

Ordre...../F.S.S.A/UAMOB/2019

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE
UNIVERSITE AKLI MOAND OULHADJ-BOUIRA



Faculté des Sciences et des Sciences Appliquées
Département : **Génie Electrique**

Mémoire de fin d'étude

Présenté par :

BOUKHERBAB Massinissa
DOUANE Anis

En vue de l'obtention du diplôme de **Master en :**

Filière : Télécommunications

Option : Systèmes des Télécommunications

Thème :

Les détecteurs de double parole dans les communications mains-libres.

Devant le jury composé de :

ALIMOHAD Abdennour	MCA	UAMOB	Président
BENZIANE Mourad	MAA	UAMOB	Encadreur
SAIDI Mohammed	MAA	UAMOB	Examinateur

Année Universitaire 2018/2019

DEDICACE

Nous dédions ce modeste travail

à nos chers parents,

à nos frères et sœurs, et toute nos

familles.

À nos amis, et à tous nos camarades.

REMERCIEMENT

Tout d'abord, nous tenons à remercier avant tout le bon dieu Le tout puissant de nous avoir donné la force et le courage, afin d'accomplir ce modeste travail.

Adressons nos plus vifs remerciements à notre promoteur Mourad Benziane pour son aide, ses conseils judicieux, ainsi que le suivi intensif dans le but de mener à bien ce travail.

Nous remercions également nos familles pour les sacrifices qu'elles ont fait pour que nous terminons nos études.

Nous remercions tous les enseignants à tous les niveaux, particulièrement ceux du département de génie électrique, pour tout le savoir que nous avons acquis grâce à eux durant notre formation.

Nous remercions également les membres de jury qui ont fait honneur d'examiner ce modeste travail.

Nous remercions aussi tous nos amis pour leurs aide, leurs patience et leurs encouragement.

Résumé

Les communications mobiles mains-libre sont très souvent perturbées par un phénomène naturel appelé écho acoustique qui est dû au couplage acoustique entre le haut-parleur et le microphone. Ce problème est classiquement résolu en utilisant les techniques de filtrage adaptatif permettant d'identifier le chemin de l'écho et d'estimer d'écho qui est ensuite soustrait du signal à transmettre, ces filtres sont susceptibles de diverger dès que la parole du locuteur locale a lieu d'où la nécessité d'utiliser des algorithmes de détection de la parole locale pour arrêter l'adaptation des coefficients du filtre et donc éviter sa divergence.

L'objectif de ce travail est de mettre en œuvre deux algorithmes DTD classiques, dans un but de montrer la nécessité des DTD dans un annuleur d'écho acoustique et de comparer les performances de ces deux algorithmes.

Mots clés : Annulation d'écho acoustique, filtrage adaptatif, détection de double parole.

Abstract

Hands-free mobile communications are very often disrupted by a natural phenomenon called acoustic echo which is due to the acoustic coupling between the speaker and the microphone. This problem is usually solved by using adaptive filtering techniques to identify the echo path and estimate echo which is then subtracted from the signal to be transmitted, these filters are likely to diverge as soon as the local speaker's speech takes place hence the need to use algorithms for local speech detection to stop the adaptation of the filter coefficients and thus avoid its divergence.

The objective of this work is to implement two classic DTD algorithms, with the aim of increasing the need for DTDs in an acoustic echo cancel and comparing the performance of these two algorithms.

Key words: Acoustic echo canceller, adaptive filtering, double-talk detection.

ملخص

غالبًا ما تزعج الاتصالات المحمولة دون استخدام اليدين بظاهرة طبيعية تسمى الصدى الصوتي، ويرجع ذلك إلى الاقتران الصوتي بين مكبر الصوت والميكروفون. يتم حل هذه المشكلة تقليديًا باستخدام تقنيات التصفية التكيفية لتحديد مسار الصدى وتقدير الصدى الذي يتم طرحه بعد ذلك من الإشارة المراد إرسالها، ومن المحتمل أن تتباعد هذه المرشحات بمجرد أن يظهر خطاب المتحدث المحلي ومنه تأتي الحاجة إلى استخدام خوارزميات الكشف عن الكلام المحلي لوقف تكيف معاملات المرشح وبالتالي تجنب التباعد.

الهدف من هذا العمل هو تنفيذ خوارزميتين كلاسيكيتين للكشف عن الكلام المزدوج، من أجل التحقق من الحاجة إليها في إلغاء الصدى الصوتي ومقارنة أداء هاتين الخوارزميتين.
الكلمات المفتاحية: إلغاء الصدى الصوتي، التصفية التكيفية، اكتشاف الكلام المزدوج.

SOMMAIRE

DEDICACE	I
REMERCIEMENTS.....	II
RESUME	III
SOMMAIRE	IV
LISTE DES TABLEAUX	VI
LISTE DES FIGURES	VII
LISTE DES ABREVIATIONS.....	VI
LISTE DES NOTATIONS	IX
INTRODUCTION GENERALE	1
Chapitre I : Généralité sur les annuleurs d'écho acoustique	
I.1. Introduction.....	3
I.2. Historique.....	3
I.3. La communication mains libres :	4
I.3.1. Problèmes de communication mains libres	5
I.4. L'écho dans le système des télécommunications.....	5
I.4.1. L'écho électrique.....	5
I.4.2. L'écho acoustique	6
I.4.3. L'écho acoustique et les systèmes téléphoniques [3].....	8
I.4.4. L'environnement acoustique [4]	9
I.4.5. Le chemin parcouru par l'écho [4].....	9
I.5. Annulation d'écho acoustique (AEC)	10
I.5.1. Principe de fonctionnement d'AEC	10
I.5.2. Filtrage adaptatif	11
I.5.2.1. L'algorithme adaptatif	13
I.5.3. Algorithmes de filtrage adaptatif standard	13
I.5.3.1. L'Algorithme LMS.....	13
I.5.3.2. L'algorithme normalisé (NLMS).....	14
I.6. Conclusion.....	16

Chapitre II : Les Détections de Double Parole

II.1. Introduction	17
II.2. La double parole	17
II.3. Détection double parole.....	18
II.3.1. Schémas génériques de la détection de double parole	19
II.4. Les méthodes de la détection de double parole	20
II.4.1. Algorithme de Geigel	20
II.4.2. Algorithme Corrélation croisée (Cross-Correlation)	20
II.4.3. Algorithme de cross-corrélation normalisé (NCC).....	22
II.4.4. DTD basé sur Fast NCC (FNCC)	23
II.4.5. DTD basé sur extraction de fonction vocale (SFE)	24
II.4.6. DTD basé sur la décomposition en valeur singulière (SVD).....	26
II.4.7. DTD basé sur filtre adaptatif auxiliaire (AAF).....	27
II.4.8. La DTD basé sur l'algorithme de réponse d'impulsion de variance (VIRE).....	27
II.5. Conclusion	28

Chapitre III : Simulation et résultats

III.1. Introduction.....	29
III.2. Présentation de logiciel de simulation	29
III.3. Structure générale du programme.....	30
III.3.1. Organigramme du system AEC	30
III.3.2. Code Matlab	31
III.3.2.1. Algorithme de filtres adaptatif	31
III.3.2.2. Algorithmes de Détection de double parole.....	32
III.4. Résultat de simulation.....	34
III.4.1. Algorithme d'annulation d'écho sans DTD.....	36
III.4.2. Algorithme d'annulation d'écho avec DTD	37
III.4.3. Etude du bruit	40
III.4.4. Etude de l'écho restant	46
III.5. Conclusion	47

CONCLUSION GENERALE	48
----------------------------------	----

BIBLIOGRAPHIE	49
----------------------------	----

LISTES DES FIGURES

Figure I.1. Configuration de base d'un système mains libres	4
Figure I.2. L'écho électrique	5
Figure I.3. Exemple de réponse impulsionnelle (4ième chemin d'écho des recommandations G168) [7].	6
Figure I.4. Exemple de communication bidirectionnelle avec écho acoustique	6
Figure I.5. Exemple de réponse impulsionnelle acoustique d'une voiture.	7
Figure I.6. Le schéma de l'écho acoustique	8
Figure I.7. Système d'annulation d'écho acoustique (AEC).....	11
Figure I.8. Principe d'un filtre adaptatif [11].	12
Figure II.1. Schéma de la double parole.....	17
Figure II.2. Structure de base de l'AEC avec DTD	18
Figure III.1. Organigramme de simulation.....	30
Figure III.2. Les signaux utilisés dans la simulation.....	34
Figure III.3. La réponse impulsionnelle	35
Figure III.4. L'influence du pas d'adaptation sur l'algorithme de NLMS en absence de signal local	36
Figure III.5. Influence du pas d'adaptation sur l'algorithme NLMS en présence du signal local.	37
Figure III.6. Evaluation de misalignement pour les méthodes de Geigel et NCC sans bruit... 38	
Figure III.7. L'évaluation de l'ERLE pour les méthodes de Geigel et NCC sans bruit..... 38	
Figure III.8. L'évaluation de MSE pour les méthodes de Geigel et NCC sans bruit. 39	
Figure III.9. Les signaux de DTD obtenus avec les méthodes de Geigel et NCC sans bruit... 40	
Figure III.10. Convergence AEC avec DTD-Geigel dans un environnement bruité. 41	
Figure III.11. Convergence AEC avec DTD-NCC dans un environnement bruité..... 41	
Figure III.12. ERLE DTD-Geigel dans un environnement bruité..... 42	
Figure III.13. ERLE du DTD NCC dans un environnement bruité. 42	
Figure III.14. MSE du DTD Geigel dans un environnement bruité..... 43	
Figure III.15. MSE du DTD NCC dans un environnement bruité. 43	
Figure III.16. Convergence avec et sans DTD (SNR=40 dB)..... 45	
Figure III.17. ERLE avec et sans DTD (SNR=40 dB)..... 45	
Figure III.18. MSE avec et sans DTD (SNR=40 dB)..... 46	
Figure III.19. Comparaison entre les signaux d'erreur des méthodes de Geigel et NCC. 46	

LISTE DES TABLEAUX

Tableau I.1. L'algorithme LMS	14
Tableau I.2. L'algorithme NLMS	15
Tableau III.1. Résumé de l'algorithme NLMS	31
Tableau III.2. Code MATLAB de l'algorithme NLMS.....	31
Tableau III.3. Résumé de l'algorithme de la DT- Geigel	32
Tableau III.4. Résumé de l'algorithme de la DT- NCC.....	32
Tableau III.5. Code MATLAB de l'algorithme de la DTD- Geigel.....	33
Tableau III.6. Code MATLAB de l'algorithme de la DTD- NCC	33
Tableau III.7. La moyenne d'ERLE des deux méthodes Geigel et NCC.....	44
Tableau III.8. La moyenne du MSE dans les deux méthodes Geigel et NCC	44

LISTE DES ABREVIATIONS

AAF	Auxiliary Adaptive Filter
AEC	Acoustic Echo Cancellation
DT	Double-Talk
DTD	Double-Talk Detection
CC	Corrélation croisé
ERLE	Echo Return Loss Enhancement
RIF	Response Impulsionnelle Finie
RLS	Recursive Least Square
MSE	Mean Square Error
NCC	Normalized Cross-Correlation
LMS	Least Mean Square
NLMS	Normalized Least Mean Square
SNR	Signal to Noise Ratio
MIC	Microphone
H-P	Haut parleur
PSTN	Public-Switched Telephone Network
FNCC	Fast Normalized Cross-Correlation
SFE	Speech Feature Extraction
SVD	Singular Value Decomposition
VIRE	Variance Impulse Response Algorithm

LISTE DES NOTATIONS

λ	Facteur d'oubli
γ	Facteur de perméabilité
δ_{NLMS}	Paramètre de régularisation de NLMS
σ^2	Variance
T	Seuil de décision des méthodes de DTD
$E[.]$	Espérance mathématique
$E_n(.)$	L'énergie
$E_s(.)$	L'énergie logarithmique
std (.)	Ecart type
$. $	Valeur absolue
$\ \mathbf{a}\ $	La norme d'un vecteur \mathbf{a}
f	Fréquence d'échantillonnage
μ_{LMS}	Pas d'adaptation de LMS
μ_{NLMS}	Pas d'adaptation de NLMS
$\mu(\mathbf{n})$	Pas d'adaptation variable de NLMS
$\xi(n)$	Décision statistique de DTD
N	Nombre d'itération totale
L	Longueur du filtre adaptatif
$x(n)$	Signal de locuteur lointain
$v(n)$	Signal de locuteur proche
$y(n)$	Signal d'écho acoustique
$\hat{y}(n)$	Signal d'écho estimé
$e(n)$	Signal d'erreur
$b(n)$	Signal de bruit ambiant
$d(n)$	Signal de microphone
$\mathbf{x}(n)$	Observation de longueur L du signal $x(n)$
\mathbf{h}	Réponse impulsionnelle du LEM
$\mathbf{w}(n)$	Coefficients du filtre adaptatif

Introduction

Générale

La révolution technologique a changé toute la dimension des systèmes de communication après l'invention du téléphone en 1876, de nombreuses augmentations ont été mises au point pour améliorer la qualité de la voix ainsi que celle des systèmes. Il y a quelques années, le terme écho fait généralement référence à l'écho hybride (ou de circuit) qui apparaît dans les lignes de transmission d'un téléphone terrestre traditionnel. Au cours des dernières décennies, nous avons bénéficié de la technologie de téléphonie sans fil. La demande des abonnés pour une qualité vocale améliorée sur les réseaux sans fil a conduit à une nouvelle technologie clé appelée annulation d'écho, qui peut fournir une qualité de voix proche de la ligne filaire sur un réseau sans fil. La popularité de la téléphonie mains libres devient de plus en plus forte et les communications personnelles et professionnelles quotidiennes sont très sollicitées. Les systèmes de téléphonie mains libres sont configurés avec un ensemble de haut-parleurs et de microphones où le microphone capte le signal émis par le haut-parleur qui est appelé écho acoustique.

Les annuleurs d'écho acoustiques sont utilisés pour supprimer le signal d'écho du signal du microphone. Le filtrage adaptatif semble la technique la plus appropriée pour résoudre ce problème. Il existe plusieurs types de filtres adaptatifs qui se différencient de leurs performances et leurs complexités de calcul, le NLMS (*Normalized Least Mean Squarre*), assurant un bon compromis, est le mieux placé pour ce genre d'application et qui est un filtre adaptatif à pas d'adaptation variable de type gradient stochastique.

La « double-parole » DT (*Double-Talk*), est un problème majeur présent dans les annuleurs d'écho acoustique (*Acoustic Echo Canceller*) AEC. Cela se produit lorsque le locuteur proche parle seul ou simultanément avec le locuteur lointain. Le problème de la DT est traité par ce qu'on l'appelle « détecteur de double-parole » (*Double-Talk Detector*) DTD qui détecte la double parole et arrête par conséquent l'adaptation des coefficients du filtre adaptatif afin d'éviter sa divergence, Ces méthodes visent à calculer une variable statistique de décision et à la comparer à une valeur de seuil constant. Avec un bon DTD, on peut éviter la divergence du filtre adaptatif et par conséquent améliorer la qualité de la communication.

L'objectif de ce travail est de mettre en œuvre deux algorithmes DTD utilisés dans un AEC basé sur un filtre adaptatif de type NLMS, l'évaluation des performances est effectuée à travers des test sur des séquences audio réelles et avec une réponse impulsionnelle modélisant le chemin réel de l'écho.

Le présent mémoire comporte trois chapitres :

- Le premier chapitre aborde le phénomène de l'écho avec les solutions utilisées pour les annuleurs d'écho acoustiques ainsi que le développement mathématique des filtres adaptatifs.
- Le deuxième chapitre traite les différentes techniques et méthodes utilisé pour la détection DT.
- Le troisième et dernier chapitre présente les résultats de simulation des deux DTDs et la discussion des résultats obtenus.
- Enfin nous terminons par une conclusion générale suivi par les perspectives envisagées.

Chapitre I

Généralités sur les annulateurs
acoustiques

Chapitre I : Généralité sur les annuleurs d'écho acoustique

I.1. Introduction

Dans un système de télécommunication, la qualité de la parole est généralement dégradée à cause de la présence d'écho acoustique et le bruit, ce phénomène est un problème bien connu sur les téléphones mains libres, systèmes de téléconférence, systèmes de sonorisation, et systèmes d'audition et celui-ci généralement indésirable mais inévitable donc l'annulation d'écho est une partie importante de système de communication.

Les supprimeurs d'écho ont été mis au point dans les années 50 en réponse à la première utilisation de satellites pour les télécommunications, mais ils ont été largement supplantés par des annuleurs d'écho plus performants. Dans tous les environnements de communication mains libres l'annulation d'écho acoustique (*Acoustic Echo Cancellations* AEC en anglais) est la méthode utilisée en téléphonie pour améliorer la qualité de la voix en empêchant la création d'écho ou en supprimant celui-ci après sa présence. En plus d'améliorer la qualité de communication (vocale).

Dans un système AEC, un signal de microphone $d(n)$ contient principalement deux signaux : le signal de parole proche $v(n)$ et le signal d'écho acoustique $y(n)$ et la question fondamentale dans ce système est alors d'estimer la réponse impulsionnelle entre le haut-parleur et le microphone de l'appareil. La solution la plus fiable à ce problème consiste à utiliser différents algorithmes de filtrage adaptatif. L'objectif est d'éliminer le signal de parole renvoyé par l'écho du microphone pour que seul le signal de parole proche soit transmis.

Dans ce chapitre nous allons présenter les concepts généraux sur le domaine de l'annulation d'écho acoustique à travers d'abord un aperçu historique sur ce domaine ainsi que la présentation de l'idée générale et de sa formulation mathématique et en fin les deux algorithmes adaptatifs utilisés dans notre application.

I.2. Historique

Les laboratoires Bell ont développé la théorie de l'annuleur d'écho au début des années 1960 ce qui a ensuite abouti à la suppression d'écho en laboratoire à la fin des années 1960 et à des annuleurs d'écho commerciaux dans les années 1980. Un annuleur d'écho génère une estimation de l'écho à partir du signal du locuteur lointain et soustrait cette estimation de la voie du signal du microphone. Cette technique nécessite un traitement adaptatif du signal pour générer un signal suffisamment précis pour annuler efficacement l'écho, où l'écho peut différer

de l'original en raison de divers types de dégradation. Depuis l'invention des algorithmes d'annulation d'écho, les entreprises AT et T Bell Labs ont améliorés ces algorithmes. Comme tous les processus d'annulation d'écho, ces premiers algorithmes ont été conçus pour estimer le signal d'écho qui est gênant pour le locuteur lointain et le soustraire du signal à transmettre afin de l'annuler [2] [3] [4].

I.3. La communication mains libres

La communication mains libres est utilisée dans le monde entier dans les systèmes de télécommunication modernes, par exemple dans les téléphones mobiles, les visiophones, les conférences audio et vidéo, les terminaux de radio mobile ou les systèmes de reconnaissance vocale, et aussi des formations à distance. Un dispositif de communication « mains-libres » peut être défini comme étant un dispositif qui permet à son utilisateur de l'utiliser sans les mains, le développement de systèmes de communication mains libres a donné naissance à un autre type d'écho appelé écho acoustique et qui est classiquement éliminé en utilisant les différents types de filtrage adaptatif [5].

La configuration de base d'un système mains libres typique est illustrée dans la figure I.1.

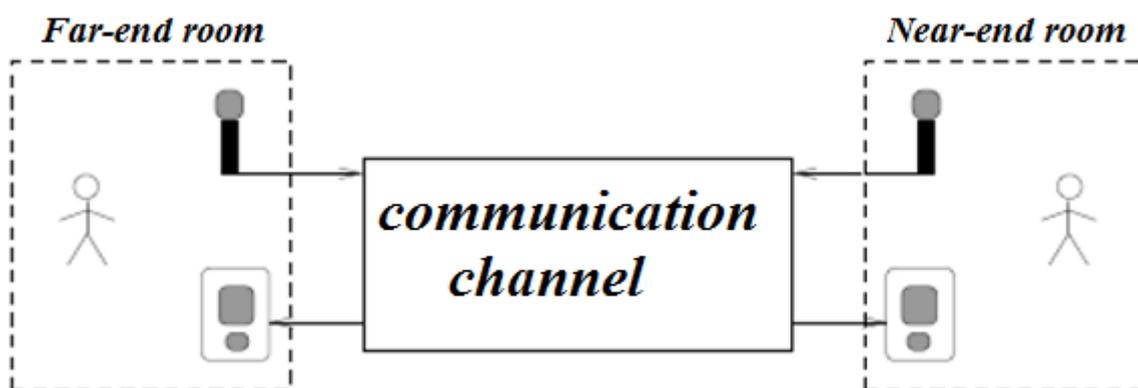


Figure I.1. Configuration de base d'un système mains libres

Chaque partie du système de communication est appelée « End ». L'extrémité distante du haut-parleur s'appelle l'extrémité lointain "Far-end" et l'extrémité proche "Near-end" où les signaux sont mesurés [6].

I.3.1. Problèmes de communication mains libres

Les problèmes majeurs rencontrés dans la plupart des applications de communications mains libres sont : le bruit de fond, la réverbération et l'écho acoustique que nous allons étudier avec plus de détails dans le paragraphe suivant.

I.4. L'écho dans le système des télécommunications

Dans les réseaux de télécommunication on distingue deux types d'échos l'écho électrique et l'écho acoustique, chacun a des origines différentes qu'il est nécessaire de clarifier et de souligner. Quelles que soient ces origines, l'écho peut être caractérisé par la réponse impulsionnelle associé à son trajet entre le point d'émission et d'observation. Cette réponse est très sensible et peut varier rapidement d'un instant à un autre en fonction du type d'écho.

I.4.1. L'écho électrique

Le réseau de télécommunication longue distance, internationale par exemple, est constitué de deux types de liaison téléphonique, liaison téléphonique "à deux fils" et liaison téléphonique "à quatre fils". Le passage d'une liaison à deux fils à une liaison à quatre fils a pour effet de générer un écho dit électrique. Ceci est dû à la désadaptation entre l'impédance présentée par le combiné téléphonique et l'impédance du canal de transmission. Nous retiendrons en outre que le problème de l'écho électrique est classiquement résolu à l'aide d'un annuleur d'écho utilisant des algorithmes du type gradient stochastique plus spécifiques, cet écho se retrouve dans les systèmes de PSTN (*Public-Switched Telephone Network*) comme illustré à la Figure I.2.

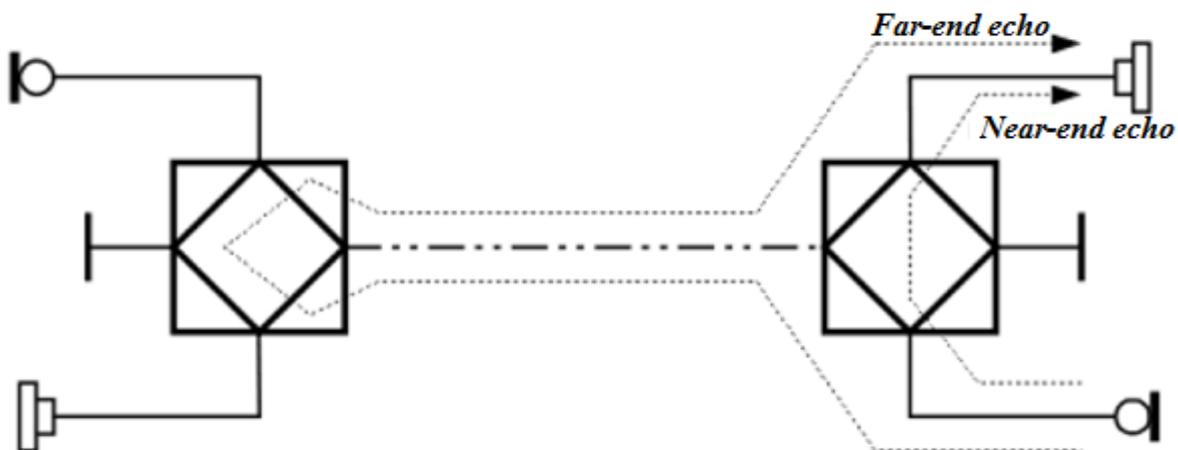


Figure I.2. L'écho électrique

La figure I.3 représente un exemple de réponse impulsionnelle modélisant l'écho électrique.

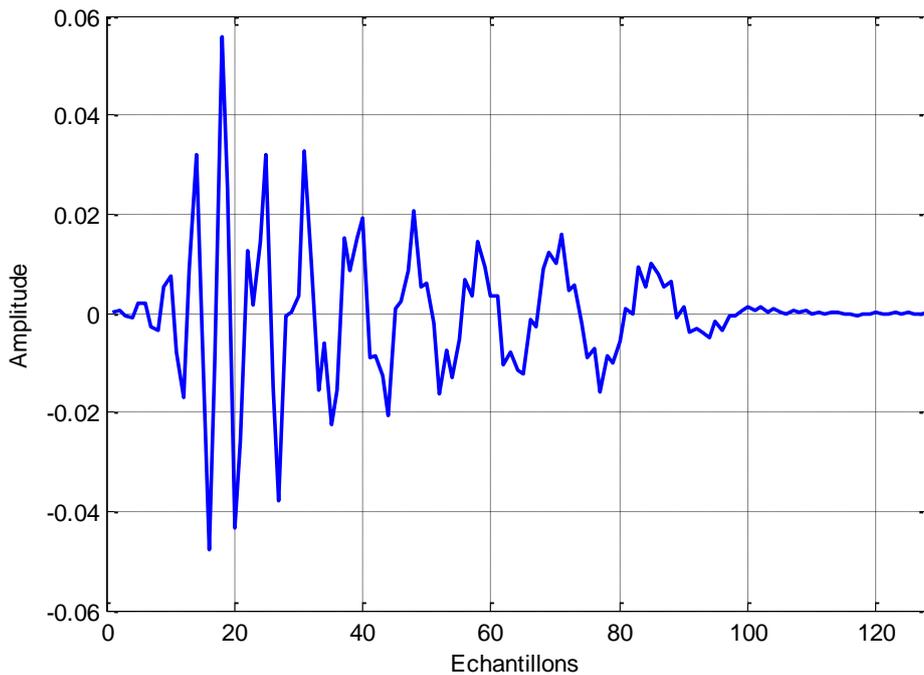


Figure I.3. Exemple de réponse impulsionnelle (4^{ème} chemin d'écho des recommandations G168) [7].

I.4.2. L'écho acoustique

L'origine de l'écho acoustique provient de l'utilisation de nouveaux systèmes de télécommunications dits "mains libres". Au début des télécommunications, l'utilisateur était obligé de coller son oreille à un combiné pour entendre son interlocuteur distant. Aujourd'hui, les nouveaux systèmes de télécommunications permettent la liberté de mouvement du locuteur en restituant le son de l'interlocuteur sur un haut-parleur (figure I.4). Le couplage acoustique, généré par l'utilisation de tels systèmes de télécommunications mains libres, provoque certains effets indésirables comme le phénomène de l'écho acoustique ou encore l'instabilité de la boucle de communication.

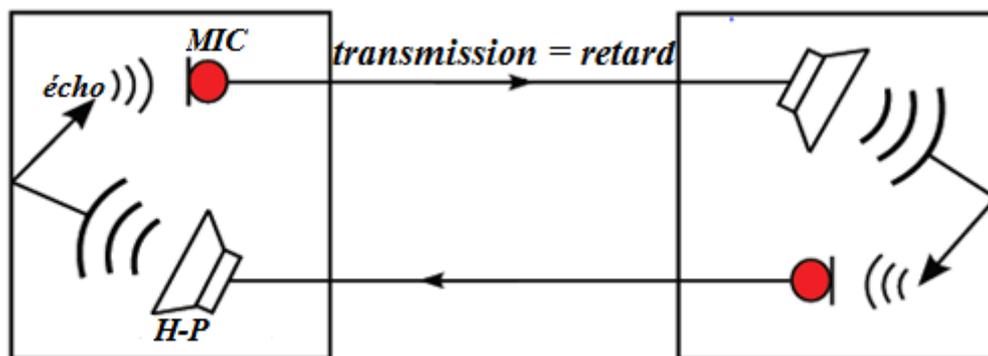


Figure I.4. Exemple de communication bidirectionnelle avec écho acoustique

Des interprétations théoriques précises peuvent être obtenues en faisant appel aux domaines de l'acoustique géométrique, ondulatoire et statistique. Néanmoins, le phénomène physique peut être décrit simplement et succinctement comme suit. Une onde sonore, émise par un émetteur, se propage suivant les lois de l'acoustique vers un récepteur. Au cours de son trajet, l'onde subit l'influence de l'environnement acoustique dans lequel elle se propage. Le phénomène se résume, dans le cas d'une propagation dans un espace libre, à l'absorption d'une onde sonore par l'air qui dépend de paramètres (température, pression atmosphérique, etc...) variant lentement dans le temps par rapport à l'échelle de stationnarité du signal sonore. S'ajoutent à cela des phénomènes de réflexion, diffraction, diffusion, et absorption, provoqués par les parois et obstacles présents dans l'espace clos. Le trajet de propagation d'une onde sonore est appelé canal acoustique. Le canal acoustique est entièrement défini par le milieu de propagation (en l'occurrence l'air), l'espace clos (sa géométrie, ses propriétés acoustiques), la source (dans le cas de l'écho, le haut-parleur) et le récepteur (dans le cas de l'écho, le microphone). Les caractéristiques de la réponse impulsionnelle associée au trajet d'écho acoustique dépendent directement de ces différents paramètres.

La réponse impulsionnelle d'un canal acoustique se présente sous la forme d'une onde directe et d'une succession d'ondes réfléchies par les parois d'une salle particulière.

La figure I.5 représente un exemple de réponse impulsionnelle acoustique d'une voiture qui échantillonnée à 16 KHz et qui est mesuré dans des conditions réelles [8].

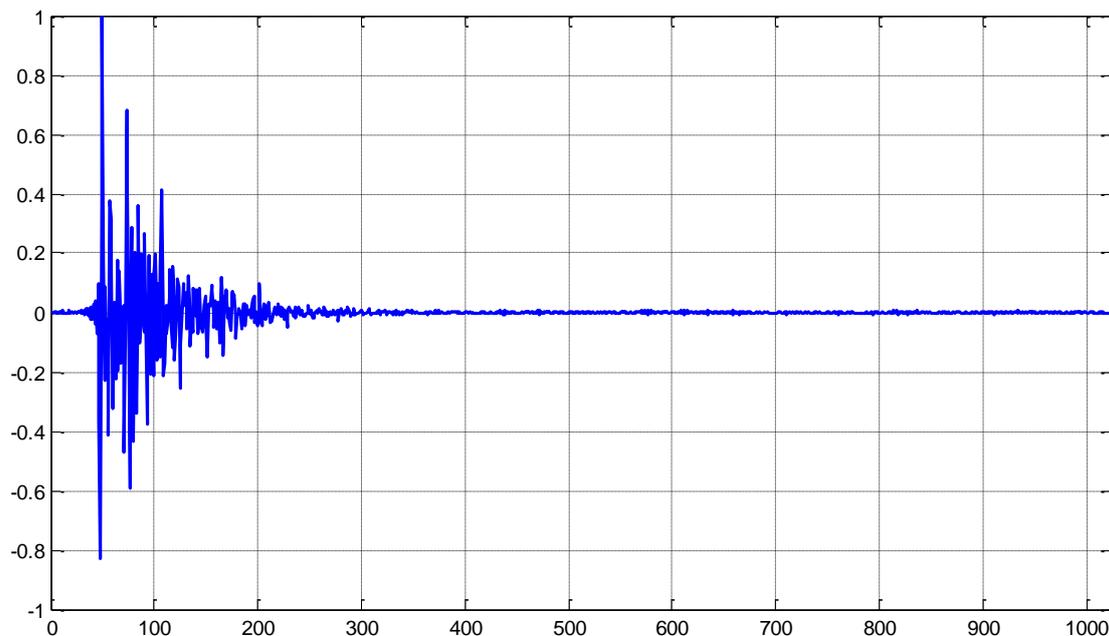


Figure I.5. Exemple de réponse impulsionnelle acoustique d'une voiture.

I.4.3. L'écho acoustique et les systèmes téléphoniques

Dans une conversation téléphonique. Le discours provenant du locuteur à l'extrémité lointaine est amplifié et reproduit par un haut-parleur, le signal vocal est transmis via un réseau de communication à un appareil équipé d'au moins un H-P et d'un microphone. Le haut-parleur joue le son du locuteur lointain au locuteur proche, tandis que le microphone enregistre la voix du haut-parleur proche. La voix du locuteur proche est ensuite transmise au locuteur lointain. Mais dans certains cas, une partie du son émis par le haut-parleur se propage dans l'environnement proche et est couplée au microphone de l'appareil. De ce fait, le locuteur lointain ne reçoit pas seulement la voix du locuteur local, mais reçoit également une version différée de sa voix, cet effet est appelé écho acoustique [9]. Comme illustré à la figure I.6, ce couplage est composé du chemin direct et des chemins réfléchis entre le haut-parleur et le microphone.

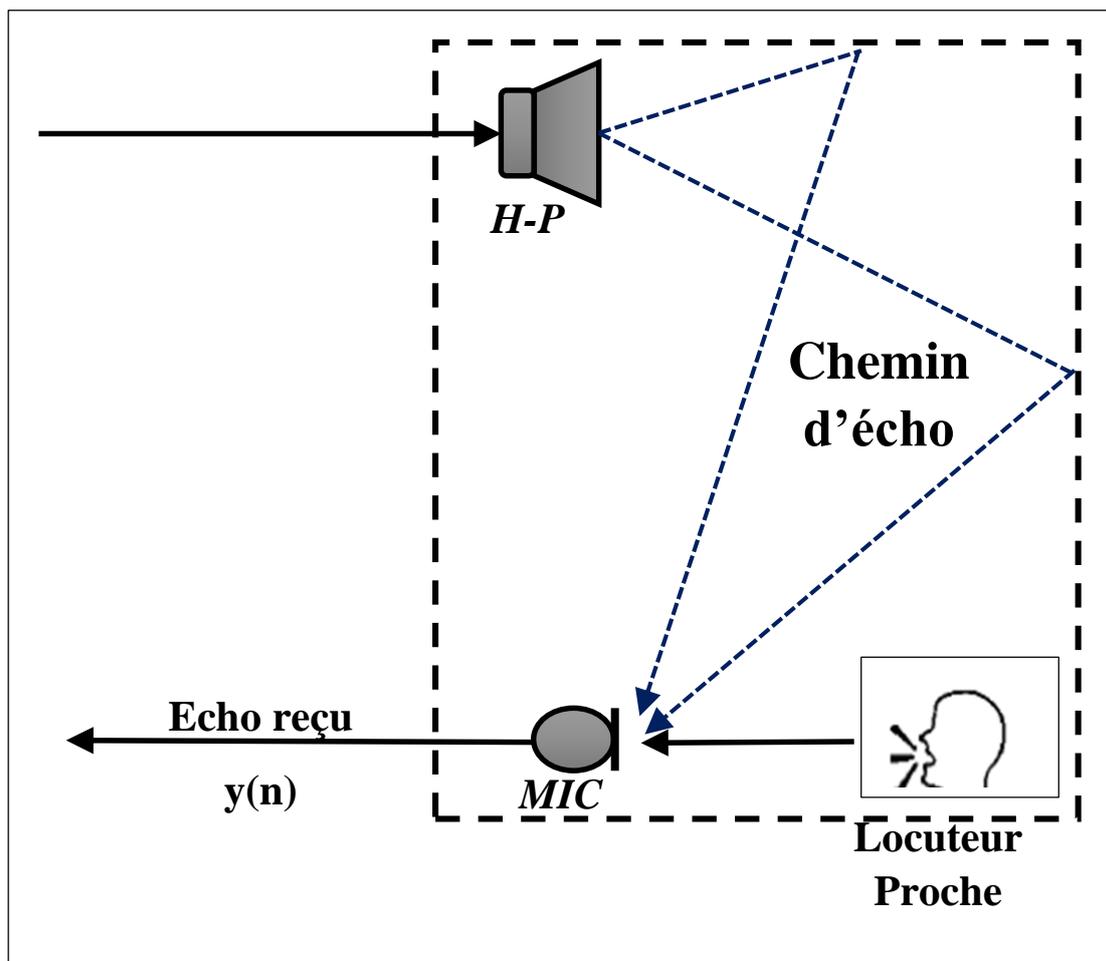


Figure I.6. Le schéma de l'écho acoustique

I.4.4. L'environnement acoustique

L'utilisation des téléphones munis d'un haut-parleur s'est généralisée ces dernières années car ce type d'appareil permet notamment d'organiser des téléconférences. Leur essor est également dû à la démocratisation des téléphones mobiles qui sont, entre autres, utilisés dans les voitures. Dans ce cas, l'aspect mains-libres est un élément primordial de sécurité alors qu'il s'agit plus de confort dans le cadre des téléconférences. Aujourd'hui, la plupart des téléphones à haut-parleur sont munies d'un annuleur d'écho similaire à ceux utilisés pour l'annulation d'écho électrique. Cependant, la nature même de l'environnement acoustique rend la suppression d'écho difficile. Les personnes qui prennent part à des téléconférences doivent observer des règles de dialogue afin de conserver une conversation intelligible [10].

Malheureusement, le contexte dans lequel apparaît l'écho acoustique est très contraignant :

- Le trajet de l'écho est très long (125 millisecondes ou plus).
- Il est extrêmement variable au cours du temps.
- Le bruit de fond du signal near-end (locuteur proche) est très important.
- Une attention particulière doit être portée sur la prévention des sifflements.

I.4.5. Le chemin parcouru par l'écho

La durée importante du trajet de l'écho est principalement due à la faible vitesse du son dans l'air. En effet, le son se déplace dans l'air à une vitesse beaucoup plus faible que celle des ondes électromagnétiques qui se propagent le long d'un fil de cuivre puisque cette dernière est proche de la vitesse de la lumière. L'onde acoustique est de plus ralentie par des réflexions sur obstacles comme des murs, des chaises, des meubles, etc. Le signal réfléchi a une énergie qui dépend des dimensions de la salle ainsi que des objets qu'elle contient qui ont un coefficient de réflexion qui leur est propre.

Dans la plupart des salles, l'intensité du son réfléchi décroît exponentiellement. Typiquement, le délai de réverbération est de l'ordre de 200 à 300 millisecondes. De plus, la réponse impulsionnelle caractérisant la salle ne reste pas constante au cours du temps et elle varie en fonction de la température, du taux d'humidité, de la pression, additionné à cela, tout mouvement d'objets ou de personnes, toute modification de l'emplacement du microphone ou des haut-parleurs provoque une variation brutale de la réponse impulsionnelle [10].

I.5. Annulation d'écho acoustique (AEC)

L'annulation d'écho consiste à éliminer l'écho dans une communication vocale de façon à améliorer la qualité de l'appel. L'annulation d'écho est souvent nécessaire car les techniques de compression de la voix et les délais dans le traitement des paquets génèrent un écho.

I.5.1. Principe de fonctionnement d'AEC

L'annulation d'écho fonctionne sur le principe de la détection du signal d'origine et du signal retransmis qui réapparaît avec un peu de délai, et en le supprimant du signal transmis ou reçu. L'annulation d'écho améliore non seulement la qualité de la communication mais elle réduit aussi l'usage de bande passante grâce à sa technique de suppression de l'écho.

La figure I.7 schématise d'une manière détaillée le système d'annulation d'écho : $x(n)$ est le signal provenant du locuteur lointain (ou distant) transmis au haut-parleur, $y(n)$ le signal d'écho perçu par le microphone. Ce signal d'écho est la résultante de la convolution entre le signal d'excitation $x(n)$ et la fonction de transfert \mathbf{h} (dans le cas de l'audioconférence c'est la salle et en téléphonie c'est le canal de transmission), dont la réponse impulsionnelle est modélisée comme un filtre à réponse impulsionnelle finie (RIF). A ce signal d'écho vient s'ajouter un signal local $z(n)$, composé du bruit ambiant $b(n)$ et du signal utile $v(n)$ qui représente la voix du locuteur proche (ou la double parole), et la résultante fournit le signal du microphone $d(n)$ le signal réverbéré dans l'environnement acoustique, et $W(n)$ représente le filtre adaptatif à l'instant n utilisé pour annuler le signal d'écho. A chaque itération, le signal d'erreur, $e(n)$, est renvoyé dans le filtre, et par conséquent les coefficients du filtre sont modifiés.

$$\begin{aligned} d(n) &= y(n) + z(n) \\ &= y(n) + v(n) + b(n) \\ &= \mathbf{h}^T \mathbf{X}_L(n) + v(n) + b(n) \end{aligned} \quad (\text{I.1})$$

Où :

$\mathbf{h} = [h_0 \ h_1 \ \dots \ h_{L-1}]$ Réponse impulsionnelle du chemin d'écho.

$\mathbf{x}_L(n) = [x(n) \ x(n-1) \ \dots \ x(n-L+1)]$ Vecteur des L derniers échantillons du signal far-end $x(n)$, avec L est l'ordre du filtre \mathbf{h} .

Nous avons aussi :

$$\hat{y}(n) = \mathbf{w}^T \mathbf{x}(n) \quad (\text{I.2})$$

Où:

$\mathbf{w} = [w_0 \ w_1 \ \dots \ w_{L-1}]$ les coefficients du filtre adaptatif.

On obtient ainsi le signal d'erreur :

$$e(n) = d(n) - \hat{y}(n) \quad (\text{I.3})$$

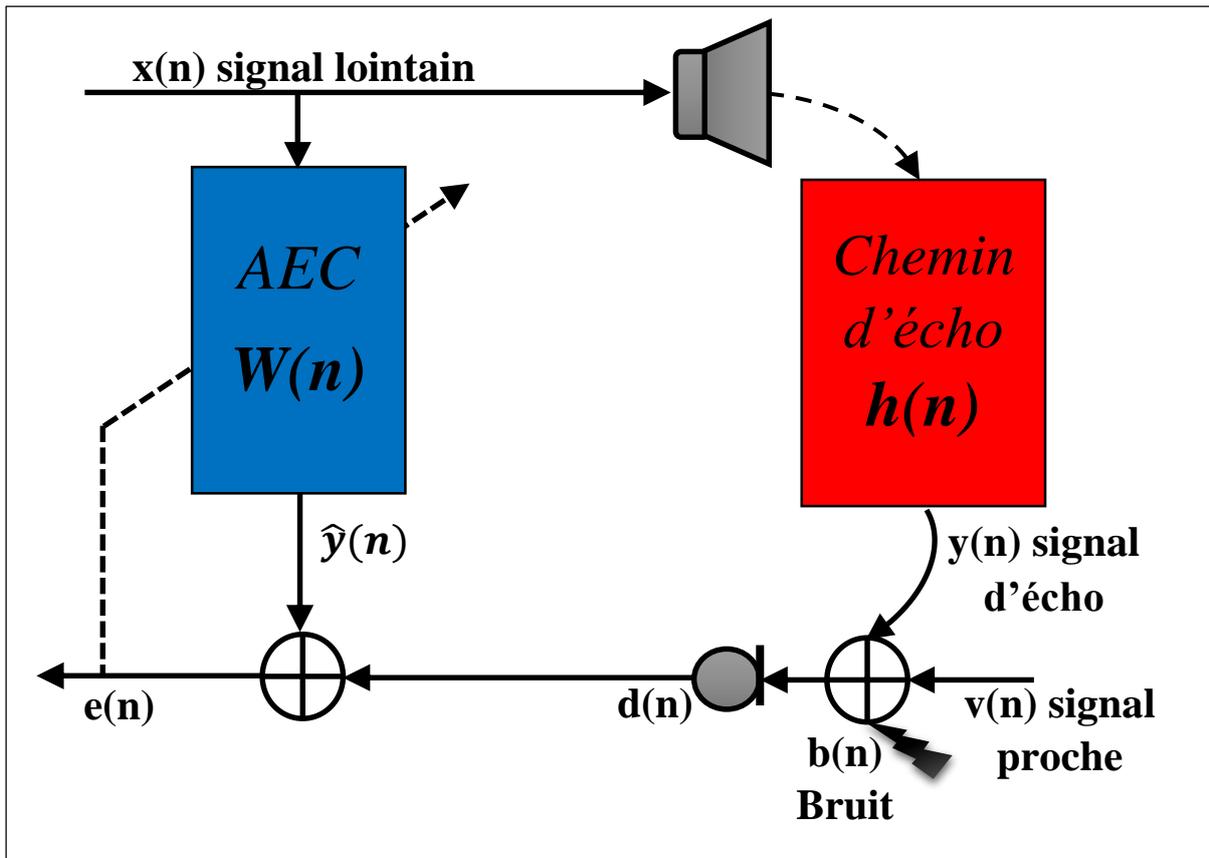


Figure I.7. Système d'annulation d'écho acoustique (AEC)

En général, tous les algorithmes d'annulation d'écho sont basés sur des filtres à coefficients adaptatifs qui fonctionnent assez bien. Cependant, la façon d'adapter les coefficients influence terriblement les performances.

I.5.2. Filtrage adaptatif

Un filtre adaptatif est un filtre numérique, dont les coefficients sont estimés selon un critère spécifique pour s'adapter aux variations des signaux reçus. Habituellement, un vecteur d'entrée et une réponse désirée, ces coefficients sont utilisés pour définir un vecteur d'erreur qui contrôle alors l'évolution des paramètres du filtre adaptatif.

Dans la figure I.8 suivante on peut voir un schéma simplifié d'un filtre adaptatif, où $d(n)$ représente le signal désiré (référence), $y(n)$ la sortie du filtre numérique, et $e(n)$ le signal d'erreur.

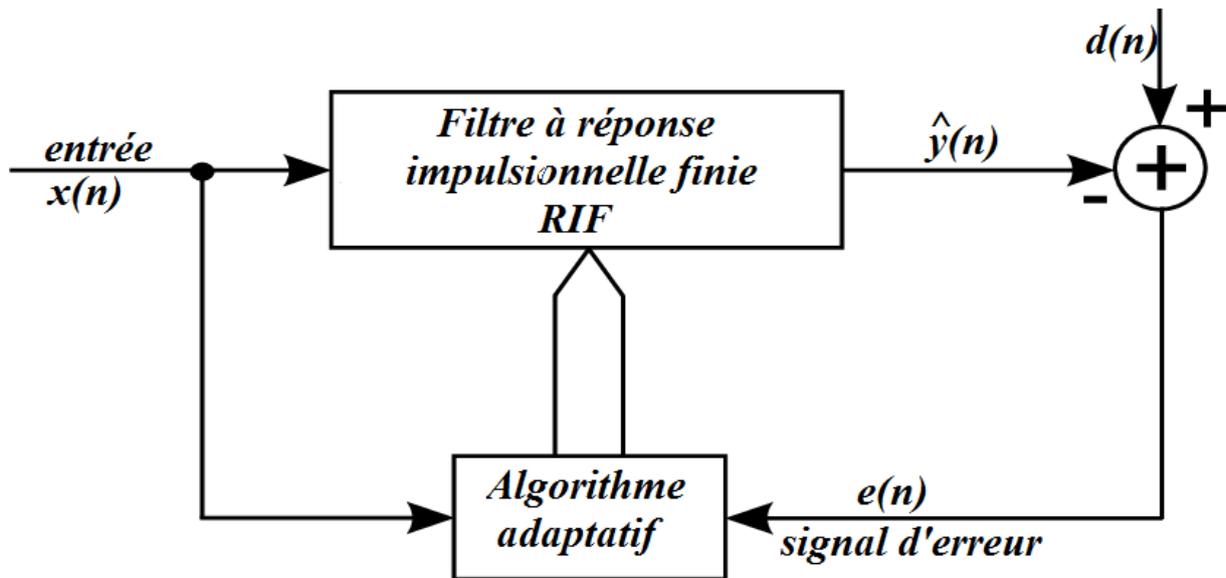


Figure I.8. Principe d'un filtre adaptatif [11].

L'algorithme adaptatif se comporte comme un opérateur humain qui a la capacité de s'adapter à un environnement changeant.

Les filtres adaptatifs peuvent être classés en fonction des choix qui sont faits sur les points suivants :

- Le critère d'optimisation.
- L'algorithme de mise à jour des coefficients.
- La structure du filtre programmable.
- Le type de signal traité, mono ou multidimensionnel.

Le filtrage adaptatif est maintenant utilisé dans de nombreux domaines d'application du traitement du signal, tels que :

- Identification des systèmes.
- Contrôle du signal.
- Annulation d'écho.
- Prédiction du signal.
- Annulation de retour adaptative.
- Suppression du bruit.
- Communications.

I.5.2.1. L'algorithme adaptatif

L'algorithme adaptatif est le cœur d'une AEC, qui décide de la convergence et du comportement de l'AEC. Le comportement de suivi indique à quelle vitesse le filtre adaptatif peut suivre les dislocations alors que le comportement de convergence est étudié comme un premier ajustement du filtre adaptatif à la réponse impulsionnelle du système.

Les algorithmes qui sont couramment utilisés sont principalement classés en deux catégories : les algorithmes de type gradient stochastique (*Least Mean Square* LMS) et les algorithmes de type moindres carrés récursives (*Recursive Least Squares* RLS) [12].

L'algorithme adaptatif de l'AEC devrait donc atteindre les caractéristiques de performance suivantes :

- Convergence rapide et capacité de suivi
- Faible ajustement à l'état d'équilibre
- Robustesse contre les variations des signaux near-end (proches)
- Faible complexité de calcul

I.5.3. Algorithmes de filtrage adaptatif standard

I.5.3.1. L'Algorithme LMS

L'algorithme LMS est initialement proposé par Widrow Hoff en 1959. Il est utilisé pour déterminer le minimum de l'erreur quadratique moyenne et est basé sur la technique de recherche par descente de gradient. Cet algorithme est utilisé en raison de son implémentation facile, sa simplicité et sa faible complexité de calcul. Si $x(n)$ est le vecteur du signal de parole d'entrée et $w(n)$ est le vecteur de poids du filtre adaptatif à l'instant n .

L'équation de mise à jour du vecteur de poids est donnée par :

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu_{LMS} \mathbf{X}_L(n)e(n) \quad (\text{I.4})$$

Où μ_{LMS} est le pas d'adaptation de l'algorithme adaptatif.

Tous les signaux sont supposés avoir une valeur réelle pour la simplicité. Une condition suffisante et nécessaire pour que l'algorithme converge consiste à faire converger tous les poids du filtre vers la solution réelle, ce qui conduit aux limites suivantes du pas d'adaptation :

$$0 < \mu_{LMS} < \frac{2}{\lambda_{\max}} \quad (\text{I.5})$$

D'où $\frac{2}{\lambda_{\max}}$ est la plus grande valeur propre de la matrice de corrélation du vecteur $x(n)$

Tableau I.1. L'algorithme LMS

Initialisation : $\mathbf{w}(0) = 0$
Paramètre : $0 < \mu_{LMS} < \frac{2}{\lambda_{\max}}$
La sortie du filtre adaptatif : $\hat{y}(n) = \mathbf{w}^T(n) \cdot \mathbf{X}_L(n)$
Erreur : $e(n) = d(n) - \hat{y}(n)$
Mise à jour du filtre : $\mathbf{w}(n+1) = \mathbf{w}(n) + \mu_{LMS} \mathbf{X}_L(n) e(n)$

I.5.3.2. L'algorithme normalisé (NLMS)

L'algorithme NLMS est une version d'extension de LMS (*Normalized LMS*). C'est l'un des algorithmes de filtrage adaptatifs les plus populaires dans les applications de traitement du signal, en raison de sa simplicité et de sa robustesse. L'algorithme NLMS montre une stabilité beaucoup plus grande avec des signaux inconnus. Le pas d'adaptation est non plus une constante, comme dans le LMS standard, mais une fonction de la puissance du signal d'entrée $x(n)$ comme suit :

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu_{NLMS}}{\mathbf{X}_L^T(n) \mathbf{X}_L(n) + \delta_{NLMS}} \mathbf{X}_L(n) e(n) \quad (\text{I.6})$$

Où δ_{NLMS} est une constante de régularisation, qui est un très petit nombre positif de valeur typique égale à une constante multipliée par la variance du signal d'entrée, elle est utilisée pour éviter la division par zéro lorsque l'entrée est égale à zéro [13].

En examinant une condition nécessaire mais non suffisante pour la stabilité de l'algorithme LMS, nous pouvons en déduire une condition nécessaire pour la stabilité de l'algorithme NLMS. Une condition intuitive pour la stabilité est d'avoir :

$$|\varepsilon(n)| < |e(n)| \quad (\text{I.7})$$

$$\text{Avec : } \varepsilon(n) = d(n) - \mathbf{X}_L^T(n) \mathbf{w}(n+1) \quad (\text{I.8})$$

$\varepsilon(n)$ Est le signal d'erreur a posteriori, calculé après l'adaptation du filtre. En remplaçant l'équation (I.4) par l'équation (I.8), et en utilisant la condition (I.7), nous trouvons :

$$0 < \mu_{LMS} < \frac{2}{\mathbf{X}_L^T(n) \mathbf{X}_L(n)} \quad (\text{I.9})$$

En comparant les termes de mise à jour des poids des algorithmes LMS et NLMS, nous constatons que la taille du pas du LMS, μ_{LMS} est comparable à $\frac{\mu_{NLMS}}{\mathbf{X}_L^T(n)\mathbf{X}_L(n) + \delta_{NLMS}}$ terme de mise à jour NLMS, si nous écartons la constante de régularisation et prenons en compte l'inégalité (I.9), nous trouvons qu'une condition nécessaire mais non suffisante pour la stabilité de l'algorithme NLMS est que le paramètre de taille de pas normalisé se situe dans les limites suivantes:

$$0 < \mu_{NLMS} < 2 \quad (\text{I.10})$$

Tableau I.2. L'algorithme NLMS

Initialisation :	$\mathbf{w}(0) = 0$
Paramètres :	$0 < \mu_{NLMS} < 2$ $\delta_{NLMS} = cst.\sigma_x^2$
La sortie du filtre adaptatif :	$\hat{y}(n) = \mathbf{w}^T(n).\mathbf{X}_L(n)$
Erreur :	$e(n) = d(n) - \hat{y}(n)$
Mise à jour du filtre :	$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu_{NLMS}}{\mathbf{X}_L^T(n)\mathbf{X}_L(n) + \delta_{NLMS}} \mathbf{X}_L(n)e(n)$

Avantages et inconvénients :

- L'algorithme NLMS a une faible complexité de calcul, avec une bonne vitesse de convergence, ce qui le rend bien pour l'annulation d'écho.
- Il y a une erreur minimale en régime permanent.
- L'amplification du bruit devient plus petite en raison de la présence d'une taille de pas normalisée [14].

Complexité de calcul :

- NLMS nécessite plus de calculs pour l'évaluation que l'algorithme LMS en raison de la présence de la puissance du signal de référence.
- L'algorithme NLMS nécessite $3L+1$ multiplications, soit L fois plus que l'algorithme LMS. Où L est la longueur du vecteur de coefficient [14].

I.6. Conclusion

Dans ce chapitre, nous avons présenté une synthèse globale sur la conception des annuleurs d'écho acoustique en utilisant des algorithmes de filtrage adaptatif. L'annulation de l'écho acoustique a une large gamme d'applications, telles que les téléphones mobiles, les haut-parleurs, les adaptations pour voitures mains libres, les accessoires Bluetooth, les têtes d'écoute et les systèmes de téléconférence multicanaux.

L'objectif principal de faire appel à des algorithmes adaptatifs est de réduire l'erreur quadratique moyenne tout en essayant de faire un compromis entre un taux de convergence plus élevé et une complexité de calcul moindre. Les algorithmes LMS sont faciles à implémenter, mais avec une vitesse de convergence moindre et les algorithmes NLMS sont caractérisés par une stabilité moins grande et une vitesse de convergence plus élevée.

Chapitre II

Les détections de double
parole

Chapitre II : Les Détections de Double Parole

II.1. Introduction

Le majeur problème d'un annuleur d'écho acoustique (*Acoustic Echo Canceller AEC*) à côté de la convergence et de la réduction de complexité de calcul, est le problème de la détection de double parole (*Double Talk Detection DTD*). La tâche principale d'une DTD consiste à arrêter la mise à jour des coefficients du filtre lors de la présence de la parole proche afin d'éviter la divergence du filtre. La détection de double parole est un élément clé dans un système d'annulation d'écho acoustique. Un DTD est donc utilisé pour détecter le moment où le signal distant est corrompu par la parole proche.

Ce chapitre présentera le concept général d'un DTD ainsi que les différentes techniques et algorithmes les plus populaires comme l'algorithme Geigel, la DTD basé sur la corrélation croisée et la corrélation croisée normalisée.

II.2. La double parole

La double parole *Double Talk* (DT) se produit pendant les périodes où le locuteur lointain et le locuteur proche parlent simultanément. Dans un système en full-duplex, il est souvent nécessaire de geler les coefficients de filtre adaptatif, de sorte qu'il ne se produise pas de divergence des poids du filtre adaptatif lors d'une DT. La forme de contrôle DT la plus radicale est le Push-To-Talk (mode half-duplex ou single-talk), qui était la norme utilisée jusqu'à l'apparition de filtres adaptatifs pour la suppression de l'écho. Avant l'apparition de l'annulation d'écho, le full duplex était impossible pour un haut-parleur, simplement parce qu'il aurait provoqué des réactions. Le « son du haut-parleur » entrant, ainsi que les effets de la pièce, ne peut pas être éliminé avant la transmission du microphone [15].

C'est la raison pour laquelle le full duplex est l'objectif de toutes les télécommunications et également lors des téléconférences. Les conversations longue distance en full-duplex sont désormais possibles grâce à une bonne technologie numérique et à l'annulation d'écho.

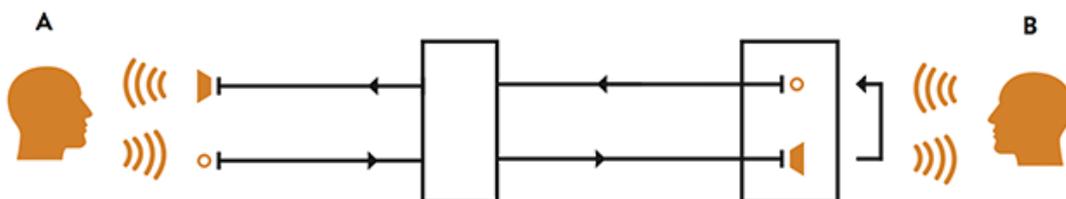


Figure I.1. Schéma de la double parole.

II.3. Détection double parole

Les périodes au cours auxquelles le signal proche $v(n)$ et le signal écho $y(n)$ sont présents dans le signal du microphone $d(n)$ sont appelées périodes de double parole. La tâche d'un DTD consiste à identifier ces périodes et à désactiver l'adaptation du filtre adaptatif en conséquence. Ainsi, l'algorithme DTD est utilisé pour empêcher le filtre adaptatif d'utiliser le signal d'erreur $e(n)$ pour l'adaptation lorsqu'il contient le signal proche. Cela empêche l'algorithme adaptatif de diverger tout en maintenant la qualité du signal sans écho [16].

Le plus difficile dans un annuleur d'écho acoustique est de savoir quand arrêter l'adaptation du filtre. Le filtre ne doit être adapté que lorsqu'il y a le signal lointain, mais pas lorsqu'il y a à la fois le signal lointain et proche c'est ce qu'on appelle la double parole. Le signal proche ferait échouer le processus d'estimation du système et produire des résultats extrêmement erronés et pour cela un détecteur de double parole (DTD) est nécessaire.

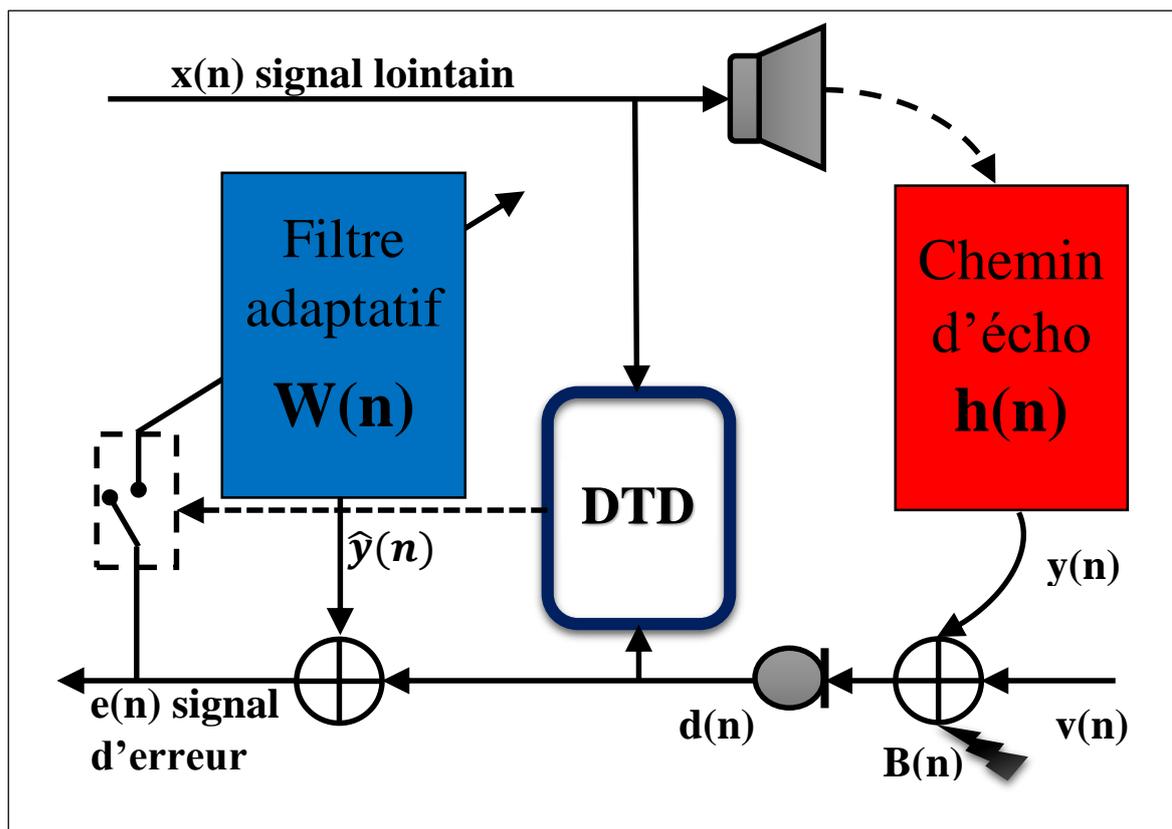


Figure I.2. Structure de base de l'AEC avec DTD

II.3.1. Schémas génériques de la détection de double parole

Généralement tous les types de détecteurs de double parole fonctionnent de la même manière. Par conséquent, la procédure générale de gestion de la double parole est décrite par les quatre étapes suivantes :

1. Une variable statistique de détection, ξ , est formée en utilisant les signaux disponibles tels que x , d et e et les coefficients de filtre estimés \mathbf{w} .
2. La variable statistique de détection est comparée à un seuil prédéfini T (une constante), et la double parole est déclarée si $\xi < T$.
3. Une fois que la double parole est détectée, elle est déclarée pour une période minimale de temps t_{hold} . Tant que la détection est maintenue, l'adaptation du filtre est désactivée.
4. Si $\xi \geq T$ pendant un temps t_{hold} , le filtre reprend son adaptation tandis que la comparaison de ξ à T continue jusqu'à ce que $\xi < T$ à nouveau.

En l'absence de double parole, la DTD n'interrompt pas le fonctionnement de l'AEC. Le temps de maintien t_{hold} des étapes 3 et 4 est essentiel pour minimiser les fausses détections dus au comportement bruyant de la variable statistique de détection. Bien qu'il y ait quelques variations possibles la plupart des algorithmes DTD conservent cette forme de base et ne diffèrent que par la façon dont ils forment la variable statistique de détection.

Une variable de décision est considérée optimale pour la détection de la double parole doit se comporter comme suit :

- si $v(n)=0$ (la double parole n'est pas présente), $\xi \geq T$
- si $v(n) \neq 0$ (la double parole est présente), $\xi < T$
- ξ est insensible aux variations du chemin d'écho.

Le contrôle du filtre adaptatif par DTD est défini comme suit :

$$\text{Contrôle} = \begin{cases} \xi(n) \geq T, & \text{DTD} = 0 \text{ adaptation } \mu \neq 0 \\ \xi(n) < T, & \text{DTD} = 1 \text{ non adaptation } \mu = 0 \end{cases}$$

Le seuil T doit être une constante, indépendante des données en plus ξ doit être insensible aux variations du chemin d'écho lorsque $v(n)=0$ [17][18].

Les sections suivantes présentent différents algorithmes de DTD, tels que l'algorithme Geigel, la méthode de corrélation croisée (CC), la méthode de corrélation croisée normalisée (NCC), la méthode basée sur l'extraction de fonction vocale (SFE), la méthode basée sur décomposition en valeur singulière (SVD) et la méthode basée sur le filtre adaptatif auxiliaire (AAF).

II.4. Les méthodes de la détection de double parole

II.4.1. Algorithme de Geigel

Geigel est l'un des algorithmes les plus populaires de DTD, qui compare l'amplitude du signal du microphone $d(n)$ à l'histoire récente du signal lointain $x(n)$ [19], où la variable de décision de ce dernier est définie comme suit :

$$\xi_G = \frac{\max\{|x(n)|, \dots, |x(n-L_g+1)|\}}{|d(n)|} \quad (\text{II.1})$$

Où L_g et T sont des constantes convenablement choisies. Ce détecteur est basé sur la comparaison des niveaux d'énergie entre le signal du microphone $d(n)$ et le signal du locuteur lointain $x(n)$, en supposant que la parole proche $v(n)$ dans le signal du microphone est plus forte que l'écho.

Le maximum, ou norme, des L_g échantillons les plus récents de $x(n)$ est choisi pour la comparaison en raison du retard incertain sur le chemin de l'écho. Le seuil T , est utilisé pour compenser le niveau d'énergie de la réponse du chemin d'écho \mathbf{h} , et est souvent réglé à 1/2 pour les anneaux d'écho électrique.

La DTD par cette méthode est définie comme suit :

$$DTD = \begin{cases} \xi_G(n) \geq T_G, & DTD = 0 \\ \xi_G(n) < T_G, & DTD = 1 \end{cases}$$

La situation DT est déclarée si $\xi_G(n) < T_G$, où T_G est une valeur de seuil de cette méthode.

Cependant, pour un AEC, il n'est pas facile de définir un seuil universel qui fonctionnera de manière fiable dans toutes les situations, car la perte par le chemin d'écho acoustique peut varier considérablement en fonction de nombreux facteurs. Pour L_g , un choix facile consiste à le régler égal à la longueur du filtre adaptatif.

II.4.2. Algorithme Corrélation croisée (*Cross-Correlation*)

Cette méthode utilise le vecteur de coefficient de corrélation croisée entre $x(n)$ et $e(n)$ comme moyen de détection de double parole [20], qui est défini comme suit :

$$\begin{aligned} \mathbf{C}_{xe}(n) &= \frac{E\{x(n)e(n)\}}{\sqrt{E\{x^2(n)\}E\{e^2(n)\}}} \\ &= [\mathbf{C}_{xe,0}, \mathbf{C}_{xe,1}, \dots, \mathbf{C}_{xe,L_h-1}]^T \end{aligned} \quad (\text{II.2})$$

Où C_{xe} est le coefficient de corrélation croisée entre $x(n-0)$ et $e(n)$; comme pour l'algorithme Geigel, le vecteur de corrélation croisée est calculé sur une longueur de $x(n)$ pour assurer la couverture des échantillons précédents L_h échantillons précédents.

Une norme appropriée, telle que la norme de l'infini, est ensuite appliquée au vecteur résultant $C_{xe}(n)$ pour produire une variable de décision scalaire, par exemple :

$$\xi_{xe}(n) = \|C_{xe}(n)\|_{\infty} \quad (\text{II.3})$$

La variable de décision $\xi_{xe}(n)$ est ensuite comparée à un seuil prédéfini ;

$$\text{si } \left\{ \begin{array}{ll} \xi_{xe}(n) \geq T, & \text{l'adaptation du filtre est autorisée} \\ \xi_{xe}(n) < T, & \text{l'adaptation du filtre est stoppée} \end{array} \right\}$$

Cette logique de décision est basée sur l'hypothèse que les signaux $v(n)$ et $x(n)$ ne sont pas corrélés.

La corrélation croisée entre $x(n)$ et $d(n)$ a également été proposée comme moyen de contrôler l'adaptation dans un AEC [21], définie comme suit :

$$\xi_{xy}(n) = \left\| \frac{E\{x(n)y(n)\}}{\sqrt{E\{x^2(n)\}E\{y^2(n)\}}} \right\|_{\infty} \quad (\text{II.4})$$

Contrairement à $\xi_{xe}(n)$, cette variable de décision indique explicitement la DT pour contrôler l'adaptation,

$$\text{si } \left\{ \begin{array}{ll} \xi_{xy}(n) < T, \text{ indiquant } v(n) > 0, & \text{l'adaptation est arrêtée} \\ \xi_{xy}(n) \geq T, \text{ indiquant } v(n) = 0, & \text{l'adaptation est autorisée} \end{array} \right\}$$

Le problème fondamental de cette méthode est que les vecteurs de coefficients de corrélation croisée ne sont pas bien normalisés. En général, on suppose que $\xi \leq 1$. Par conséquent, si $v(n)=0$, cela ne signifie pas que $\xi = 1$ ou toute autre valeur connue. La valeur de ξ n'est pas connue en général.

L'importance de la corrélation dépend grandement des statistiques du signal et du chemin d'écho. En conséquence, la meilleure valeur de T varie d'une expérience à l'autre. Il n'y a pas de niveau de seuil naturel associé à la variable lorsque $v(n)=0$.

Ces complexités mènent à un autre algorithme DTD, appelé corrélation croisée normalisée, cette méthode est simplement une modification de la méthode existante de corrélation croisée.

II.4.3. Algorithme de cross-corrélation normalisé (NCC)

L'algorithme de corrélation croisée normalisé (*Normalized Cross-Correlation, NCC*), calcule la variable de décision en fonction des relations entre le signal du microphone $d(n)$ et le signal d'erreur $e(n)$. Elle peut être approchée en considérant les valeurs de la variance du signal proche et la corrélation croisée entre le signal d'erreur et le signal du microphone. La NCC détecte la DT et possède une variable de décision convenablement normalisée, en ce sens que pour $v(n) = 0$, $\xi(n) \approx 1$ et pour $v(n) \neq 0$, $\xi(n) < 1$ [22]. La variable de décision NCC est expliquée par la dérivation suivante : Supposons que $v(n) = 0$

$$\text{Donc} \quad \sigma_y^2 = E\{y^2(n)\} = \mathbf{h}^T \mathbf{R}_{xx} \mathbf{h} \quad (\text{II.5})$$

$$\text{Tel que :} \quad y(n) = \mathbf{h}^T x(n), \text{ et } \mathbf{r}_{xx} = \mathbf{R}_{xx} \mathbf{h}$$

(II.5) peut-être réécrite comme suit :

$$\sigma_y^2 = \mathbf{r}_{xy}^T \mathbf{R}_{xx} \mathbf{r}_{xy} \quad (\text{II.6})$$

En générale pour $v(n) \neq 0$, (II.4) devient :

$$\sigma_y^2 = \mathbf{r}_{xy}^T \mathbf{R}_{xx}^{-1} \mathbf{r}_{xy} + \sigma_v^2 \quad (\text{II.7})$$

En divisant (II.6) par (II.7) et en prenant la racine carrée, on obtient la variable de détection normalisée de corrélation croisée suivante :

$$\xi = \sqrt{\mathbf{r}_{xy}^T (\sigma_y^2 \mathbf{R}_{xx})^{-1} \mathbf{r}_{xy}} \quad (\text{II.8})$$

En remplaçant (II.5) et (II.7) par (II.8), la variable de détection peut être exprimée par

$$\xi = \frac{\sqrt{\mathbf{h}^T \mathbf{R}_{xx} \mathbf{h}}}{\sqrt{\mathbf{h}^T \mathbf{R}_{xx} \mathbf{h} + \sigma_v^2}} \quad (\text{II.9})$$

On peut déduire de l'équation (II.6) que pour $v(n) = 0$, $\xi = 1$ et pour $v(n) > 0$, $\xi < 1$, démontrant que la variable de décision NCC est correctement normalisée.

Théoriquement, cette variable de décision est également immunisée contre les modifications du chemin d'écho.

NCC est le plus souvent implémenté en remarquant que (II.8) (II.9) peuvent être simplifiés en supposant que $h \approx \hat{h}$, et par l'incorporation d'expression $\mathbf{R}_{xx}^{-1} \mathbf{r}_{xy} = \mathbf{h}$, pour donner une variante efficace en calcul de la variable de décision NCC

$$\xi_{NCC} = \frac{\sqrt{\mathbf{r}_{xy}^T \mathbf{h}}}{\sigma_y} \approx \frac{\sqrt{\mathbf{r}_{xy}^T \hat{\mathbf{h}}}}{\sigma_y} \quad (\text{II.10})$$

Le DTD basé sur cette méthode est défini comme :

$$DTD = \begin{cases} \xi_{NCC}(n) \geq T_{NCC}, & DTD = 0 \\ \xi_{NCC}(n) < T_{NCC}, & DTD = 1 \end{cases}$$

La situation DT est déclarée si $\xi_{NCC}(n) < T_{NCC}$, où T_{NCC} est une valeur de seuil de cette méthode.

En général, il a été démontré que la DTD-NCC est plus performante par rapport à la technique de corrélation croisée.

II.4.4. DTD basé sur Fast NCC (FNCC)

Les quantités estimées de corrélation croisée et la puissance du signal proche sont introduites pour la dérivation du DTD FNCC, L'équation de NCC est donnée par :

$$\xi = \sqrt{\mathbf{r}_{xd}^T (\sigma_d^2 \mathbf{R}_{xx})^{-1} \mathbf{r}_{xd}} \quad (\text{II.11})$$

L'équation de NCC peut être réécrite comme suit :

$$\xi_{FNCC}(n) = \xi^2(n) = \frac{\mathbf{r}^T(n) \mathbf{R}^{-1}(n) \mathbf{r}(n)}{\sigma_d^2(n)} = \frac{\eta^2(n)}{\sigma_d^2(n)} \quad (\text{II.12})$$

où nous avons appliqué le carré pour la variable de décision pour plus de simplicité, la variable de corrélation est estimée comme suite :

$$\mathbf{r}(n) = \lambda \mathbf{r}(n-1) + \mathbf{X}(n)d(n) \quad (\text{II.13})$$

$$\mathbf{R}(n) = \lambda \mathbf{R}(n-1) + \mathbf{X}(n)\mathbf{X}^T(n) \quad (\text{II.14})$$

$$\eta^2(n) = \lambda \eta^2(n-1) + d^2(n) - \varphi(n)e^2(n) \quad (\text{II.15})$$

$$\sigma_d^2(n) = \lambda \sigma_d^2(n-1) + d^2(n) \quad (\text{II.16})$$

On doit distinguer entre l'estimation du chemin d'écho du « background » calculé dans le DTD FNCC; on le note désormais $\mathbf{h}_b(n)$, et l'estimation calculée dans l'annuleur d'écho, $\mathbf{h}(n)$, regardons maintenant la statistique $\eta^2(n)$:

$$\begin{aligned}
\eta^2(n) &= [\lambda \mathbf{r}^T(n-1) + d(n)\mathbf{X}^T(n)]\mathbf{R}^{-1}(n) \times [\lambda \mathbf{r}(n-1) + \mathbf{X}(n)d(n)] \\
&= \lambda^2 \mathbf{r}^T(n-1)\mathbf{R}^{-1}(n)\mathbf{r}(n-1) + \lambda \mathbf{r}^T(n-1)\mathbf{R}^{-1}(n)\mathbf{X}(n)d(n) \\
&\quad + \lambda d(n)\mathbf{X}^T(n)\mathbf{R}^{-1}(n)\mathbf{r}(n-1) + d^2(n)\mathbf{X}^T(n)\mathbf{R}^{-1}(n)\mathbf{X}(n) \\
&= \lambda^2 \mathbf{r}^T(n-1)[\lambda^{-1}\mathbf{R}^{-1}(n-1) - \lambda^{-1}\alpha(n)\mathbf{k}(n)\mathbf{k}^T(n)]\mathbf{r}(n-1) \\
&\quad + 2\lambda d(n)\mathbf{k}^T(n)\mathbf{r}(n-1) + d^2(n)\mathbf{X}^T(n)\mathbf{k}(n) \\
&= \lambda \mathbf{r}^T(n-1)\mathbf{R}^{-1}(n-1)\mathbf{r}(n-1) - \lambda \alpha(n)[\mathbf{k}^T(n)\mathbf{r}(n-1)]^2 \\
&\quad + 2\lambda d(n)\mathbf{k}^T(n)\mathbf{r}(n-1) + d^2(n)\left(1 - \frac{\lambda}{\alpha(n)}\right) \\
&= \lambda \eta^2(n-1) - \frac{\lambda}{\alpha(n)} \hat{d}^2(n) + 2 \frac{\lambda}{\alpha(n)} d(n) \hat{d}(n) \\
&\quad + \left(1 - \frac{\lambda}{\alpha(n)}\right) d^2(n) \\
&= \lambda \eta^2(n-1) - \varphi(n) \hat{d}^2(n) + 2\varphi(n) \hat{d}(n) d(n) \\
&\quad + [1 - \varphi(n)] d^2(n) \\
&= \lambda \eta^2(n-1) + d^2(n) - \varphi(n) e^2(n) \tag{II.17}
\end{aligned}$$

Où la variable de vraisemblance $\varphi(n) = \lambda/\alpha(n)$ et $e(n)$ est l'erreur résiduelle,

$e(n) = d(n) - \hat{y}(n)$ on constate donc que les quantités nécessaires à partir de la statistique de test de la DTD de la FNCC sont données par l'équation récursive du premier ordre (II.14).

La DTD par cette méthode est défini comme :

$$DTD = \begin{cases} \xi_{FNCC}(n) \geq T_{FNCC}, & DTD = 0 \\ \xi_{FNCC}(n) < T_{FNCC}, & DTD = 1 \end{cases}$$

La situation DT est déclarée si $\xi_{SFE}(n) < T_{SFE}$, où T_{SFE} est une valeur de seuil de cette méthode.

II.4.5. DTD basé sur extraction de fonction vocale (SFE)

La méthode DTD basée sur l'extraction des caractéristiques de parole (SFE : *Speech Feature Extraction*) est présentée dans [23]. Cette méthode utilise une trame f de K historique des observations récentes des signaux normalisés du far-end et du microphone, respectivement, où $K \leq L$. Ces trames sont définies comme suit :

$$\mathbf{X}_K(n) = [x(n)x(n-1), \dots, x(n-K+1)]^T \quad (\text{II.18})$$

$$\mathbf{d}_K(n) = [d(n)d(n-1), \dots, d(n-K+1)]^T \quad (\text{II.19})$$

Un vecteur caractéristique $\mathbf{V}_{ch}(n)$ est calculé pour chaque trame et qui contient quatre caractéristiques de la parole :

- L'énergie $\mathbf{E}_n(\cdot)$
- L'écart type $\mathbf{std}(\cdot)$
- La valeur maximale $\mathbf{max}(\cdot)$
- L'énergie logarithmique $\mathbf{E}_s(\cdot)$

Est défini comme suit :

$$\mathbf{V}_{ch}(n) = [\alpha_1 E_n(f) \quad \alpha_2 std(f) \quad \alpha_3 \max(f) \quad (\alpha_4 + E_s(f))/\alpha_4] \quad (\text{II.20})$$

Où $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ sont des paramètres de normalisation positifs, $E_n(f)$ et $E_s(f)$ sont respectivement, l'énergie et l'énergie logarithmique de la $n^{\text{ième}}$ trame f .

Pour détecter les périodes de double parole, une variable de décision $\xi_{SFE}(n)$ indique (décision d'extraction de fonction vocale) est calculée pour chaque itération. Elle est définie par la distance euclidienne entre les vecteurs caractéristiques $\mathbf{V}_{chx}(n)$ et $\mathbf{V}_{chd}(n)$ des trames (far-end) et microphone, respectivement :

$$\xi_{SFE}(n) = 1 - [\|\mathbf{V}_{chx}(n) - \mathbf{V}_{chd}(n)\| - \beta \hat{\sigma}_b^2] \quad (\text{II.21})$$

Où $\|\mathbf{V}_{chx}(n) - \mathbf{V}_{chd}(n)\|$ désigne la distance euclidienne entre les deux vecteurs $\mathbf{V}_{chx}(n)$ et $\mathbf{V}_{chd}(n)$,

β est une constante positive.

$\hat{\sigma}_b^2$ est la valeur de la variance estimée du bruit de fond.

Ensuite, un seuil dynamique est utilisé en fonction de l'activité vocale du signal near-end (VAD *Voice Activity Detection*), le module de VAD basé sur l'énergie du near-end est développé comme suit :

$$VAD = \begin{cases} 1, & \text{si } E_n(n) \geq T_{VAD} \\ 0, & \text{si } E_n(n) < T_{VAD} \end{cases}$$

Où T_{VAD} est la valeur de seuil de VAD.

La valeur du seuil T_{SFE} pour la décision DT est définie par :

$$T_{SFE} = \begin{cases} T_{SFE1}, & \text{si } VAD = 1 \\ T_{SFE2}, & \text{si } VAD = 0 \end{cases}$$

Ici T_{SFE1} et T_{SFE2} sont des constantes de seuil.

La DTD par cette méthode est défini comme suit :

$$DTD = \begin{cases} \xi_{SFE}(n) \geq T_{SFE}, & DTD = 0 \\ \xi_{SFE}(n) < T_{SFE}, & DTD = 1 \end{cases}$$

La situation DT est déclarée si $\xi_{SFE}(n) < T_{SFE}$, où T_{SFE} est une valeur de seuil de cette méthode.

II.4.6. DTD basé sur la décomposition en valeur singulière (SVD)

Une approche de DTD utilisant la décomposition en valeur singulière SVD (*Singular Value Decomposition*) du signal lointain est proposée dans [24], dans laquelle la SVD est calculée pour chaque matrice de remodelage $\mathbf{X}_{R \times R}(n)$ de $\mathbf{X}_M(n)$, l'historique de longueur-M de $x(n)$, où $R \times R = M$ et $M < L$. La variable de décision de cette méthode est exprimée comme suit :

$$\xi_{SVDX}(n) = \frac{\varphi}{|\hat{r}_{de}|} \quad (\text{II.22})$$

Où $r_{de} = E\{d(n)e(n)\}$, $E\{\cdot\}$ représente une espérance mathématique. Cette décision statistique est mise en œuvre à l'aide de l'estimation de la corrélation croisée \hat{r}_{de} déterminée par l'équation récurrente suivante :

$$\hat{r}_{de}(n) = \lambda \hat{r}_{de}(n-1) + (1-\lambda)e(n)d(n) \quad (\text{II.23})$$

Où λ est le facteur de pondération exponentiel ($0,9 < \lambda < 1$).

La variable φ est définie par : $\varphi = \max\{\text{diag}\Sigma_X\}$ (II.24)

Où Σ_X est la matrice diagonale résultant de la SVD appliquée à une matrice de remodelage $\mathbf{X}_{R \times R}(n)$ du vecteur $\mathbf{X}_M(n)$, La valeur seuil T_{SVD} est comparée à ξ_{SVDX} pour la déclaration des périodes DT.

La DTD par cette méthode est défini comme suit :

$$DTD = \begin{cases} \xi_{SVDX}(n) \geq T_{SVD}, & DTD = 0 \\ \xi_{SVDX}(n) < T_{SVD}, & DTD = 1 \end{cases}$$

La situation DT est déclarée si $\xi_{SVDX}(n) < T_{SVD}$, où T_{SVD} est une valeur de seuil de cette méthode.

II.4.7. DTD basé sur filtre adaptatif auxiliaire (AAF)

Une nouvelle méthode de DTD utilisant un filtre adaptatif auxiliaire (AAF : *Auxiliary Adaptive Filter*) [25]. L'idée principale de cette méthode est d'améliorer la capacité de discrimination entre le signal far-end $x(n)$ et le signal du microphone $d(n)$. Une valeur maximale de chaque vecteur d'historique de longueur-M est utilisée pour contrôler les mises à jour des deux filtres adaptatifs. Le filtre adaptatif auxiliaire est contrôlé par la première décision, où une variable de décision statistique $\xi_A(n)$ est calculée comme suit :

$$\xi_A(n) = 1 - |N_{m\hat{y}}(n) - N_{mx}(n)| \quad (\text{II.25})$$

Où $N_{m\hat{y}}(n)$ et $N_{mx}(n)$ sont les valeurs max du vecteur d'écho estimé $\hat{y}_m(n)$ et du vecteur du far-end $x_M(n)$, respectivement.

La mise à jour du filtre adaptatif auxiliaire est arrêtée si $\xi_A(n) < T_A$,

Où $T_A < 1$ est une valeur seuil.

Le filtre adaptatif fondamental est contrôlé par la deuxième décision, où une variable décision statistique $\xi_{AAF}(n)$ est calculée comme suit :

$$\xi_{AAF}(n) = 1 - |N_{md}(n) - N_{m\hat{y}}(n)| \quad (\text{II.26})$$

Où $N_{md}(n)$ est le maximum du signal du microphone.

$$DTD = \begin{cases} \xi_{AAF}(n) \geq T_{AAF}, & DTD = 0 \\ \xi_{AAF}(n) < T_{AAF}, & DTD = 1 \end{cases}$$

La situation DT est déclarée si $\xi_{AAF}(n) < T_{AAF}$, où T_{AAF} est une valeur de seuil de cette méthode.

II.4.8. La DTD basé sur l'algorithme de réponse d'impulsion de variance (VIRE)

Une autre méthode de DTD utilisant l'algorithme de variance de réponse impulsionnelle (VIRE : *Variance Impulse Response Algorithm*) [26] et qui calcule la variance de la valeur maximale des coefficients récents dans le filtre adaptatif.

Si la variance dépasse un certain seuil, qui peut varier dans le temps, nous avons une DT. En d'autres termes, si la réponse impulsionnelle estimée de la pièce change beaucoup, on

suppose que ce n'est pas la pièce qui a changé, mais qu'une autre source de son est apparue. La formule mathématique est un peu compliquée et elle s'écrit comme suit :

$$\xi_{VIRE}(n) = \lambda \cdot \xi(n-1) + (1-\lambda) \cdot [\gamma - \bar{\gamma}]^2 \quad (\text{II.27})$$

$$\sigma_{\bar{\gamma}}^2(t) = \lambda \sigma_{\bar{\gamma}}^2(t-1) + (1-\lambda)(\gamma - \bar{\gamma}(t))^2 \quad (\text{II.28})$$

$$\bar{\gamma}(t) = \lambda \bar{\gamma}(t-1) + (1-\lambda)\gamma \quad (\text{II.29})$$

$$\gamma = \max\{\hat{h}_1, \dots, \hat{h}_n\} \quad (\text{II.30})$$

Avec $\bar{\gamma}(t)$ la valeur moyenne de λ et un facteur d'oubli, le calcul est simple car il ne nécessite que cinq multiplications.

La DTD par cette méthode est défini comme suit :

$$DTD = \begin{cases} \xi_{VIRE}(n) \geq T_{VIRE}, & DTD = 0 \\ \xi_{VIRE}(n) < T_{VIRE}, & DTD = 1 \end{cases}$$

La situation DT est déclarée si $\xi_{VIRE}(n) < T_{VIRE}$, où T_{VIRE} est une valeur de seuil de cette méthode.

II.5. Conclusion

Dans ce chapitre, nous avons présenté le concept de détection de double parole notamment les algorithmes basés sur la corrélation croisée entre les différents signaux disponibles aux entrées du filtre adaptatif, l'objectif principal de DTD est d'interrompre la mise à jour des coefficients de filtre pendant les périodes de double parole. Les détecteurs de double parole peuvent être classés en deux catégories : ceux basés sur l'énergie (algorithme de Geigel) et ceux basés sur la corrélation croisée. Les algorithmes basés sur l'énergie ont l'avantage sur les algorithmes basés sur la corrélation croisée dans la complexité de calcul donc ils nécessitent très peu de mémoire et qui ont été utilisés avec succès dans l'annulation d'écho électrique, cependant, ils ne fournissent pas toujours des performances fiables dans un environnement d'écho acoustique, d'autre part, les algorithmes basés sur la corrélation montrent des performances de détection améliorées dans un tel environnement variant dans le temps.

Chapitre III

Simulation et Résultats

Chapitre III : Simulation et résultats

III.1. Introduction

Dans les chapitres précédents, nous avons présenté les concepts généraux ainsi que les développements mathématiques relatifs au domaine de l'annulation d'écho acoustique, en particulier les détecteurs de double parole. Ce dernier chapitre représente le fruit de cette étude théorique dans lequel nous allons présenter les résultats de simulation, pour se faire nous allons utiliser des signaux réels afin de mettre en œuvre dans un premier temps l'algorithme de filtrage adaptatif et les deux DTD choisis suivie d'une évaluation des performances dans différentes situations et on finira par une comparaison des deux DTDs et qui est basée sur des critères standards.

III.2. Présentation de logiciel de simulation

MATLAB est un environnement informatique numérique particulièrement efficace pour calculer et simuler les problèmes techniques. Ce langage de programmation très puissant permet la manipulation matricielle, le traçage de fonctions et de données, la mise en œuvre d'algorithmes, la création d'interfaces utilisateur et l'interfaçage avec d'autres langages de programmation (C, C ++, Fortran et Java). L'une des caractéristiques les plus bénéfiques est la visualisation graphique, qui nous aide à avoir confiance en nos résultats en surveillant et en analysant les graphiques résultants. De plus, MATLAB implémente Simulink, les modèles de paquets, simule, et analyse des systèmes dynamiques. Cela nous permet de poser une question sur un système, de le modéliser et de voir ce qui se passe.

Pour notre objectif de simulation, MATLAB est en réalité un logiciel nécessaire et efficace pour obtenir des résultats convaincants pour les raisons suivantes :

- Enregistrement facile des signaux audio des locuteurs lointains et proches. Ces données sont indispensables à la simulation.
- Le calcul matriciel est très important car les données ont été traitées sous forme de formats matriciels.
- Facile à surveiller les résultats en traçant les graphiques désirés. Nous devons en particulier entendre les sons résultants, par MATLAB, c'est simple à réaliser.
- La structure des commandes convient au calcul avec le traitement du signal.

III.3. Structure générale du programme

III.3.1. Organigramme du system AEC

L'organigramme de l'algorithme AEC de la figure III.1 nous montre la façon avec laquelle le programme se déroule, et les différentes étapes suivies dans l'algorithme d'adaptation du filtre et dans les DTDs.

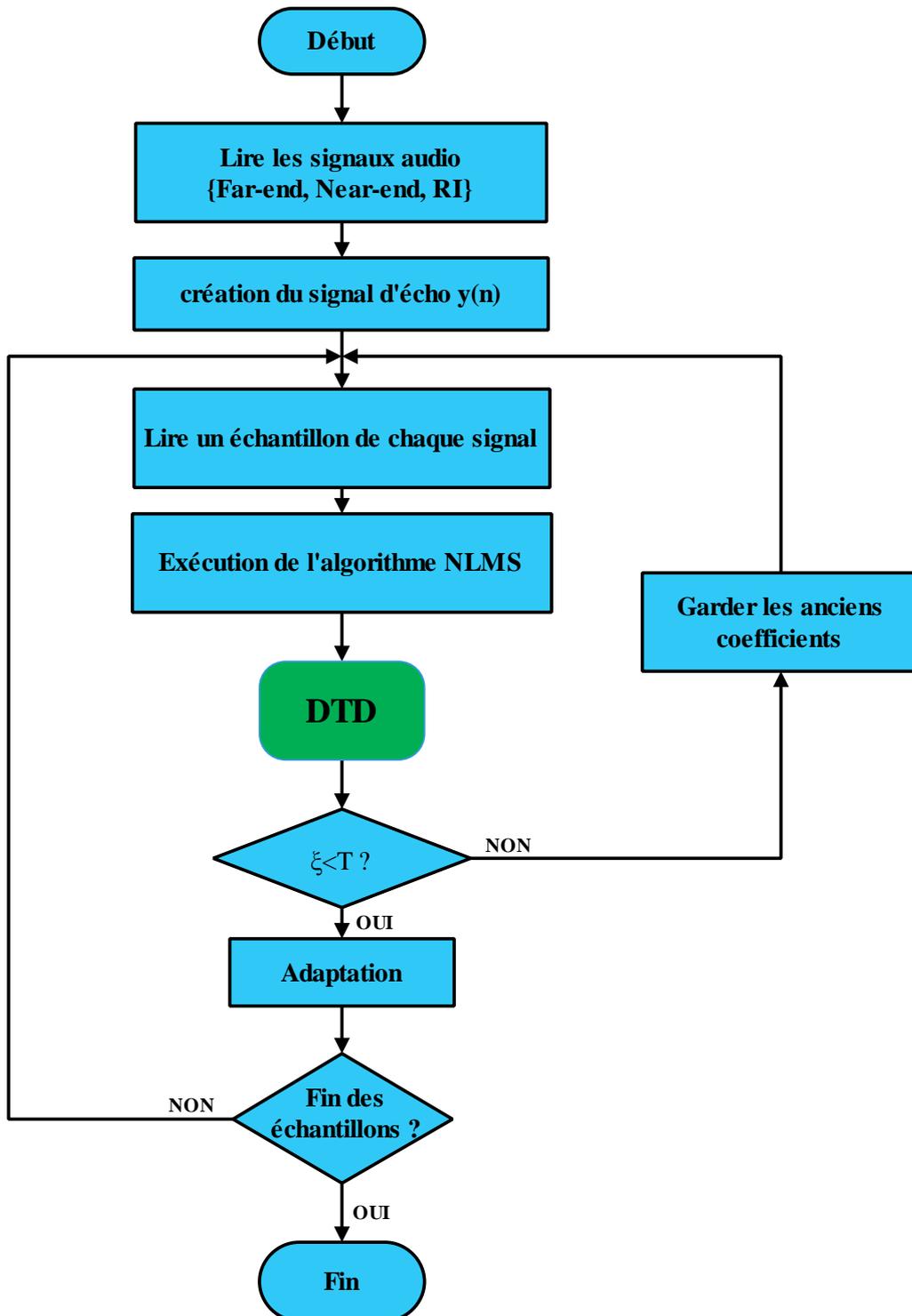


Figure III. 1. Organigramme de simulation.

III.3.2. Code Matlab

D'après l'organigramme précédent, le programme se divise en deux parties essentielles : la première sur l'algorithme de filtrage adaptatif et la deuxième sur le DTD.

III.3.2.1. Algorithme de filtres adaptatif

Dans cette partie, nous allons présenter sous formes de tableaux le calcul matriciel issu du développement mathématique déjà expliqué dans le premier chapitre ainsi que le code MATLAB correspondant, le choix de l'algorithme NLMS s'est basé sur son compromis entre ses bonnes performances et sa complexité de calcul relativement faible.

Tableau .III.1. Résumé de l'algorithme NLMS

NLMS algorithme	
Conditions initiales :	$0 < \beta < 2$ β : est la taille de pas normalisée C : petit constant Vecteur d'entrée : $X_{L,1} = [0,0 \dots, 0]^T$ Vecteur de poids : $w_{L,1} = [0,0 \dots, 0]^T$
Pour chaque instant, $n = 1, 2, \dots$, calculer :	
Signal de la sortie	$\hat{y}(n) = \mathbf{w}^T(n) \cdot \mathbf{X}_L(n)$
Signal d'erreur :	$e(n) = d(n) - \hat{y}(n)$
Mise à jour du filtre :	$\mu(n) = \frac{\beta}{\mathbf{X}^H(n) \mathbf{X}(n)}$ $\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu_{NLMS}}{\mathbf{X}_L^T(n) \mathbf{X}_L(n) + \delta_{NLMS}} \mathbf{X}_L(n) e(n)$

Tableau .III.2. Code MATLAB de l'algorithme NLMS

Code MATLAB	
Conditions initiales :	<code>beta=0.7 ; %béta</code> <code>c=0.0001 ; %Une petite constante</code> <code>N=length(x) ; %Nombre d'itérations</code> <code>L=length(h) ; %Longueur du filtre adaptatif</code> <code>w=zeros (L, 1) ; %Vecteur du poids initial</code> <code>xin=zeros (L, 1) ; %Signal d'entrée initial</code>
Pour chaque instant, $T = 1, 2, \dots$, calculer :	
Signal de la sortie :	<code>y(k)=w'*xin ;</code>
Signal d'erreur :	<code>e(k)=d(k)-w'*xin ;</code>
Mise à jour du filtre :	<code>mu=beta/(c+xin'*xin) ;</code> <code>wtemps=w+mu*e(k)*xin ;</code>

III.3.2.2. Algorithmes de Détection de double parole

Dans notre étude, nous avons utilisé deux méthodes de détection de double parole qui sont déjà détaillées dans le deuxième chapitre, il s'agit de l'algorithme Geigel et la corrélation croisée normalisée (NCC). Rappelons que Geigel est basé sur une comparaison de niveau entre le signal du microphone $d(n)$ et le signal lointain $x(n)$ tout en supposant que le signal de parole proche $v(n)$ dans le signal du microphone est plus fort que le signal d'écho $y(n)$.

En revanche dans l'autre méthode (NCC) on calcul les estimations à l'aide de l'algorithme de pondération exponentiel récursive pour obtenir les valeurs de la corrélation croisée $r_{ed}(n)$ entre le signal d'erreur et le signal du microphone et la variance $\hat{\sigma}_d^2(n)$ du signal du microphone et calculer par conséquent la variable statistique de décision (DTD) à partir de ces valeurs. Enfin, le résultat est comparé au seuil (T) pour prendre la décision du détecteur de double parole.

Le résumé et le code MATLAB ci-dessous illustrent la théorie et la simulation du problème de détection de double parole.

Tableau .III.3. Résumé de l'algorithme de la DT- Geigel

L'algorithme de Geigel	
Conditions initiales :	Seuil ≈ 1
Pour chaque instant, $n = 1, 2, \dots$, calculer :	
Statistique de décision	$\xi_G = \frac{\max\{ x(n) , \dots, x(n - L_g + 1) \}}{ d(n) }$
Décision de la DTD	Si $\xi_G < T$ DTD Si $\xi_G > T$ NO-DTD

Tableau .III.4. Résumé de l'algorithme de la DT- NCC

L'algorithme NCC	
Conditions initiales :	$0 < \lambda < 1$ et $\lambda \approx 1$ λ : le facteur de pondération exponentiel Seuil ≈ 1
Pour chaque instant, $n = 1, 2, \dots$, calculer :	
Cross-corrélation	$r_{ed}(n) = \lambda r_{ed}(n - 1) + (1 - \lambda)e(n)d^T(n)$
La variance de la parole proche	$\hat{\sigma}_d^2(n) = \lambda \hat{\sigma}_d^2(n - 1) + (1 - \lambda)d(n) d^T(n)$
Statistique de décision	$\xi_{NCC} = 1 - \frac{\hat{r}_{ed}}{\hat{\sigma}_d^2}$
Décision de la DTD	Si $\xi_{NCC} < T$ DTD Si $\xi_{NCC} > T$ NO-DTD

Dans le code MATLAB, on introduit la nouvelle variable *wtemp* (valeur temporaire) permettant de mettre à jour les coefficients du filtre adaptatif.

- Si $\xi < T$, le filtre adaptatif doit être gelé, c'est-à-dire arrêter de mettre à jour les coefficients du filtre adaptatif. La parole proche est détectée (mode double parôle).
- Si $\xi > T$, continuez alors à exécuter la boucle de l'algorithme de filtre adaptatif, c'est-à-dire que la mise à jour des coefficients du filtre adaptatif doit continuer. Pas de parole proche détectée : parole lointaine uniquement.

Tableau .III.5. Code MATLAB de l'algorithme de la DTD- Geigel

Code MATLAB de Geigel	
Conditions initiales :	T=0.4 ; % Seuil M=100 ;
Pour chaque instant, n = 1, 2, ..., calculer :	
Statistique de décision :	ksi(k) = abs(max(x(N)))/abs (d(k));
Décision de la DTD :	if k>M DTD(k)=1; if ksi(k)>T w=w+mu*e(k)*xin; DTD(k)=0; else w=w ; DTD(k)=1; End

Tableau .III.6. Code MATLAB de l'algorithme de la DTD- NCC

Code MATLAB de NCC	
Conditions initiales :	T=0.7 ; % Seuil Lambda_DTD=0.95 ; % Constant DTD_start=50*L ; % Le temps d'activer la DTD
Pour chaque instant, n = 1, 2, ..., calculer :	
Cross-corrélation	r_ed(k)=lambda_DTD*(r_ed(k-1))+(1-lambda_DTD)*e(k)*d(k)';
La variance de la parole proche	varMIC(k)=sqrt (lambda_DTD*(varMIC (k-1)^2) + (1-lambda_DTD)*d(k)*d(k)') ;
Statistique de décision	deci_stat(k)=1-(r_ed(k)/varMIC(k)^2);
Décision de la DTD	if (deci_stat(k)> Seuil (k)) w=wtemps ; end

III.4. Résultat de simulation

Dans cette partie nous allons présenter les résultats obtenus et qui concernent les performances d'annuleur d'écho acoustique avec l'algorithme NLMS en utilisant des DTDs.

Nous allons d'abord représenter les différents signaux utilisés dans les simulations :

- Le signal lointain $x(n)$ et le signal local $v(n)$ qui sont deux enregistrements vocaux composés de 36000 échantillons avec une fréquence d'échantillonnage de 8 KHz,
- Le signal d'écho $y(n)$ est obtenu par couplage du signal lointain avec la réponse impulsionnelle choisie.
- Le signal du microphone $d(n)$ qui est obtenu à partir de la somme de $y(n)$ et $v(n)$.
- La réponse impulsionnelle utilisée dans toutes les simulations et qui contient 128 échantillons.

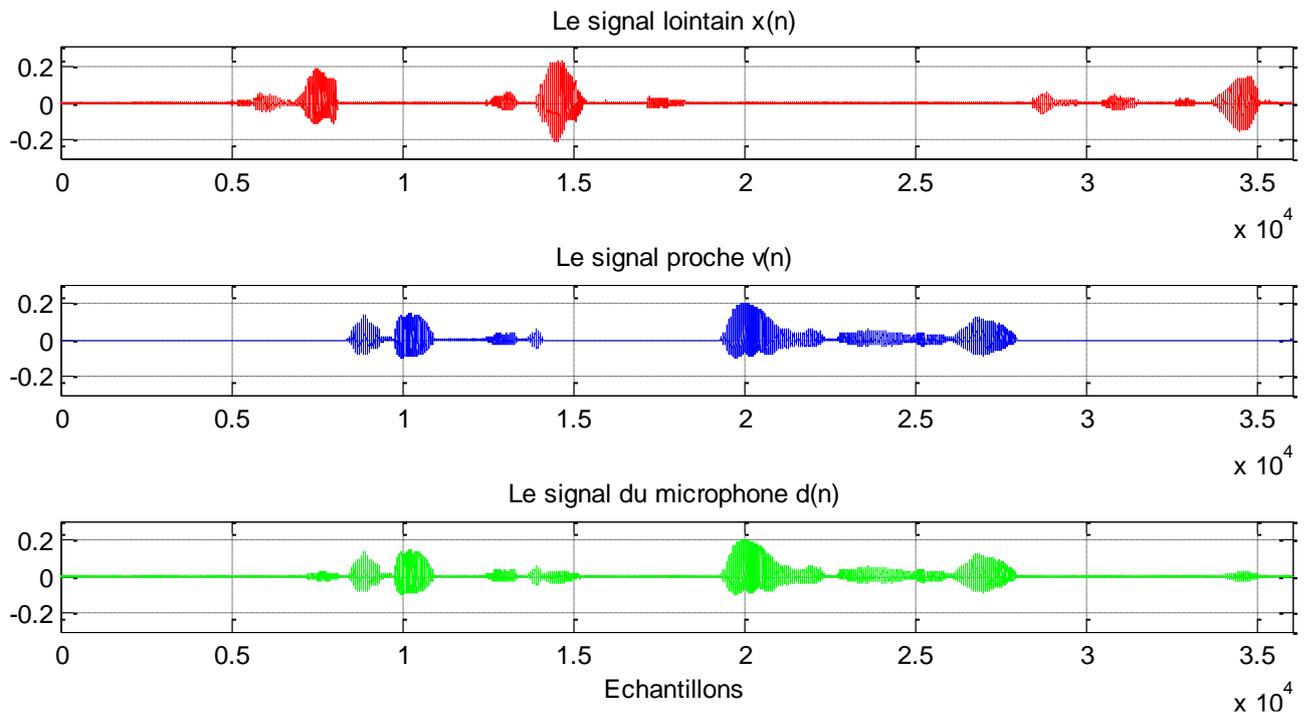


Figure III.2. Les signaux utilisés dans la simulation.

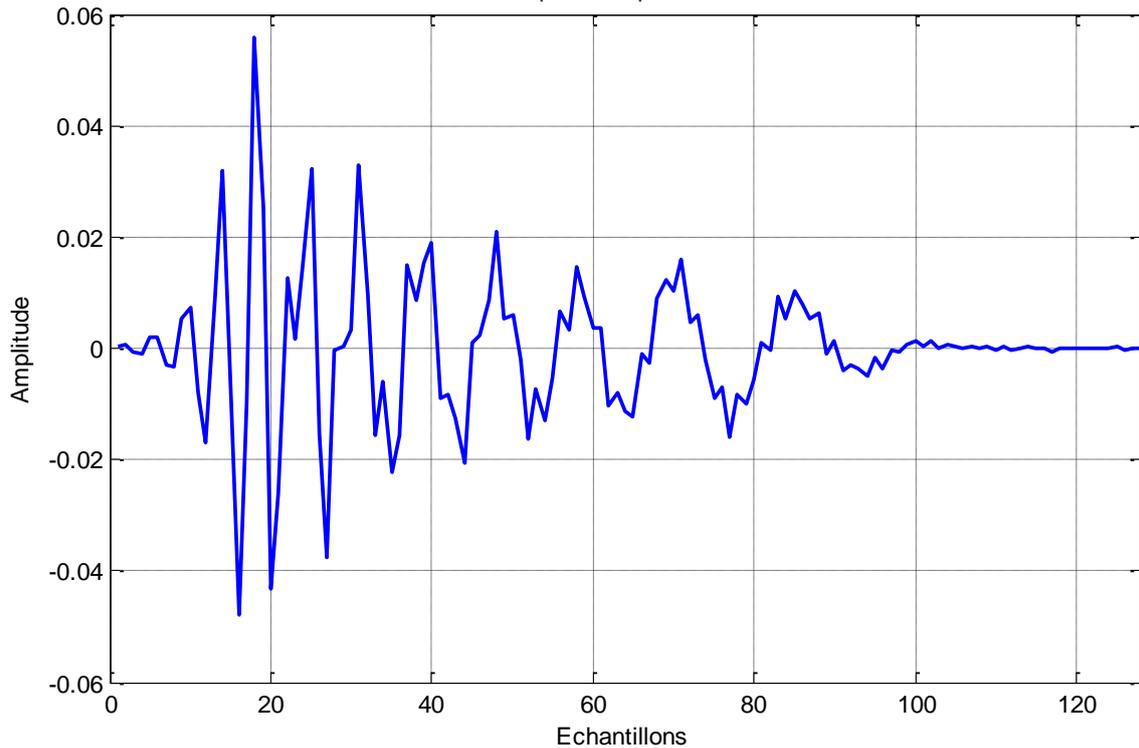


Figure III.3. La réponse impulsionnelle

Pour analyser les performances des algorithmes, nous nous intéressons dans le cadre de cette étude, à trois critères principaux à savoir, la convergence du filtre (Malignement), la réduction d'écho par la mesure (*Echo Return Loss Enhancement ERLE*) et enfin l'erreur quadratique moyenne (*Mean Square Error MSE*).

Malignement :

La convergence des filtres adaptatifs, mesuré en dB est défini comme suit :

$$Misalignment(dB) = 10 \log_{10} \left[\frac{\|h - \hat{w}\|^2}{\|h\|^2} \right] \quad (III.1)$$

h et \hat{w} désignent respectivement les réponses impulsionnelles, mesurées et estimées du chemin d'écho. L'étude de l'évolution de la convergence en cours du temps permet de nous renseigner sur la vitesse de convergence de l'algorithme d'adaptation.

ERLE :

L'ERLE est l'un des paramètres les plus importants est couramment utilisé pour évaluer les performances de l'algorithme d'annulation d'écho. Ce critère mesure l'atténuation de l'écho supprimée du signal de microphone, il est également mesuré en dB et défini comme le rapport entre la puissance du signal de microphone $d(n)$ et celle du signal de l'erreur résiduelle $e(n)$.

$$ERLE(dB) = 10 \log_{10} \frac{P_d(n)}{P_e(n)} = \frac{E[d^2(n)]}{E[e^2(n)]} \quad (III.2)$$

Erreur quadratique moyenne (MSE) :

L'erreur quadratique moyenne représente la puissance du signal d'erreur :

$$MSE(dB) = 10 \log_{10}\{E\{|e(n)|^2\}\} \quad (III.3)$$

$e(n)$ et $E[.]$ désignent respectivement le signal d'erreur et l'espérance mathématique

III.4.1. Algorithme d'annulation d'écho sans DTD

L'objectif de cette simulation est d'évaluer les performances de l'algorithme de filtrage adaptatif NLMS et de voir l'influence de la valeur du pas d'adaptation β . Dans un premier temps on suppose qu'il n'existe pas de parole locale ($v(n)=0$), nous n'avons donc pas besoin d'utiliser un DTD, la figure III.4 illustre les courbes de convergence obtenues pour différentes valeurs du pas d'adaptation. Nous pouvons constater que la vitesse de convergence est proportionnelle au pas d'adaptation β .

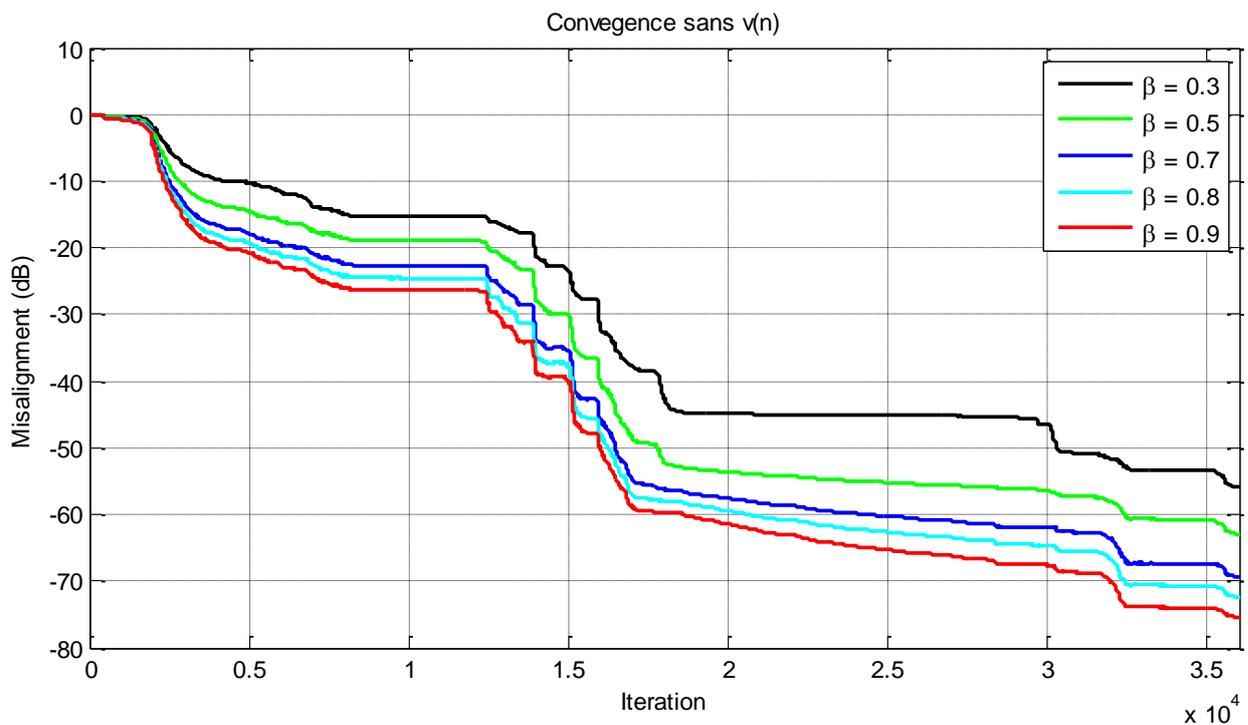


Figure III.4. L'influence du pas d'adaptation sur l'algorithme de NLMS en absence de signal local

Dans une deuxième simulation, nous avons refait presque le même travail mais avec la présence de la parole locale. L'objectif est de voir l'utilité du DTD dans un AEC et l'influence du pas d'adaptation en présence de $v(n)$.

Le résultat obtenu est présenté dans la figure III.5. Nous pouvons constater que plus le filtre est rapide plus la divergence est importante. Autrement dit, on peut choisir un pas

d'adaptation plus important pour avoir une grande vitesse de convergence mais au détriment de la robustesse à la présence de la parole local, il est important donc de choisir un pas d'adaptation assurant une bonne vitesse de convergence toute en gardant la robustesse en cas de présence de parole locale.

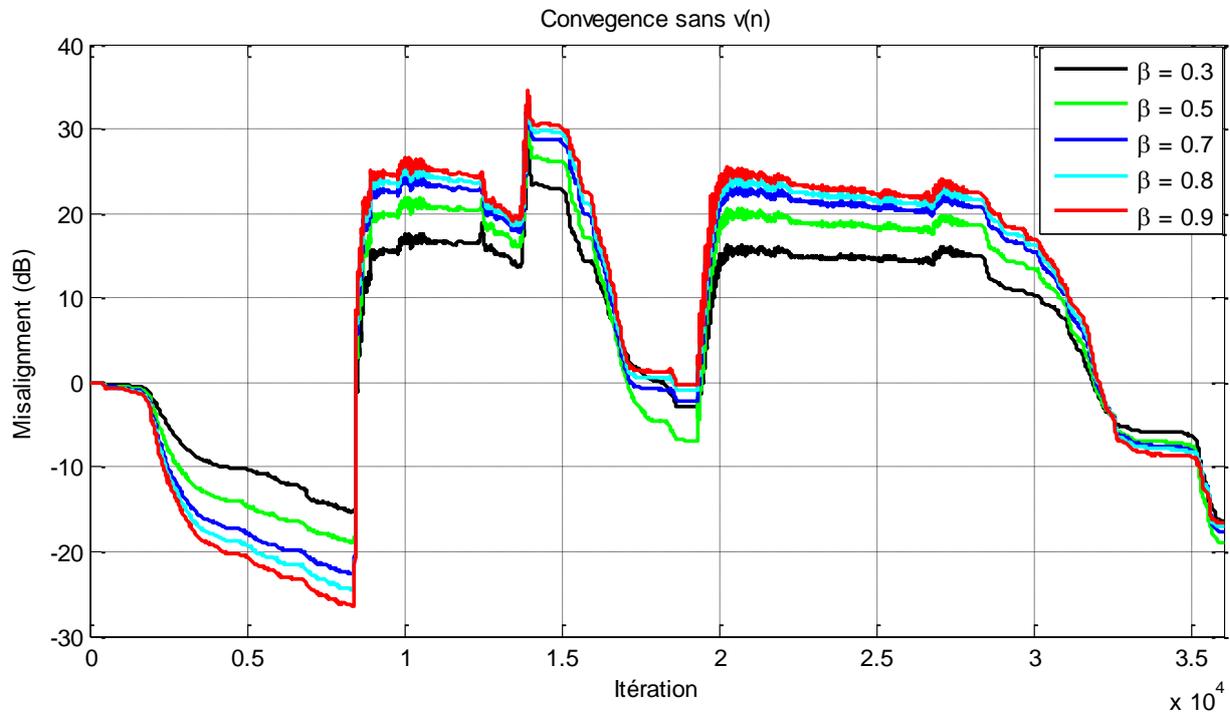


Figure III.5. Influence du pas d'adaptation sur l'algorithme NLMS en présence du signal local.

III.4.2. Algorithme d'annulation d'écho avec DTD

Pour voir les performances des deux DTDs choisis, nous avons mis en œuvre les deux algorithmes (Geigel, et NCC), dans un premier temps en supposant qu'il n'existe pas de bruit ($b(n)=0$), avec les paramètres utilisés sont comme suit :

- NLMS: $N=36000$, $L=128$, $\beta=0.7$, $c=0.001$.
- Geigel: $T_G=0.4$, $M=100$, $DTD_start=1000$.
- NCC: $T_{NCC}=0.15$, $\lambda_{DTD}=0.95$, $DTD_start=1000$.

Les Figures III.6, III.7 et III.8 illustrent respectivement l'évaluation de misalignement, d'ERLE et de la MSE pour un système d'annulation d'écho acoustique sans et avec DTD. Les courbes obtenues montrent une dégradation dans les performances de l'AEC sans le DTD dans les situations de DT. Cette dégradation est indiquée par la divergence du filtre et la décroissance des valeurs d'ERLE. Dans l'autre côté, une nette amélioration lorsque on utilise un DTD.

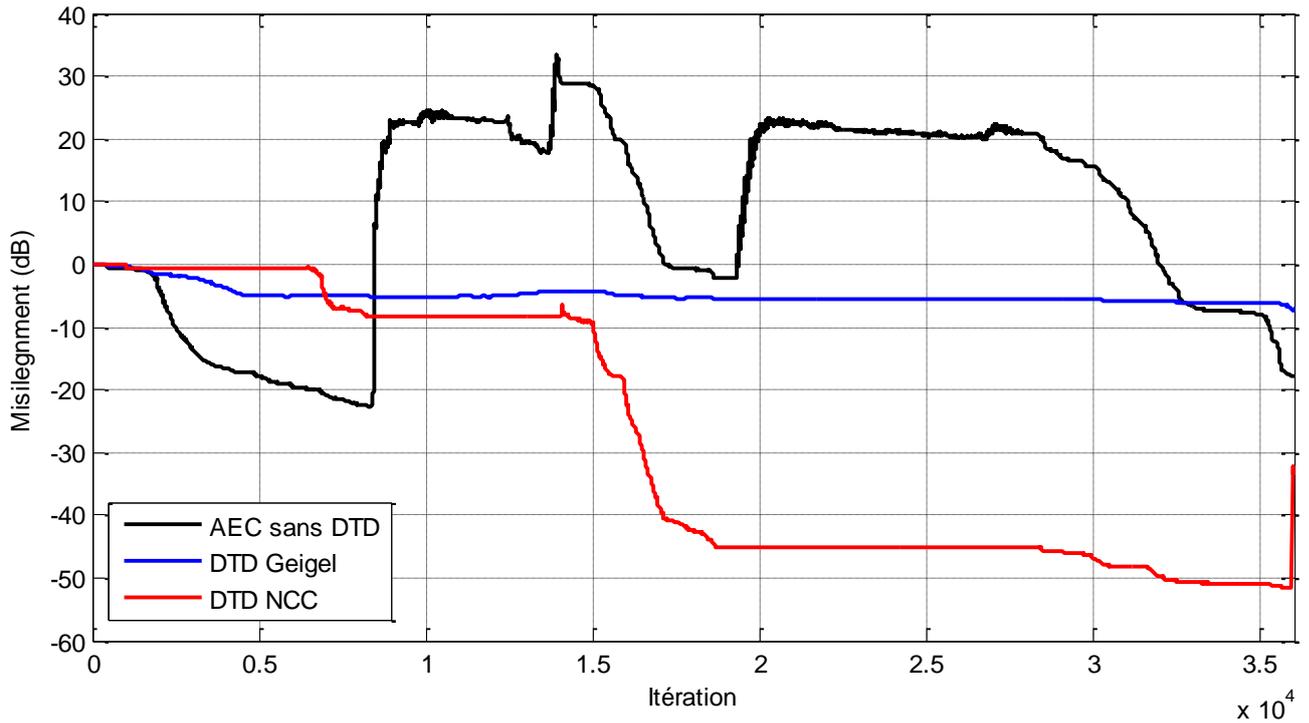


Figure III.6. Evaluation de misalignement pour les méthodes de Geigel et NCC sans bruit.

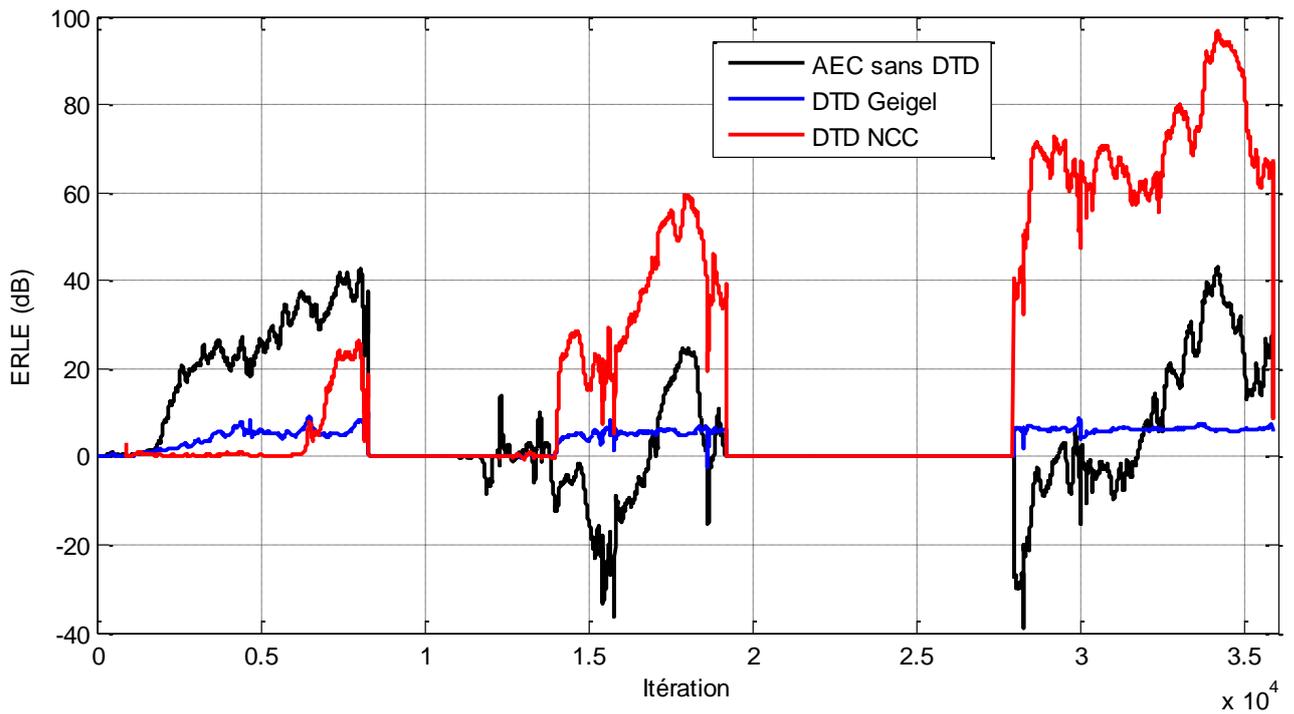


Figure III.7. L'évaluation de l'ERLE pour les méthodes de Geigel et NCC sans bruit.

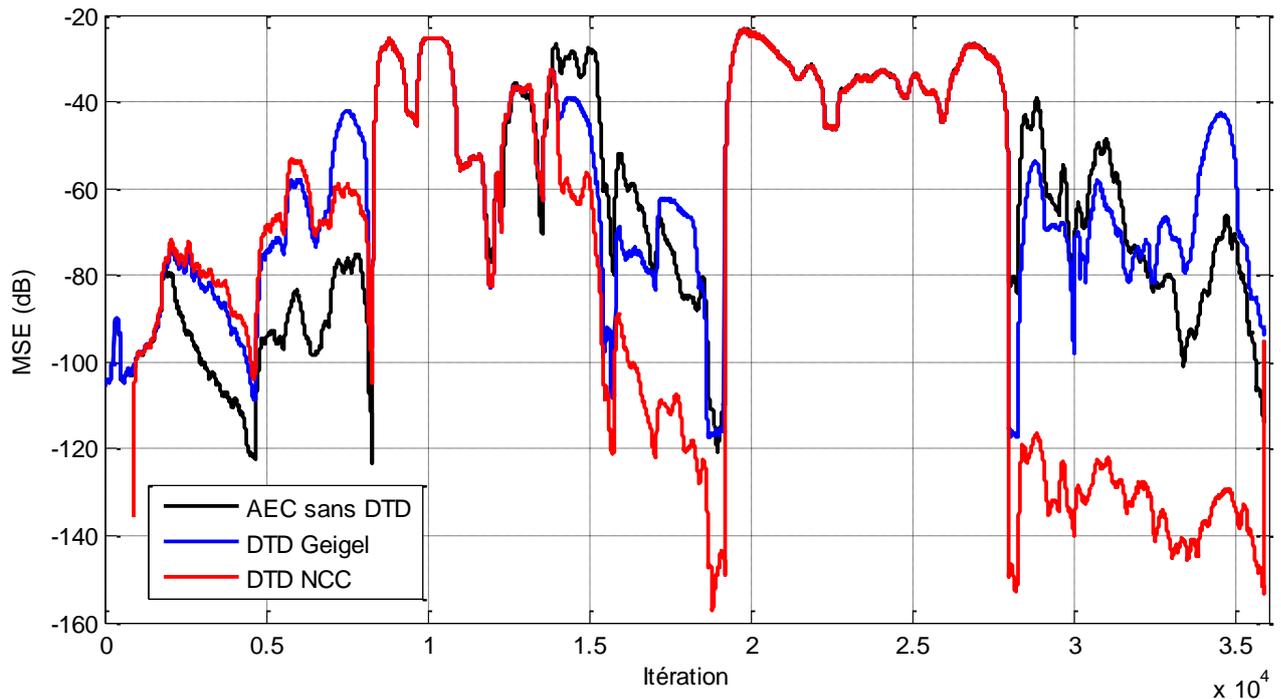


Figure III.8. L'évaluation de MSE pour les méthodes de Geigel et NCC sans bruit.

Ces résultats s'expliquent encore quand on représente les variables de décisions pour les deux DTD (Geigel et NCC) dans la figure III.9. Les variables de décision des deux DTDs indiquent l'activité vocale du locuteur proche dans le signal du microphone $d(n)$ qui contient un mélange d'écho $y(n)$ et du signal proche $v(n)$. D'après ces résultats on remarque que NCC est beaucoup plus performante que Geigel car sa variable de décision suit parfaitement la présence et l'absence du signal local alors que celle de Geigel fait beaucoup de fausses détections et une instabilité claire.

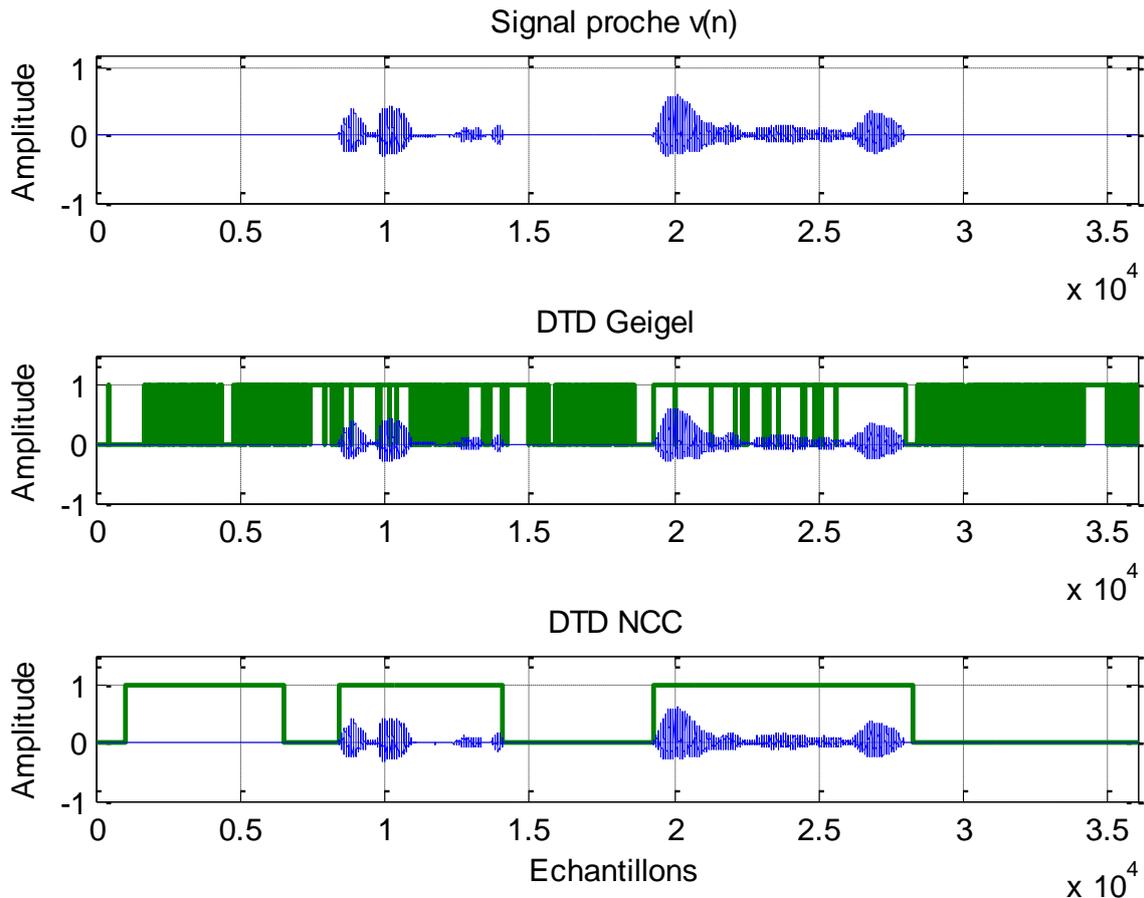


Figure III.9. Les signaux de DTD obtenus avec les méthodes de Geigel et NCC sans bruit.

III.4.3. Etude du bruit

Dans les simulations précédentes, nous avons supposé que la communication se déroule dans un environnement non-bruité, dans cette section l'environnement réel de communication est modélisé par un bruit blanc Gaussien $b(n)$ qui est ajouté au signal d'écho $y(n)$ avec différentes valeurs du rapport signal sur bruit (*Signal-to-Noise Ratio* SNR), où :

$$SNR(dB) = 10 \log_{10} \left\{ \frac{\left[E |y(n)|^2 \right]}{\left[E |b(n)|^2 \right]} \right\} \quad (III.4)$$

Les figures III.10 à III.15 représentent les résultats obtenus pour les trois critères d'évaluation utilisés des deux DTDs choisis. On remarque bien l'effet d'instabilité du bruit sur les performances de l'AEC, on voit également que le DTD NCC garde toujours sa suprématie par rapport à Geigel dans les niveaux faibles du bruit mais pour les niveaux élevés (SNR=10 par exemple) Geigel semble proche du NCC.

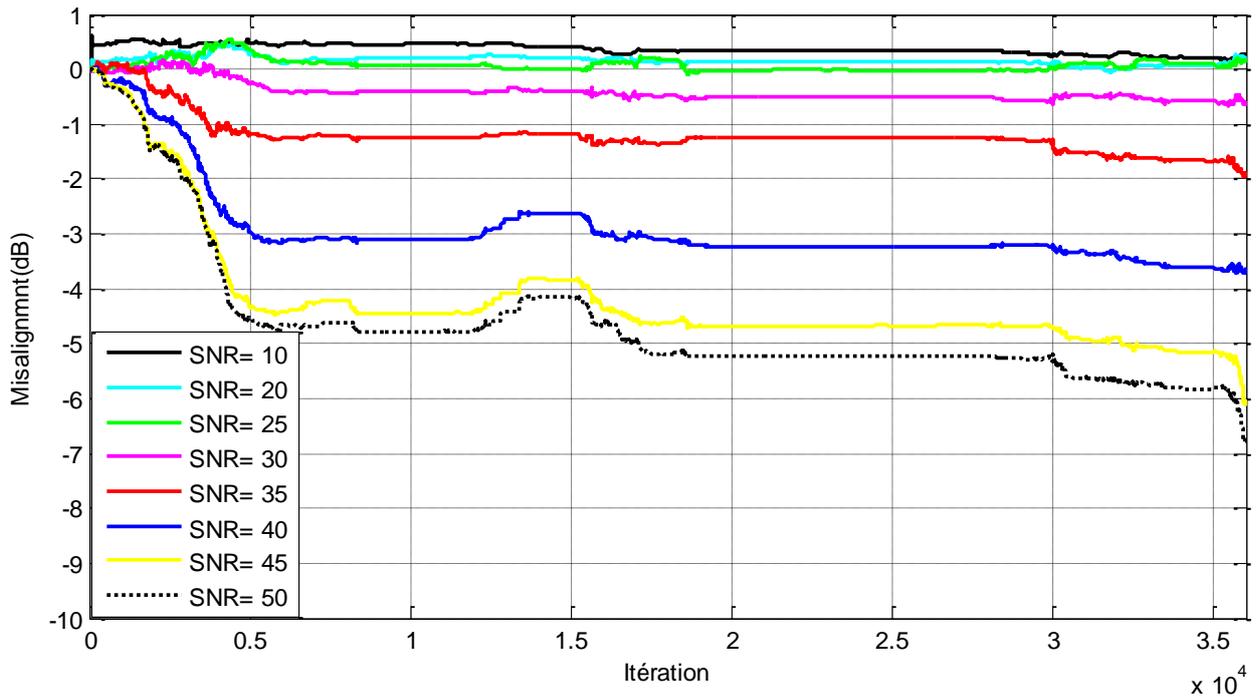


Figure III.10. Convergence AEC avec DTD-Geigel dans un environnement bruité.

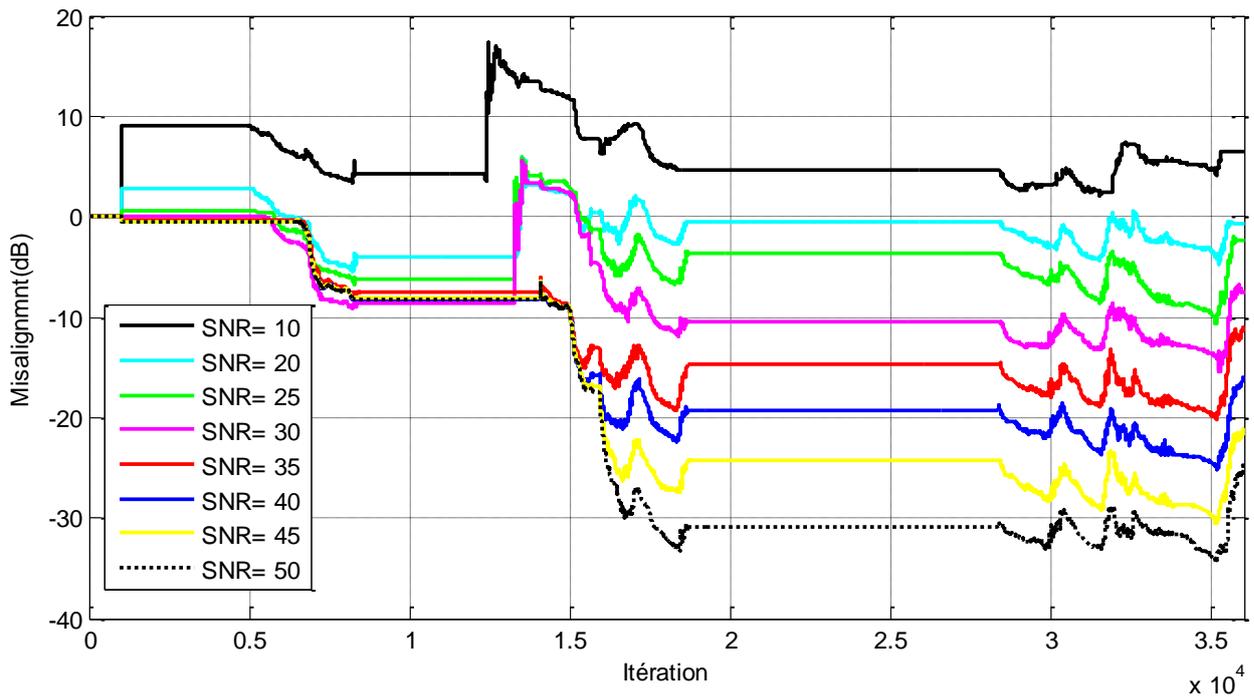


Figure III.11. Convergence AEC avec DTD-NCC dans un environnement bruité.

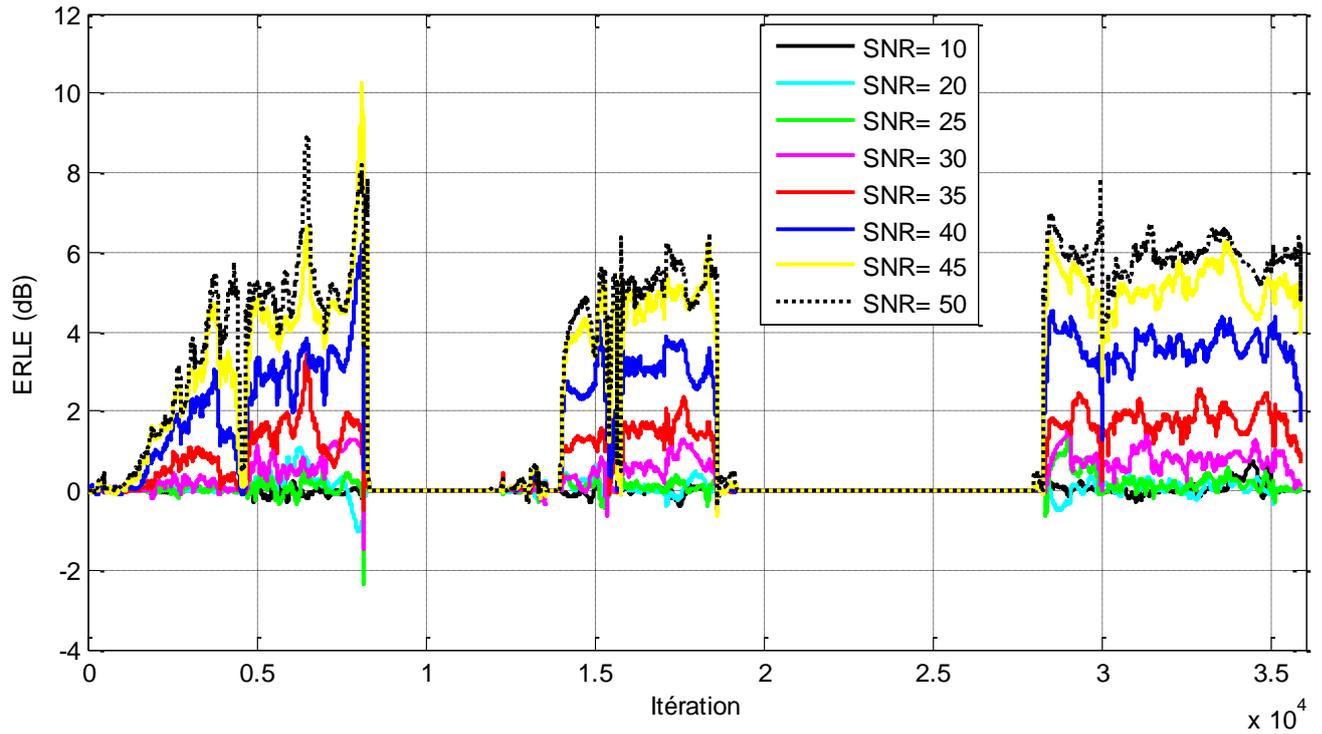


Figure III.12. ERLE DTD-Geigel dans un environnement bruité.

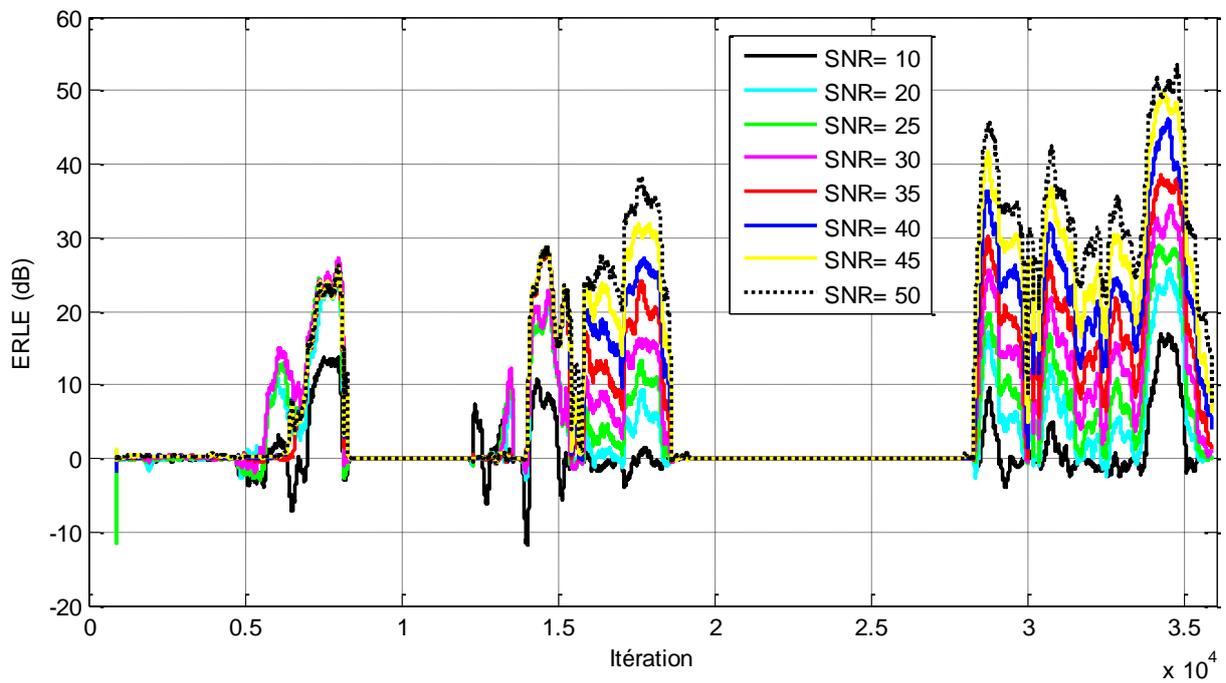


Figure III.13. ERLE du DTD NCC dans un environnement bruité.

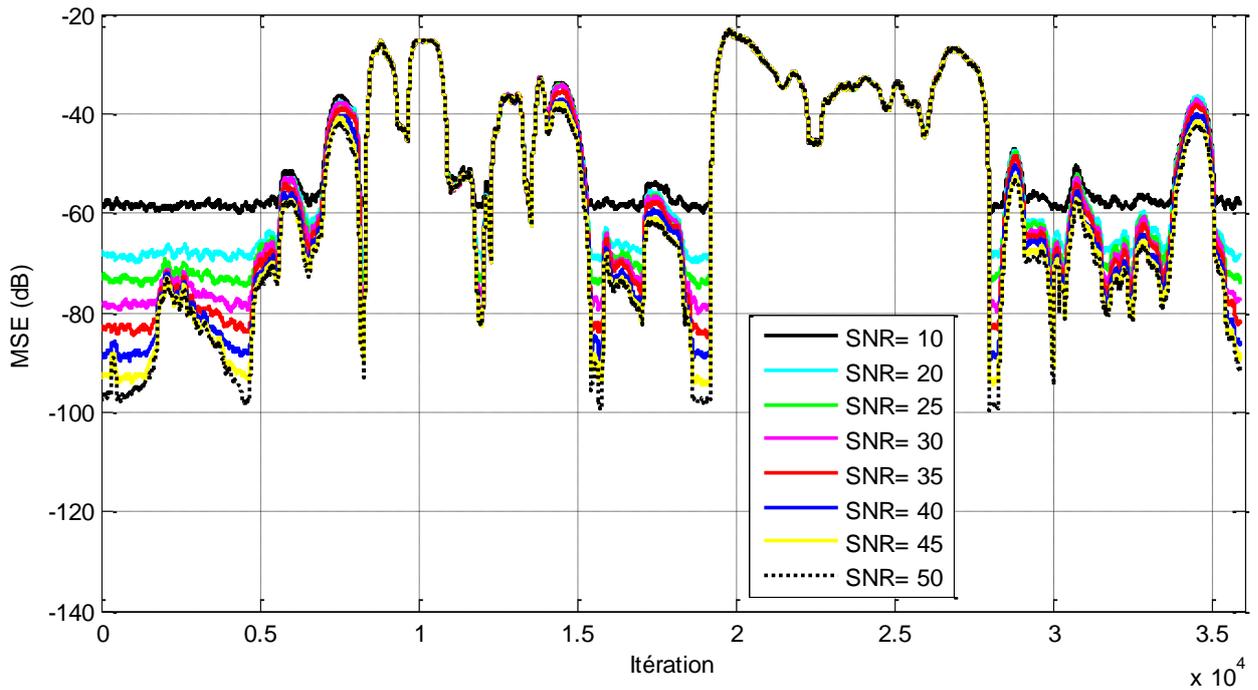


Figure III.14. MSE du DTD Geigel dans un environnement bruité.

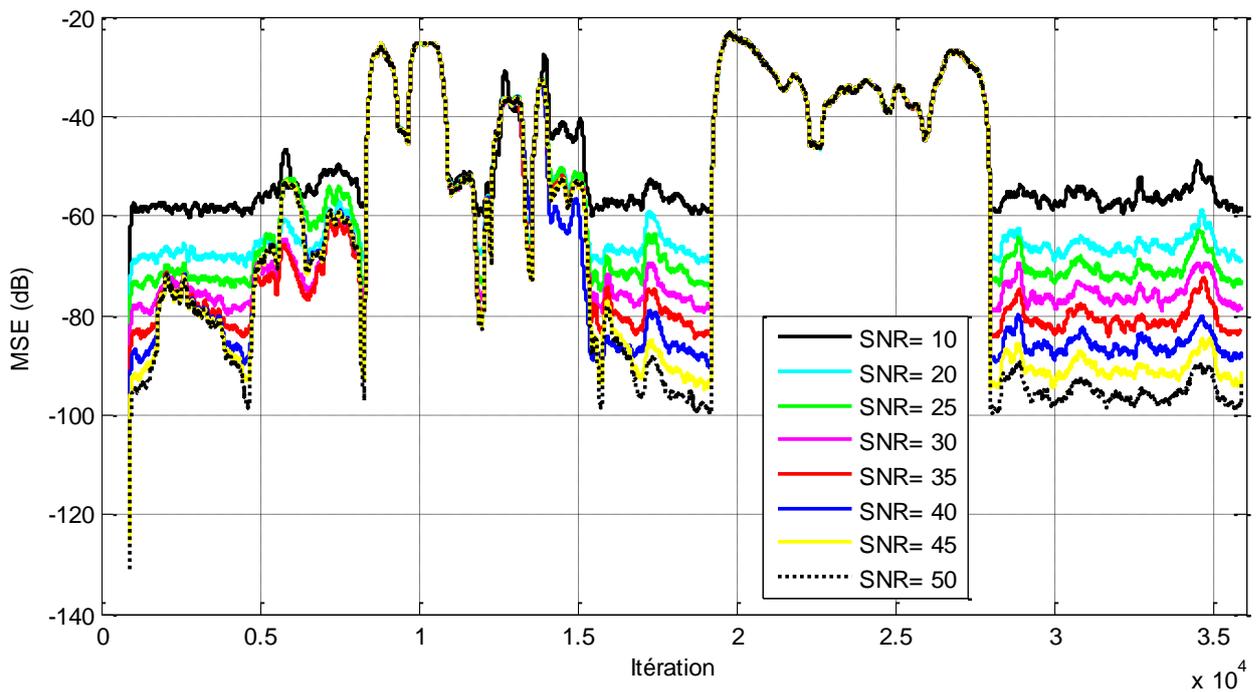


Figure III.15. MSE du DTD NCC dans un environnement bruité.

Les tableaux III.7 et III.8 résument respectivement la moyenne d'ERLE et de MSE obtenues pour les deux DTDs pour différents niveaux du SNR.

Tableau III.7. La moyenne d'ERLE des deux méthodes Geigel et NCC

SNR	Moyenne d'ERLE	
	Geigel	NCC
10	-0.0107	1.2050
20	-0.0094	2.6558
25	0.0549	4.3671
30	0.2426	5.7150
35	0.9359	6.8848
40	1.5027	8.3293
45	2.3379	9.7703
50	2.7423	11.2018

Tableau III.8. La moyenne du MSE dans les deux méthodes Geigel et NCC

SNR	Moyenne de MSE	
	Geigel	NCC
10	-47.0835	-48.8773
20	-50.6818	-52.6894
25	-52.0626	-54.5805
30	-53.2672	-56.7788
35	-54.5103	-58.6410
40	-55.9505	-60.5725
45	-57.1971	-62.3427
50	-57.8468	-64.0060

Les figures III.16, III.17 et III.18 représentent une comparaison en utilisant les trois critères d'évaluation des performances de l'AEC avec et sans DTD pour un niveau SNR de 40 dB, ces comparaisons confirment l'indispensabilité d'un DTD dans un AEC et montrent également que NCC est beaucoup plus performant que Geigel.

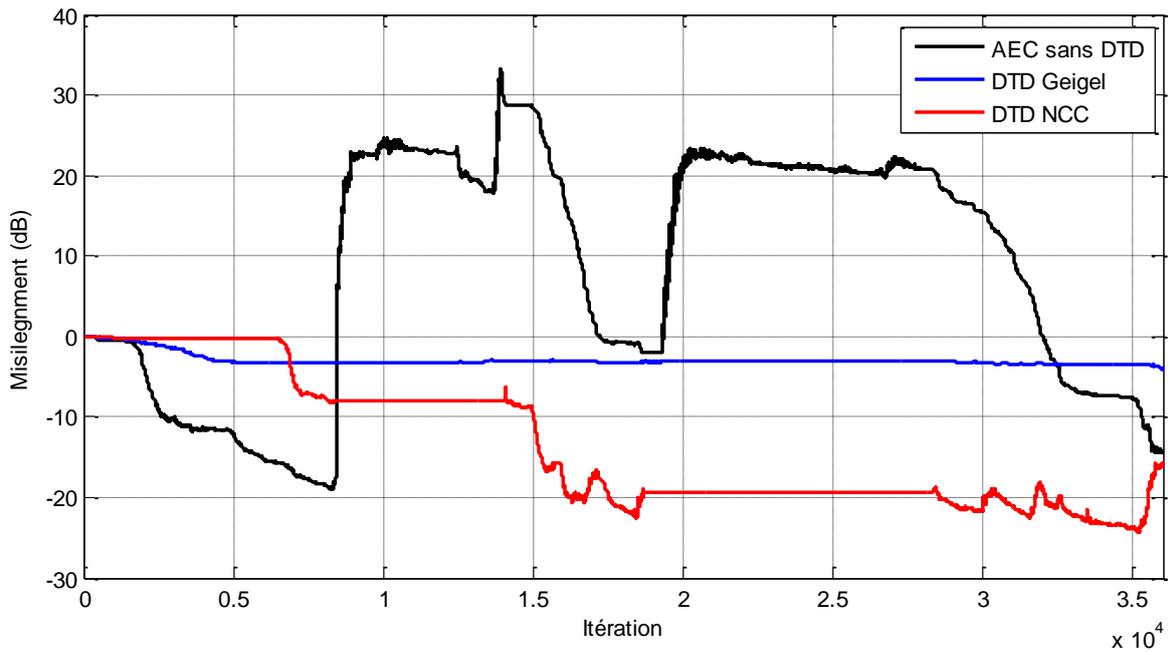


Figure III.16. Convergence avec et sans DTD (SNR=40 dB).

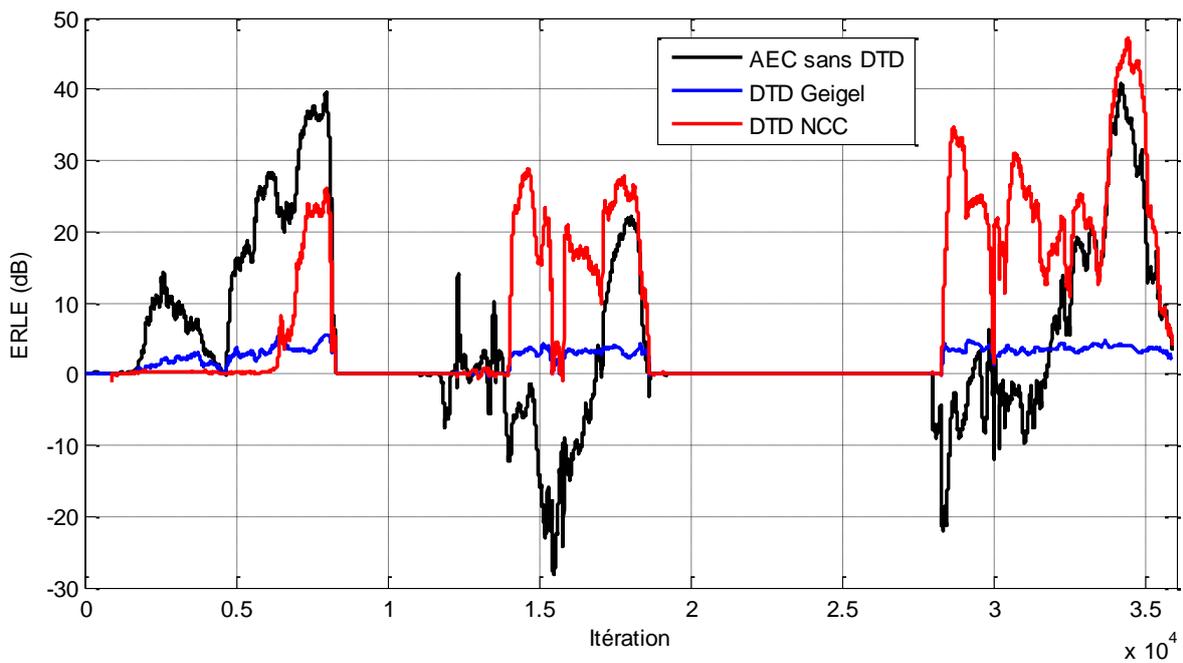


Figure III.17. ERLE avec et sans DTD (SNR=40 dB).

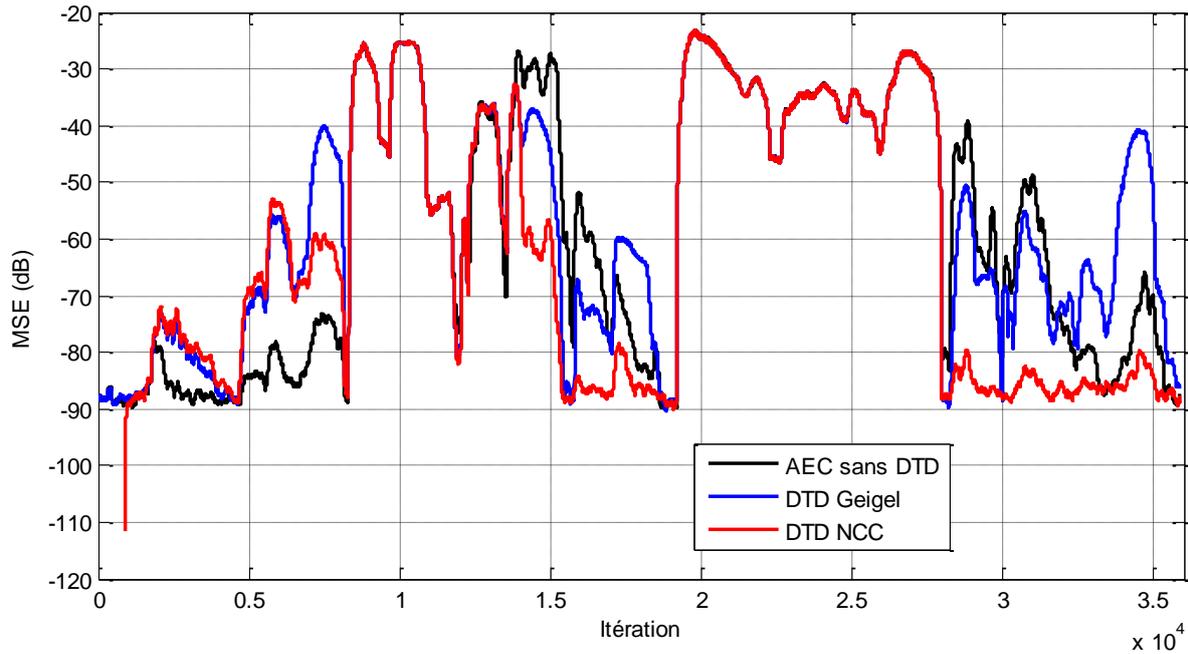


Figure III.18. MSE avec et sans DTD (SNR=40 dB).

III.4.4. Etude de l'écho restant

Dans une dernière étape nous allons étudier l'écho restant qui est le résultat de soustraction de $v(n)$ du signal d'erreur $e(n)$:

$$Echo_restant = e(n) - v(n) \tag{III.5}$$

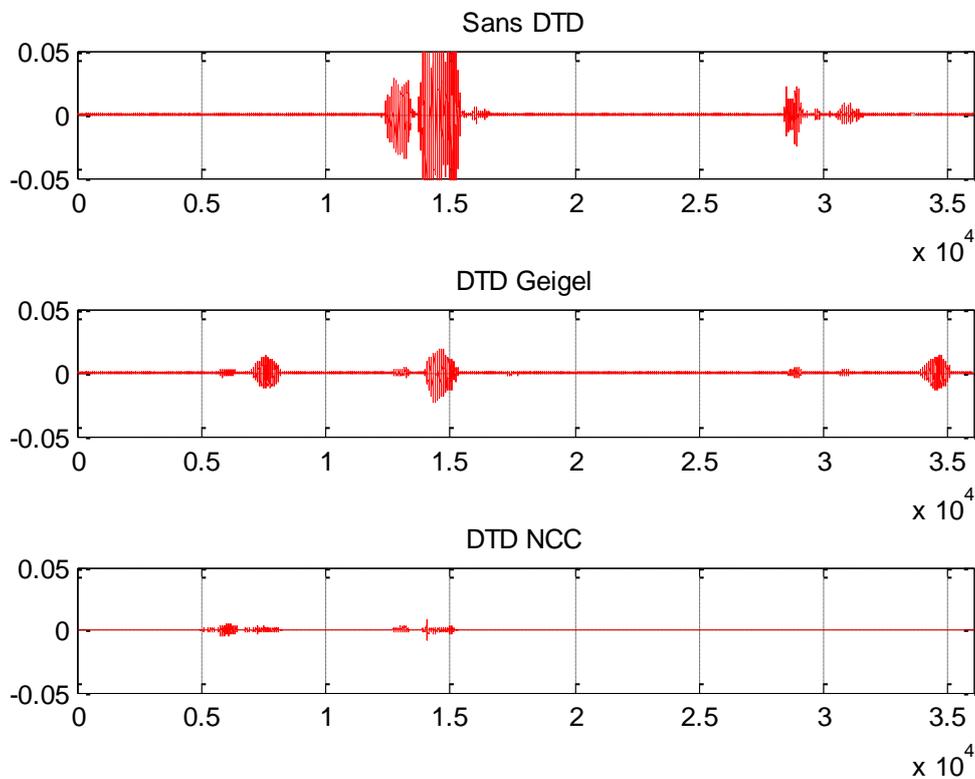


Figure III.19. Comparaison entre les signaux d'erreur des méthodes de Geigel et NCC.

Rappelons que $e(n)$ est le signal à transmettre au locuteur lointain, il doit contenir donc la voix du locuteur local dépourvue de l'écho acoustique. La figure III.19 représente l'écho restant sans DTD et en utilisant les deux DTDs. L'énergie totale de cet écho restant est de : 1.70 sans DTD, 0.15 en utilisant Geigel et 0.006 en utilisant NCC.

III.5. Conclusion

Dans ce chapitre nous avons présenté les résultats obtenus de la mise en œuvre de deux algorithmes de détection de double parole dans un système de communication mains-libres, l'objectif est de montrer l'utilité de la présence d'un DTD dans un AEC d'une part et d'une autre part la comparaison ces deux algorithmes suivant des critères connus et souvent utilisés dans ce domaine.

Les résultats obtenus montrent qu'un DTD est indispensable dans un AEC notamment quand on a à faire avec un filtre adaptatif plus rapide en termes de convergence car cette vitesse influe directement sur sa sensibilité à la présence de double parole. D'une autre part ces résultats ont montrés que le DTD NCC qui est basé sur le calcul de la crosse-corrélation est plus performant que le DTD Geigel qui est basé sur un simple calcul d'un rapport des énergies et ce suivant les critères d'évaluation choisis et même quand on a introduit le bruit avec différents niveaux SNR.

Conclusion

Générale

Dans ce travail nous avons traité le problème d'annulation d'écho acoustique qui est considéré comme l'un des problèmes majeurs dans les communications mains libres, l'objectif est de mettre en œuvre deux algorithmes de détection DT, le premier est l'algorithme Geigel qui est le plus ancien et qui a montré son efficacité avec l'écho hybride, et le deuxième est le NCC qui est considéré comme l'un des meilleurs DTDs assurant d'un côté de bonnes performances et d'un autre côté une complexité de calcul faible. Concernant l'algorithme de filtrage adaptatif, nous avons opté pour le célèbre NLMS connu pour son compromis entre la vitesse de convergence et la complexité de calcul.

Les résultats de simulation sont obtenus en utilisant des séquences audio réelles représentant la voix des deux locuteurs ainsi qu'une réponse impulsionnelle modélisant le chemin d'écho, l'évaluation des performances est effectuée suivant plusieurs critères : la convergence du filtre (misalignement), ELRE, MSE ainsi que l'écho résiduel, des tests ont été effectués dans différentes situations dans un premier temps sans prise en compte du bruit ensuite en ajoutant du bruit avec différents niveaux SNR. Ces résultats montrent d'un côté l'indispensabilité de la présence d'un DTD dans un annuleur d'écho acoustique et d'un autre côté la supériorité du NCC par rapport à Geigel.

En perspective nous souhaitons dans l'avenir :

- Utiliser d'autres DTDs, à savoir ceux qui sont basés sur l'analyse fréquentielle.
- Utiliser d'autres algorithmes de filtrage adaptatif plus rapides (le RLS entre autres)
- Implémentation de la totalité de l'annuleur d'écho pour évaluer la complexité de calcul et de vérifier la contrainte temps réel.

Références bibliographiques

RÉFÉRENCES BIBLIOGRAPHIQUES

- [1] J. Benesty, T. Gänslér, D. R. Morgan, M. M. Sondhi, and S. L. "Gay, Advances in network and acoustic echo cancellation," Berlin, Germany: Springer-Verlag, 2001.
- [2] M. M. Sondhi, "An Adaptive Echo Canceller," Bell Syst. Tech. J., vol., 46, No. 3, pp 497511, Mar. 1967.
- [3] M.M. Sondhi, "Closed loop adaptive echo canceller using generalized filter networks." U.S. Patent 3 499 999, 1970.
- [4] J.L. Kelly and B.F. Logan, "Self-adaptive echo canceller," U.S. Patent 3 500 000, 1970.
- [5] N. Swaroopa Adapa, S. Bollu, " Performance Analysis of different Adaptive Algorithms based on Acoustic Echo Cancellation ", mémoire de master, université de Blekinge, Suède ,2010.
- [6] L. Lu, "Implementation of Acoustic Echo Cancellation For PC Applications Using MATLAB", mémoire de master, université de Stockholm, Suède, 2007.
- [7] Digital network echo cancellers, ITU-T Rec. G.168, 2002.
- [8] M. Djendi, A. Benallal, A. Guessoum, D. Berkani, "Three new versions for the Newton type adaptive filtering algorithm, in: Proceedings of the Seventh International Symposium on Signal Processing and Its Applications," ISSPA, pp.559–562, 2003.
- [9] W.LAIFAOU, "Implémentation des filtres adaptatifs sur une carteFPGA pour l'annulation d'écho acoustique", mémoire de master, université de Mohamed Larbi Ben M'hidi, Oum El bouaghi, Algérie, juin 2018.
- [10] F. YKHLEF, F. YKHLEF, H. AITSAADI, A .GUESSOUM, "Annulation d'Echo Acoustique et Suppression de Bruit pour les Communications Mains-libres", 5th International Conference: Sciences of Electronic, March 22-26, 2009, Tunisia.
- [11]. S. Haykin, "Adaptive filter theory," 3 Ed. New York: Prentice Hall Inc, 1996.

- [12] V. Das, A. Kar ,C.Mahesh , “ Advanced Adaptive Algorithms for Double Talk Detection in Echo Cancellers”, these de doctotat ,université de Bhubaneswar et Mesra, Inde, January 2015.
- [13] A.A.M.Muzahid, “Study on Efficient Adaptive Filtering Algorithms for Acoustic Echo Cancellation in Full-duplex Channel”,memoire de master , université de Chongqing, chine , May 2016.
- [14] BENVENISTE A., METIVIERM., PRIOURET P., “Algorithmes Adaptatifs et Approximations Stochastiques”, Masson, Paris, 1987.
- [15] E. Hänsler, G. Schmidt, “Speech and Audio Processing in Adverse Environments”, Springer Science & Business Media, 2008.
- [16] G. Szwoch, A. Czyżewski, M. Kulesza, “A low complexity double-talk detector based on the signal envelope,” Elsevier, Signal Processing, Vol. 88(11), pp. 2856-2862, 2008.
- [17] T. Gänsler, J. Benesty, The fast-normalized cross-correlation double-talk detector, Signal Process. 86 (2006) 1124–1139.
- [18] J. H. Cho, D. R. Morgan, and J. Benesty, "An objective technique for evaluating doubletalk detectors in acoustic echo cancelers," IEEE Trans. Speech Audio Processing, vol. 7, Nov. 1999, pp. 718-724.
- [19] D. Duttweiler, “A twelve-channel digital echo canceler,” IEEE Trans on Communications”, vol. 26, no. 5, pp. 647–653, 1978.
- [20] Y. Hua and W. Bo-Xiu, "A new double-talk detection algorithm based on the orthogonality theorem," Communications, IEEE Transactions on, vol. 39, pp. 1542-1545, 1991.
- [21] R. D. Wesel, "Cross-correlation vectors and double-talk control for echo cancellation," Unpublished, 1994.
- [22] J. Benesty, D. R. Morgan, and J. H. Cho, "A new class of doubletalk detectors based on cross-correlation," Speech and Audio Processing, IEEE Transactions on, vol. 8, pp. 168-172, 2000.

- [23] M HAMIDIA, A AMROUCHE, “Double-talk Detector Based on Speech Feature Extraction for Acoustic Echo Cancellation”, université Bab Ezzouar, Algeria , 2013.
- [24] M. Hamidia, A. Amrouche, “Double-talk detection using the singular value decomposition for acoustic echo cancellation,” in: IEEE International Conference on Communications, ICC, IEEE, Budapest, pp.4745–4749, 2013.
- [25] M. Hamidia, A. Amrouche,” A new structure for acoustic echo cancellation in double-talk scenario using auxiliary filter,” in: 14th International Workshop on Acoustic Signal Enhancement, IWAENC, IEEE, Juan-les-Pins, France, 2014, pp.253–257.
- [26] Sonika and Sanjeev Dhull, “Double Talk Detection in Acoustic Echo Cancellation based on Variance Impulse Response”, International Journal of Electronics and Communication Engineering, vol. 4, No. 5 , pp. 537-542, 2011.