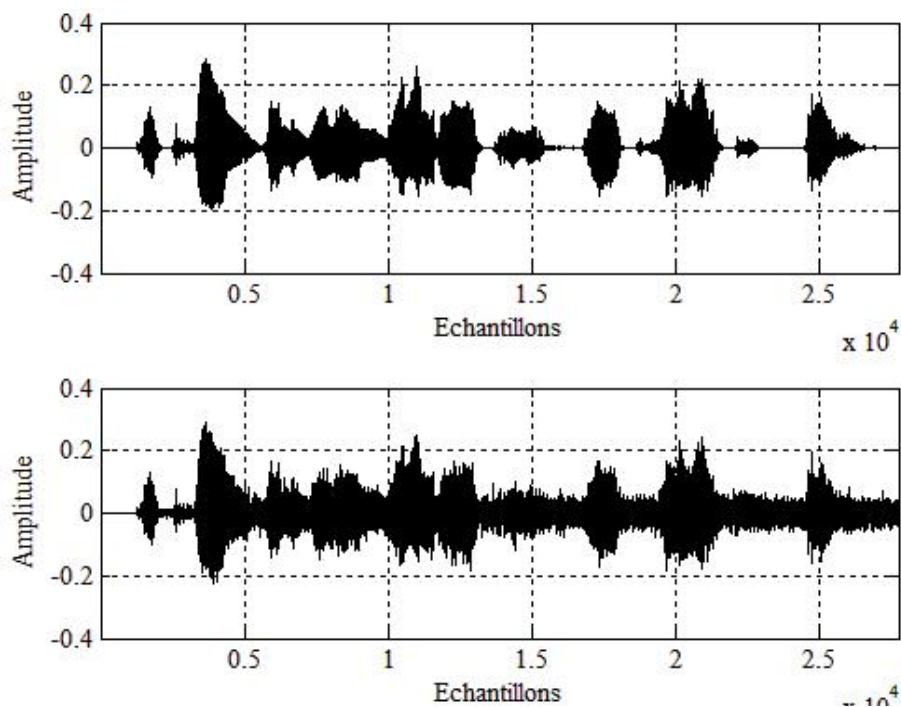


Fig.4.6. En haut : le signal de parole original, En bas : le signal de parole bruité par un bruit blanc et un SNR : 5dB.

La figure 4.6 montre le signal de parole original ainsi que le signal de parole bruité avec un bruit blanc avec un SNR de 0 dB.

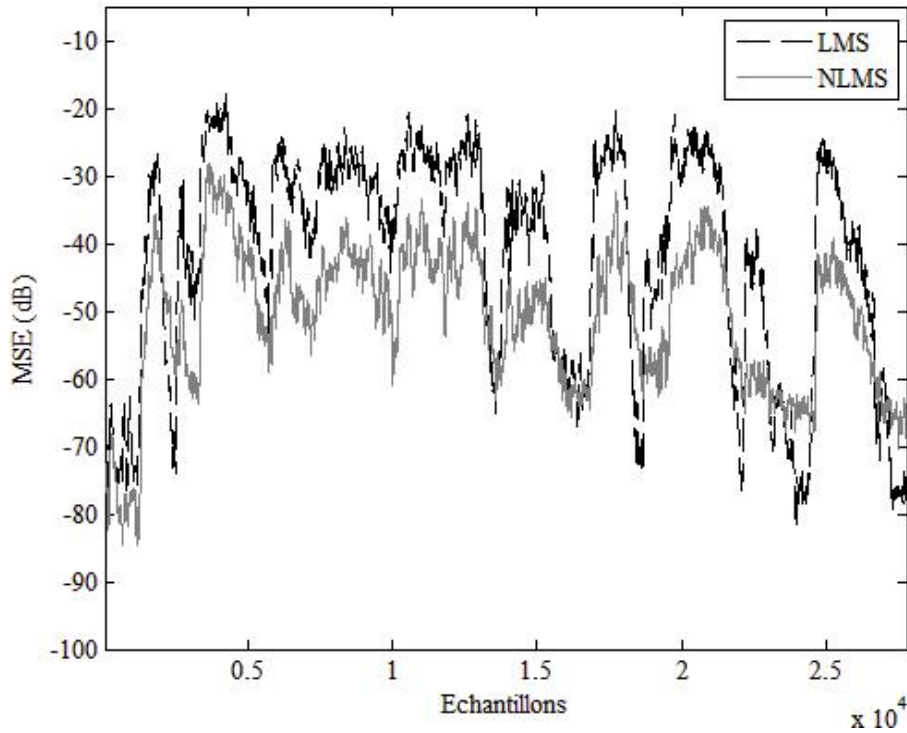


Fig. 4.7. La convergence du MSE de l’algorithme NLMS et LMS dans le contexte de l’annulation du bruit pour un signal de parole.

La figure 4.7 montre la convergence du MSE des algorithmes LMS et NLMS dans le contexte de l’annulation adaptatif du bruit pour un signal de parole. Nous remarquons de cette figure que l’algorithme NLMS dépasse l’algorithme LMS.

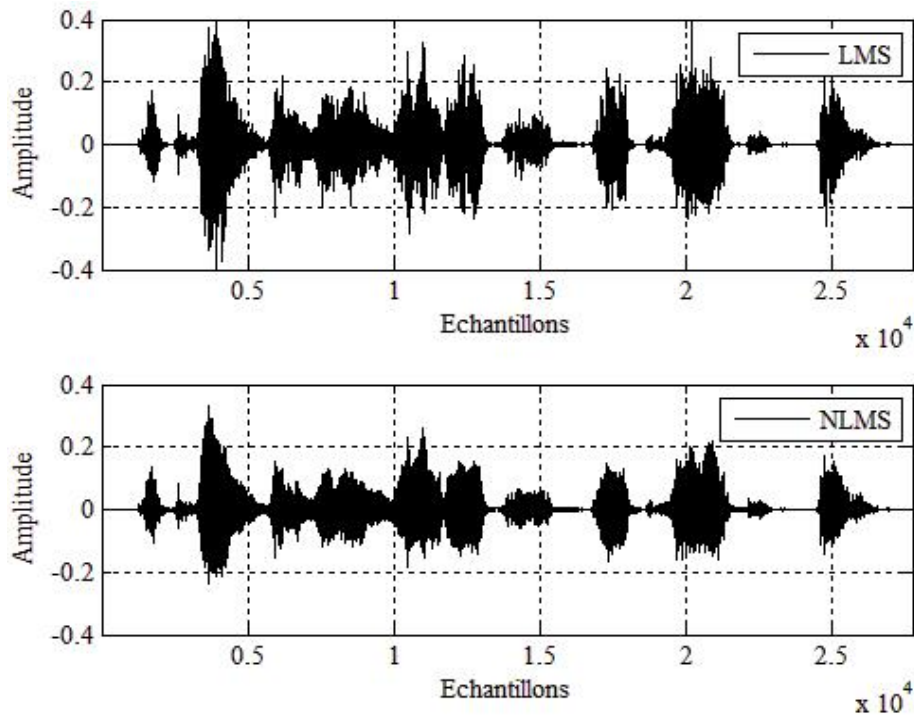


Fig. 4.8. En haut : le signal de parole récupéré (filtré) a la sortie de l'annulateur adaptatif basé sur l'algorithme LMS, En bas : le signal de parole récupéré (filtré) a la sortie de l'annulateur adaptatif basé sur l'algorithme NLMS

La figure 4.8 montre que le signal récupéré a la sortie du filtre NLMS est plus proche dans sa forme au signal originale que le signal récupéré a la sortie du filtre LMS.

Chapitre IV : Test et expérimentations

4.6. Mesure de l'amélioration du SNR après filtrage :

Tableau 4.1 : Comparaison mesure de l'amélioration du SNR après filtrage de LMS et NLMS.

	SNR de l'entrée (dB)	SNR de sortie (dB)
LMS	0	1.11
NLMS		15.22

Le tableau 4.1 montre le SNR de sortie mesuré pour les signaux filtrés par les algorithmes LMS et NLMS. Nous remarquons que le SNR du signal à la sortie du filtre NLMS est plus grand que celui du signal à la sortie du filtre LMS. Alors on peut dire que le signal à la sortie du filtre NLMS contient moins de bruit à cette qualité et plus haute que celle du signal à la sortie du filtre LMS.

4.7. Conclusion :

Dans ce chapitre on a appliqué le filtrage adaptatif basé que les deux algorithmes LMS et NLMS dans le contexte de l'annulation de bruit pour un signal de avec différents niveaux du SNR. L'algorithme NLMS est plus performant que l'algorithme LMS et le signal récupéré à sa sortie présente une qualité plus élevée que celle du signal à la sortie du filtre LMS.

Conclusion Générale

Conclusion Générale

Conclusion générale :

Dans ce mémoire, nous avons étudié quelques techniques utilisées pour le débruitage du signal parole. On a présenté quelques transformées discrètes orthogonales telles que la TFD, la DHT et la DCT.

On a fait une étude comparative entre les algorithmes LMS et NLMS, utilisés dans le filtrage adaptatif, dans le contexte du débruitage du signal parole.

Ces algorithmes nécessitent une bonne configuration pour qu'ils assurent des bonnes performances. La comparaison de ces deux algorithmes dans le contexte de l'annulation adaptatif de bruit a montré que l'algorithme NLMS est plus performant que l'algorithme LMS en terme de convergence du MSE.

Références Bibliographiques

Les références bibliographie :

- [1] A. Amehraye, Débruitage perceptuel de la parole, Ecole Nationale Supérieure des Télécommunications de Bretagne, thèse de doctorat. (2009).
- [2] Site Web : (<https://fr.wikipedia.org/wiki/Cepstre>), consulté le 28 avril 2016, à 15:28.
- [3] M.M. Rezak, Généralités sur la transmission des données, Notes de cours. (2006).
- [4] Site Web :(http://fr.wikipedia.org/wiki/Bruit_blanc) , consulté le 28/05/2013 à 17 :10.
- [5] Site Web :(http://fr.wikipedia.org/wiki/Rapport_signal_sur_bruit) , consulté le 28/05/2013.
- [6] Site Web :(https://www.greyc.fr/sites/default/files/users/mfrikel/Chap2_TFD.pdf)
- [7] Site Web : (http://perso.esiee.fr/~bercherj/New/polys/poly_tfd.pdf)
- [8] Site eb : (http://www.tsi.enst.fr/pages/enseignement/ressources/mti/KL/site/principe_KL.html source de tkf).
- [9] F. PICCININ, « The fast Hartley transform as an alternative to the fast Fourier transform »
Surveillance research laboratory, South Australia, technical memorandum; SRL-0006-TM, February 1988
- [10] Site Web: (https://fr.wikipedia.org/wiki/Analyse_en_composantes_principales)
- [11] linaire optimal université Joseph fourer – polytech 'Grenoble cour de cinquième année du département 3i option image et signaux automatique –(2010)
- [12] algorithme du gradient stochastique « least-mean square-LMS »INRS-EMT J.benesty
- [13] Site Web : (https://fr.wikipedia.org/wiki/analyse-en_composantes_principales)
- [14] Site Web :(www.tsi.enst.fr/pages/enseignement/ressources/mti/kl/site/principe-kl.html source de tkf)
- [15] Site Web :(https://books.google.fr/books?id=livre_communication_audiovisuel_titre_la_transformee_en_cosinus_discrete_ou_TCD).
- [16] mémoire d'amélioration des performances d'un annulateur adaptatif de bruit pour les signaux paroles. Réalisé par : Ounissi Ridha et Maatoub Djaber, Présenté le : 27 – Juin – 2013
- [17] S. V. Vaseghi , Advanced Digital Signal Processing and Noise Reduction, 3 rd ed., Wiley. (2006).