

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne démocratique et populaire

وزارة التعليم العالي والبحث العلمي
Ministère de l'enseignement supérieur et de la recherche scientifique

جامعة سعد دحلب البلدية
Université SAAD DAHLAB de BLIDA

كلية التكنولوجيا
Faculté de Technologie

قسم الإلكترونيك
Département d'Électronique



Mémoire de Master

Mention Électronique
Spécialité Télécommunications et Réseaux

présenté par

SAFI Meriem

&

BENSAFIA Sara

Rehaussement du signal de parole par un algorithme fréquentiel de type Backward

Proposé par : Pr. DJENDI Mohamed

Année Universitaire 2015-2016

Remerciements

Je tiens à remercier en tout premier lieu Dieu tout puissant, de m'avoir permis de poursuivre mes objectifs et surmonter toutes les dures épreuves et moments difficiles pour aboutir à la fin de ce travail.

J'adresse mes sincères remerciements à tous les professeurs, en particulier mon encadreur Pr. DJENDI Mohamed pour m'avoir fait confiance et m'avoir dirigée durant cette année. Je tiens particulièrement à le remercier pour sa patience tout au long de ce travail, ainsi que pour la richesse de ses idées.

Je remercie mes très chers parents, qui ont toujours entaieut là pour moi, « Vous avez tout sacrifié pour votre fille n'épargnant ni santé ni efforts. Vous m'avez donné un magnifique modèle de labeur et de persévérance. Je suis redevable d'une éducation dont je suis fière ».

Que soient également remerciés tous les membres du jury, pour l'honneur et l'intérêt qu'ils m'ont accordé, en acceptant d'examiner mon manuscrit.

Mes remerciements les plus sincères à mon binôme Sara, et à tous mes amis qui ont contribué de près ou de loin à l'élaboration de ce mémoire.

SAFI Meriem

Remerciements

Je tiens à remercier en tout premier lieu Dieu tout puissant, de m'avoir permis de poursuivre mes objectifs et surmonter toutes les dures épreuves et moments difficiles pour aboutir à la fin de ce travail.

J'adresse mes sincères remerciements à tous les professeurs, en particulier mon encadreur Pr. DJENDI Mohamed pour m'avoir fait confiance et m'avoir dirigée durant cette année. Je tiens particulièrement à le remercier pour sa patience tout au long de ce travail, ainsi que pour la richesse de ses idées.

Je remercie mes très chers parents, ainsi que ma petite sœur Serina, je vous aime très fort.

Que soient également remerciés tous les membres du jury, pour l'honneur et l'intérêt qu'ils m'ont accordé, en acceptant d'examiner mon manuscrit.

Mes remerciements les plus sincères à mon binôme Maya, et à tous mes amis qui ont contribué de près ou de loin à l'élaboration de ce mémoire.

BENSAFIA Sara

ملخص: في هذا العمل، نأخذ بعين الاعتبار مشكلة الرفع من إشارة الكلام المطبقة على أنظمة الهاتف اللاسلكي المستعمل في السيارات، في هذا السياق قمنا بإنجاز هيكلية ثنائية القناة عن مصدر أعمى، المطبقة على ميكروفونين متباعدين نسبيا. و بناء على هذا قمنا باقتراح خوارزمية اللارتباط المتمائل بتكيف، و التي تعتمد على طريقة الفصل المتكرر للقضاء على التشويش، هذه الخوارزمية الجديدة تظهر نتائج إيجابية و مستحبة للغاية

كلمات المفاتيح: القضاء على التشويش، هيكلية ثنائية القناة، الاستشعار، فصل مصدر أعمى.

Résumé : Dans ce manuscrit, nous considérons une problématique de rehaussement du signal de parole, appliquée aux systèmes de téléphonie mains-libre en voiture. Dans ce contexte, nous mettons en œuvre une structure bi-voies de séparation aveugle de sources, impliquant deux microphones relativement espacés. On propose un nouvel algorithme (BFD-SAD) de decorrélation symétrique adaptative, qui se base sur la méthode de séparation récursive (Backward), afin d'éliminer le bruit. Ce nouvel algorithme montre des résultats particulièrement favorables.

Mots clés : Débruitage ; Bi-capteurs ; Filtrage adaptatif ; SAS.

Abstract: In this manuscript, we consider a problem of enhancing speech signal applied to the hands-free phone systems drive. In this context, we are implementing a two-channel blind source separation structure, involving two relatively spaced microphones. It proposes a new algorithm (BFD-SAD) adaptive symmetric decorrelation, which is based on the recursive method of separation (Backward) to eliminate noise. This new algorithm shows particularly favorable results.

Keywords: Noise cancellation; Bi-sensors; Adaptive filtering; BSS.

Listes des acronymes et abréviations

- RIF : Réponse Impulsionnelle Finie.
- MA : Moyenne Ajustée.
- RII : Réponse impulsionnelle Infinie.
- ARMA : Auto Régressif à Moyenne Ajustée.
- MIPS : Million Instructions Per Second.
- VLSI : Very Large Scale Integration.
- LMS : Least Mean Square.
- NLMS : Normalised Mean Square.
- IEEE: Institute of Electrical and Electronics Engineers.
- ANC : Adaptive Noise Canceller.
- AR : Autorégressif.
- DSP : Densité Spectrale de Puissance.
- VAD : Voice Activity Detector.
- LPC : Linear Prediction Coding.
- DFT : Discrete Fourier Transform.
- DCT : Discrete Cosine Transform.
- MMSE: Minimum Mean Square Error.
- MSE: Mean Square Error.
- BMI: Blind Mixture Identification.
- IAM: Identification Aveugle de Mélanges.
- SAS: Séparation Aveugle de Sources.
- BSS : Blind Source Separation.
- OLA : Overlap and Add.
- OLS : Overlap and Save.
- FFT : Fast Fourier Transform.
- IFFT : Inverse Fast Fourier Transform.
- BTD-SAD: Backward Time Domain-Symmetric Adaptive Decorrelation.
- BFD-SAD: Backward Frequency Domain-Symmetric Adaptive Decorrelation.
- SNR: Signal to Noise Ratio.
- USASI: United State of America Standard Institute.
- CD : Cepstral Distance.
- SM : System Mismatch.
- ACR : Absolut Category Rating.
- DCR : Degradation Category Rating.
- CCR : Comparison Category Rating.
- CMOS: Comparative Mean Opinion Score.

Table des matières

Introduction générale.....	1
----------------------------	---

Chapitre 1 : Filtres et Filtrage adaptatif.

1.1	Introduction	3
1.2	Notions sur le filtrage numérique	3
b	Invariance dans le temps	4
1.2.2	La convolution	4
1.2.3	Transformée en z	5
1.2.4	Causalité et stabilité	5
a	Causalité.....	5
b	Stabilité	5
1.2.5	Spécification d'un filtre numérique	6
1.2.6	Classification des filtres numériques	7
a	Filtres à réponse impulsionnelle finie (RIF) [MA].....	7
b	Filtres à réponse impulsionnelle infinie (RII) [ARMA]	8
1.3	Filtrage adaptatif.....	9
1.3.2	Le filtre adaptatif.....	9
1.3.3	Principe du filtrage adaptatif	10
1.3.4	Critères de comparaison des algorithmes adaptatifs	10
1.3.5	Choix de l'algorithme	11
1.4	Algorithmes du gradient stochastiques	11
1.4.1	Filtrage linéaire optimal	11
1.4.2	Algorithme du gradient stochastique LMS.....	13
1.4.3	Algorithme du gradient stochastique normalisé NLMS	14
1.5	Applications.....	15
1.6	Conclusion	16

Chapitre 2 : Rehaussement du signal de parole par filtrage adaptatif.

2.1	Introduction	17
2.2	La parole et le bruit.....	17

2.2.1	Définition de la parole.....	17
2.2.2	Mécanisme de phonation	17
2.2.3	Sons voisées et non-voisées.....	18
2.2.4	Modélisation autorégressive (AR) du signal de parole	19
2.2.5	Bruit acoustique	19
2.2.6	Bruit Johnson (bruit blanc).....	20
2.2.7	Bruit coloré.....	20
2.2.8	Silences.....	21
2.3	Principe de la réduction de bruit	21
2.3.1	Estimation du bruit.....	21
2.3.2	Les types de mélanges	23
a	Mélanges linéaires instantanés.....	23
b	Mélanges linéaires à atténuations et retards	23
c	Mélanges linéaires convolutifs.....	23
2.4	Méthodes mono/bi/multi-capteurs pour la réduction du bruit et le rehaussement du signal de parole	24
2.4.1	Méthodes d'annulation du bruit mono-capteurs	25
a	Réduction de bruit basée sur le modèle autorégressif.....	25
b	Approche par soustraction spectrale.....	26
2.4.2	Méthodes de débruitage bi-capteurs	27
a	Séparation aveugle de source	28
b	Annulation adaptative du bruit.....	29
2.4.3	Méthodes de débruitage multi-capteurs.....	31
2.5	Conclusion.....	31

Chapitre 3 : Proposition d'un nouvel algorithme de decorrélation fréquentiel symétrique.

3.1	Introduction	32
3.2	Séparation aveugle de source (SAS).....	32
3.3	Solution de la séparation de sources	33
3.3.1	Structure directe (Forward)	33
3.3.2	Définition de la Backward	34
3.3.3	Structure recursive (Backward).....	35
3.4	Overlap and add (OLA)	37

3.5	Proposition d'un nouvel algorithme de séparation de source aveugle (BFD-SAD)	38
3.6	Conclusion	41

Chapitre 4 : Résultats de simulations.

4.1	Introduction	42
4.2	Modèle expérimental du mélange	42
4.3	Signaux de simulations utilisés	43
4.4	Simulations des réponses impulsionnelles	45
4.5	Signaux de mélanges	46
4.6	Détection d'activité vocale	48
4.7	Critères objectifs de qualité	49
4.7.1	Le critère du rapport Signal a Bruit (SNR)	49
4.7.2	Le critère de la Distance Cepstrale (CD)	50
4.7.3	Le critère du désajustement	50
4.8	Résultats de simulations du nouvel algorithme BFD-SAD	51
4.9	Influence des paramètres L et SNR sur l'algorithme BFD-SAD	54
4.9.1	Effet de la taille du filtre adaptatif L	55
4.9.2	Effet du SNR du mélange	56
4.10	Comparaison entre les algorithmes BFD-SAD et BTD-SAD	58
4.11	Critères subjectifs de qualité	62
4.12	Conclusion	65
	Conclusion général	66
	Bibliographie	68

Liste des figures

Figure 1.1. filtre numérique.....	3
Figure 1.2. Réponse fréquentielles idéales des 4 filtres de base	6
Figure 1.3. Principe du filtrage adaptatif	10
Figure 1.4. Filtre de Wiener	12
Figure 2.1. Représentation de l'appareil phonatoire humain	18
Figure 2.2. Exemple de son voisé [e].....	19
Figure 2.3. Exemple de son non-voisé [s]	19
Figure 2.4. Schéma explicatif du principe de la réduction de bruit.....	21
Figure 2.5. Principe de l'estimation du bruit	22
Figure 2.6. Principe du mélange avec deux sources et deux capteurs.....	24
Figure 2.7. Principe de l'approche par soustraction spectrale	27
Figure 2.8. Configuration générale de la séparation aveugle de source	29
Figure 2.9. Structure d'annulation adaptative du bruit.....	30
Figure 3.1. Structure de séparation de sources directe(Forward)	33
Figure 3.2. Structure de séparation de sources récursive (Backward)	35
Figure 3.3. Nouvel algorithme fréquentiel proposé	38
Figure 3.4. Algorithme proposé (FD-SAD) contrôlé par un système de détection manuel d'activité vocal (DAV)	41
Figure 4.1. : Modèle de mélange convolutif.....	43
Figure 4.2. Signal de parole utilisé, $F_s = 8$ KHz	44
Figure 4.3. Signal du bruit USASI	44
Figure 4.4. Signal du bruit blanc	44
Figure 4.5. Exemple d'une réponse impulsionnelle simulée h_{12}	45
Figure 4.6. Exemple d'une réponse impulsionnelle simulée h_{12}	45
Figure 4.7. Signal de mélange p_1 avec son spectrogramme	46
Figure 4.8. Signal de mélange p_2 avec son spectrogramme	47
Figure 4.9. Signal de parole utilisé avec la segmentation (DAV) manuelle	48
Figure 4.10. Signal de parole originale avec son spectrogramme	52
Figure 4.11. Signal de bruit USASI avec son spectrogramme	53
Figure 4.12. Signal de bruit USASI avec son spectrogramme	53
Figure 4.13. Comparaison entre la réponse impulsionnelle réelle h_{21} et estimée w_{21}	54

Figure 4.14. Effet de la taille du filtre sur le nouvel algorithme BFD-SAD en utilisant un bruit USASI dans le mélange.....	55
Figure 4.15. Effet de la taille du filtre sur le nouvel algorithme BFD-SAD en utilisant un bruit blanc dans le mélange.....	56
Figure 4.16. Effet du RSB d'entrée sur le nouvel algorithme BFD-SAD en utilisant un bruit USASI dans le mélange.....	57
Figure 4.17. Effet du RSB d'entrée sur le nouvel algorithme BFD-SAD en utilisant un bruit blanc dans le mélange	57
Figure 4.18. Effet du RSB d'entrée sur le nouvel algorithme BFD-SAD en utilisant un bruit USASI dans le mélange.....	58
Figure 4.19. Signal de parole originale (Haut), Signal $S_1(n)$ de l'algorithme BFD-SAD (Milieu) signal $U_1(n)$ de l'algorithme BTD-SAD (Bas)	60
Figure 4.20. Evaluation du critère de la distance cepstrale des deux algorithmes BFD-SAD et BTD-SAD	61
Figure 4.21. Evolution du critère du rapport signal sur bruit pour les deux algorithmes BFD-SAD [Magenta], et BTD-SAD [Noir]	61
Figure 4.22. Comparaison du désajustement (SM) obtenu avec les deux algorithmes BFD-SAD [Magenta] et BTD-SAD [NOIR]	62

Liste des tableaux

Tableau 1.1. Les différentes classes de bruit.....	20
Tableau 4.1. Les différentes classes de bruit.....	63
Tableau 4.2. Comparaison subjective entre BFD-SAD et BTD-SAD	64

Introduction générale

Un bon nombre d'algorithmes appliqués à des applications temps-réel comme la téléphonie se base sur l'estimation du bruit lors des temps de silence. Le rehaussement de la parole est un domaine de la télécommunication dont les recherches prennent de plus en plus d'ampleur. En effet, dans un monde où la télécommunication connaît un véritable essor, les technologies se doivent d'être de plus en plus performantes afin de plaire au plus grand nombre. Les applications qui nécessitent un rehaussement de la parole sont très nombreuses, la plus évidente étant sans doute celle de la téléphonie mobile. En effet il n'est pas rare qu'un locuteur se retrouve dans un environnement bruyant et hostile à sa communication. Les techniques de rehaussement de la parole vont alors lui être utiles afin d'améliorer la clarté et l'intelligibilité de sa voix [1].

Les kits pour téléphonie mains-libres en voiture sont un équipement qui devient de plus en plus standard dans les véhicules actuels. Ces accessoires répondent à un besoin de communiquer tout en conduisant, pour des raisons professionnelles ou personnelles. Or, la voiture s'avère un environnement acoustique particulièrement difficile. En effet, un habitacle de voiture présente une forte réverbération, du fait de la présence de nombreuses surfaces vitrées, et il est aussi très bruyant. Ce bruit vient de multiples sources, comme le moteur, le roulement de pneu sur la route, le vent, la circulation environnante...et ces sources varient fortement d'une condition de conduite à l'autre. La réduction du bruit ambiant au niveau de la prise de son dans l'habitacle constitue donc un élément majeur dans le confort des utilisateurs de ce type d'équipement. L'objectif essentiel des travaux effectués dans le cadre de ce mémoire de recherche est de fournir une solution efficace de réduction de bruit pour cette application, en utilisant deux capteurs sonores.

Parmi les méthodes de débruitage bi-capteurs, on retrouve la séparation aveugle de source de type récursive, qui constitue l'objet de ce travail.

Nous présentons ci-après les différents chapitres de ce mémoire :

- ♣ **Le chapitre 1** : On a décrits les filtres numériques et introduit le filtrage adaptatif.
- ♣ **Le chapitre 2** : Nous avons présenté les différentes méthodes de débruitage en fonction du nombre de microphones utilisés.
- ♣ **Le chapitre 3** : On a proposé un nouvel algorithme de decorrélation fréquentiel symétrique, utilisant la structure récursive.
- ♣ **Le chapitre 4** : Nous illustrons les résultats de simulations de notre nouvel algorithme.

Chapitre 1 Filtres numériques

1.1 Introduction

Le filtrage adaptatif intervient quand il faut réaliser, simuler ou modéliser un système dont les caractéristiques évoluent dans le temps. Il conduit à la mise en œuvre de filtres à coefficients variables dans le temps. Les variations des coefficients sont définies par un critère d'optimisation et réalisées suivant un algorithme d'adaptation, qui est déterminé en fonction de l'application. Il existe une grande variété de critères et d'algorithmes possibles. Dans le présent chapitre, on présente les filtres numériques et filtrage adaptatif.

1.2 Notions sur le filtrage numérique

1.2.1 Définition

On appelle filtre numérique un système utilisé pour modifier la distribution fréquentielle d'un signal numérique d'entrée, en le transformant en un signal numérique désirée en sortie. Les filtres numériques sont, pour les signaux échantillonnés, les équivalents des filtres analogiques pour les signaux continus. En raison du développement des circuits intégrés rapides, les filtres numériques deviennent plus intéressants que les filtres analogiques en apportant de nombreux avantages : précision, fiabilité, stabilité, adaptabilité et facilité de commande [2]. La figure ci-dessous est une représentation d'un filtre numérique à une entrée et une sortie.

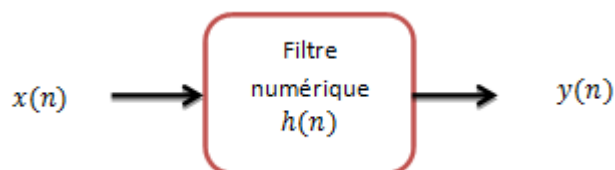


Figure 1.1. Filtre numérique.

Le filtre numérique est un système discret linéaire invariant dans le temps, il relie la sortie $y(nT_e) = y(n)$ à l'entrée $x(nT_e) = x(n)$ à chaque instant nT_e , T_e étant la période d'échantillonnage du signal.

a Linéarité

Soit deux suites $x_1(n)$ et $x_2(n)$ avec les sorties correspondantes $y_1(n)$ et $y_2(n)$. Dire que le système est linéaire signifie que [3] :

$$\begin{cases} \forall \lambda_1 \in \mathcal{R} \\ \forall \lambda_2 \in \mathcal{R} \end{cases} \Rightarrow \lambda_1 x_1(n) + \lambda_2 x_2(n) \Rightarrow \lambda_1 y_1(n) + \lambda_2 y_2(n) \quad (1.1)$$

b Invariance dans le temps

Soit la suite $x(n)$ et la sortie correspondante $y(n)$, dire que le système est invariant dans le temps signifie qu'à la suite $x(n - n_0)$ correspond la sortie $y(n - n_0)$, et ceci quel que soit n_0 [3].

La sortie d'un filtre numérique est le produit de convolution de l'entrée par sa réponse impulsionnelle $h(n)$.

$$y(n) = h(n) * x(n) \quad (1.2)$$

$$y(n) = \sum_{k=-\infty}^{k=+\infty} h(k)x(n - k) = \sum_{k=-\infty}^{k=+\infty} h(k)h(n - k) \quad (1.3)$$

1.2.2 La convolution

Une impulsion brève, injectée à l'entrée d'un système de transmission linéaire, continu et stationnaire, donne en sortie un signal de durée finie. Cette réponse est appelée « réponse impulsionnelle » (ou percussionnelle) du filtre et notée $h(t)$. Dans le cas général, c'est-à-dire pour un signal d'entrée quelconque, nous avons une relation mathématique qui lie le signal d'entrée $e(t)$ et le signal de sortie $s(t)$ pour un système de transmission Linéaire Stationnaire, et Continue ; soit :

$$s(t) = \int_{-\infty}^{+\infty} e(\tau) \cdot h(t - \tau) d\tau = e(t) * h(t) \quad (1.4)$$

Cette opération, appelée « convolution » et notée $*$, exprime la réponse à un signal quelconque à partir de la réponse à un signal type (réponse impulsionnelle) ; la

réponse dépend du filtre, caractérisé par $h(t)$, et du signal. Le calcul de la convolution est complexe. Il nécessite de nombreuses étapes de calculs : pour chaque point de la réponse $s(t)$, il est nécessaire d'élaborer la fonction $h(t - \tau)$, symétrique de la réponse impulsionnelle par rapport à l'axe des ordonnées et décalée temporellement, puis le produit par le signal d'entrée $e(\tau)$ et enfin l'intégration sur la variable τ .

Les filtres, qui sont définis comme des systèmes de transmission linéaires, continus et stationnaires, sont des systèmes de convolution [2].

1.2.3 Transformée en z

Dans le cas des signaux analogiques, on dispose de transformées (transformée de Fourier ou transformée de Laplace) permettant d'étudier et de traiter les signaux dans des domaines plus aisés (respectivement plan fréquentiel ou domaine de Laplace). Dans le cas des signaux discrets, ces transformées sont très limitées, en particulier pour les signaux possédant une infinité d'échantillons. Pour cela, une transformée des signaux discrets a été introduite : la transformée en z . Cette transformée en z , notée $S(z)$, d'un signal $s(t)$ causal, échantillonné à la période T_e , s'exprime à l'aide de la variable complexe z selon la relation suivante [2] :

$$S(z) = \sum_{k=0}^{+\infty} s(kT_e) z^{-k} = \sum_{k=0}^{+\infty} s(k) z^{-k} \quad z \in \mathbb{C} \quad (1.5)$$

1.2.4 Causalité et stabilité

a Causalité

Un filtre numérique est dit causal, si sa réponse impulsionnelle $h(n)$ est nulle pour $n < 0$.

Un filtre numérique est dit anti causal, si sa réponse impulsionnelle $h(n)$ est nulle pour $n \geq 0$ [4].

b Stabilité

On dit qu'un filtre numérique est stable, si à toute entrée bornée $x(n)$ correspond une sortie $y(n)$ bornée [4].

1.2.5 Types de filtres numériques

Avant qu'un filtre numérique ne soit conçu et implanté, nous avons besoin de définir ses spécifications.

Un filtre doit laisser passer certaines fréquences, alors qu'il doit en atténuer (voire éliminer) d'autres. Nous devons donc pouvoir représenter ces contraintes. Il y a quatre filtres de bases [5] :

1. les filtres passe-bas laissent passer les fréquences inférieures à une fréquence de coupure Ω_c et bloquent celles qui lui sont supérieures.
2. les filtres passe-haut bloquent les fréquences inférieures à une fréquence de coupure Ω_c et laissent passer celles qui lui sont supérieure.
3. les filtres passe-bande laissent passer les fréquences autour d'une fréquence centrale Ω_0 (ou comprises entre Ω_1 et Ω_2) et bloquent les autres.
4. les filtres réjecteur-de-bande bloquent les fréquences autour d'une fréquence centrale Ω_0 (ou comprises entre Ω_1 et Ω_2) et laissent passer les autres.

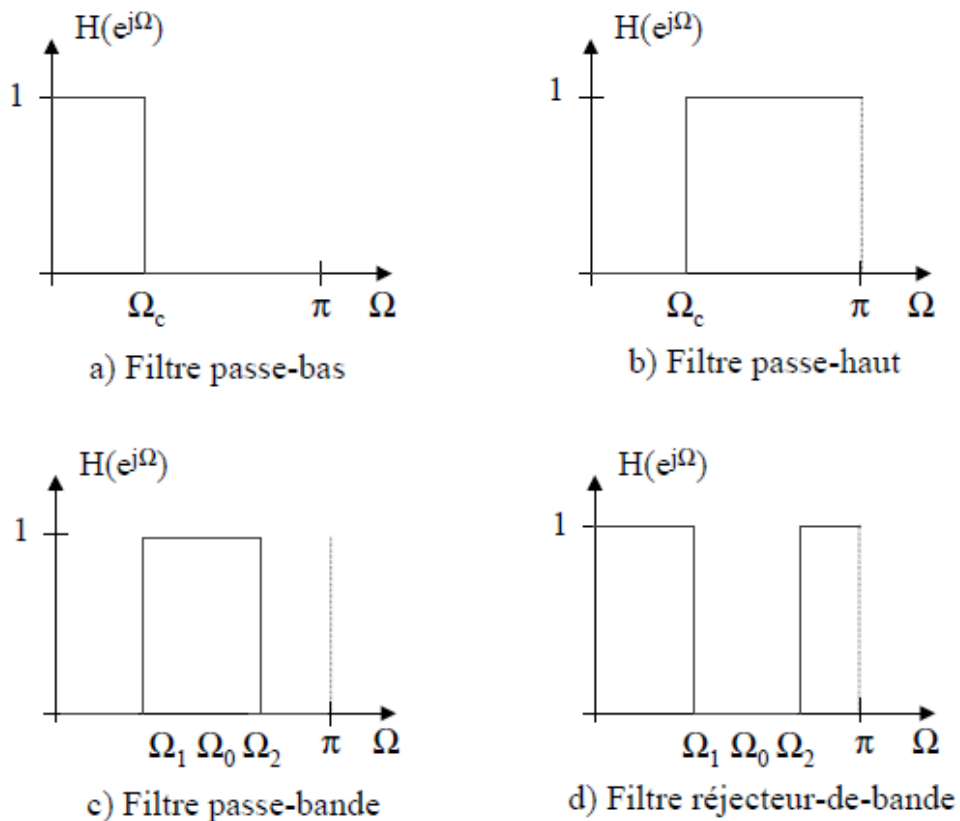


Figure 1.2. Réponse fréquentielles idéales des 4 filtres de base [5].

Les filtres représentés en figure (1.2) sont idéaux. Dans un cas réel il ne peut y avoir de discontinuités. Le passage entre zones passantes et zones atténuées se fait par des zones dites "de transition" dont la largeur va exprimer la sélectivité du filtre. Les bandes passantes et atténuées ne sont également pas idéales, elles contiennent des ondulations dont l'amplitude est exprimée par les paramètres d'ondulation en bande passante et d'atténuation [5].

Pour toutes ces raisons, la spécification d'un filtre est habituellement réalisée à partir d'un gabarit fréquentiel, défini entre 0 et π [5].

1.2.6 Classification des filtres numériques

Les filtres numériques se divisent en deux catégories selon la durée de la réponse impulsionnelle.

a Filtres à réponse impulsionnelle finie (RIF)

Les filtres RIF ne peuvent pas être dérivés des filtres analogiques. Ils sont cependant très largement utilisés car ils possèdent des propriétés uniques (phase linéaire, stabilité, flexibilité).

Les équations (1.6) et (1.7) montrent la fonction de transfert en z et l'équation aux différences correspondante de la forme générale d'un filtre RIF. N est appelé ici la longueur de la réponse impulsionnelle du filtre.

$$H(z) = \sum_{i=0}^{N-1} b_i \cdot z^{-i} \quad (1.6)$$

$$y(n) = \sum_{i=0}^{N-1} b_i \cdot x(n-i) = \sum_{i=0}^{N-1} h(i) \cdot x(n-i) \quad (1.7)$$

On remarque en exploitant l'équation (1.6) que les coefficients b_i du filtre sont également les valeurs de la réponse impulsionnelle $h(n)$, qui se trouve donc être limitée dans le temps [5].

$$H(z) = \sum_{i=0}^{N-1} b_i \cdot z^{-i} \Leftrightarrow h(n) = \sum_{i=0}^{N-1} b_i \cdot \delta(n-i) \quad (1.8)$$

$$h(n) = \begin{cases} b_n & \text{pour } 0 \leq n \leq N-1 \\ 0 & \text{ailleurs} \end{cases} \quad (1.9)$$

- **Les principales caractéristiques des filtres RIF sont [5]:**

1. une bande de transition qui sera toujours plus large qu'un filtre RII ayant le même nombre de coefficients ;
2. des méthodes de synthèse permettant de dériver n'importe quelle réponse fréquentielle ;
3. une stabilité inhérente ;
4. une plus grande stabilité numérique que les RII ;
5. une phase qui peut être exactement linéaire, par conséquent un temps de propagation de groupe constant et une absence de distorsion harmonique dans le signal ;
6. une plus grande facilité d'implantation dans un système numérique de traitement.

b Filtres à réponse impulsionnelle infinie (RII)

Les filtres analogiques ont nécessairement une réponse impulsionnelle infinie. Les filtres numériques RII se comportent de manière similaire, mis à part les effets dus à la discrétisation [5].

Cette catégorie de filtre est également caractérisée par une fonction de transfert en z contenant des pôles, et une équation aux différences récursives, c'est à dire lorsque la sortie $y(n)$ dépend à la fois des entrées et des sorties précédentes [5].

Les équations (1.10) et (1.11) montrent la fonction de transfert en z et l'équation aux différences correspondante de la forme générale d'un filtre RII. N est appelé ici l'ordre du filtre [5].

$$H(z) = \frac{N(z)}{D(z)} = \frac{\sum_{i=0}^N b_i \cdot z^{-i}}{1 + \sum_{i=1}^N a_i \cdot z^{-i}} \quad (1.10)$$

$$y(n) = \sum_{i=0}^N b_i \cdot x(n-i) - \sum_{i=0}^N a_i \cdot y(n-i) \quad (1.11)$$

A partir de l'équation (1.9), deux cas se présentent [5]:

1-si $N(z)$ n'est pas divisible par $D(z)$, on a un nombre infini de termes dans la division polynomiale de $N(z)$ par $D(z)$:

$$H(z) = \sum_{n=0}^{\infty} c_n \cdot z^{-n} \quad (1.12)$$

$$h(n) = c_n \text{ pour } n = 0 \dots \infty \quad (1.13)$$

$H(z)$ est un filtre RII.

2-si $N(z)$ est divisible par $D(z)$, on a un nombre fini de termes dans la division polynomiale de $N(z)$ par $D(z)$ [5]:

$$H(z) = \sum_{n=0}^N c_n \cdot z^{-n} \quad (1.14)$$

$$h(n) = c_n \text{ pour } n = 0 \dots N - 1 \quad (1.15)$$

$H(z)$ est un filtre RIF.

- **Les principales caractéristiques des filtres RII sont [5]:**

1. une bande de transition qui peut être étroite ;
2. des méthodes de synthèse par transposition des méthodes pour les filtres analogiques;
3. une instabilité potentielle due à des pôles situés en dehors du cercle unité (i.e. $|p_i| \geq 1$ quel que soit i) ;
4. une instabilité numérique (i.e. après quantification des coefficients et du signal) potentielle due au rebouclage.

1.3 Filtrage adaptatif

1.3.1 Définition

Le filtrage adaptatif est un outil puissant en traitement du signal, communications numériques, et contrôle automatique. Les applications sont diverses mais présentent les caractéristiques suivantes: on dispose d'une entrée $x(n)$ ainsi que de la réponse désirée (référence) $d(n)$ et l'erreur $e(n)$, qui est la différence entre $d(n)$ et la sortie du filtre $y(n)$, sert à contrôler (adapter) les valeurs des coefficients du filtre [6].

1.3.2 Le filtre adaptatif

Un filtre adaptatif est un filtre numérique dont les coefficients se modifient eux-mêmes en fonction des signaux extérieurs. Il est utilisé chaque fois qu'un environnement est mal connu ou est changeant, ou pour supprimer des perturbations situées dans le domaine des fréquences du signal utile, ce que les filtres classiques ne peuvent pas faire [7].

Le rôle primordial d'un filtre adaptatif est d'ajuster un vecteur de paramètre W pour un objectif bien défini (minimisation de l'erreur quadratique moyenne).

1.3.3 Principe du filtrage adaptatif

La technique du filtrage adaptatif se décompose classiquement en deux étapes [8] :

1. Une étape de filtrage qui permet d'obtenir une estimation du signal inconnu en convoluant le signal d'entrée $x(n)$ avec les coefficients du filtre adaptatif w .
L'erreur d'estimation $e(n) = y(n) - \hat{y}(n)$ est ensuite utilisée dans la partie adaptation pour mettre à jour les coefficients du filtre.
2. Une étape d'adaptation qui permet d'ajuster les coefficients du filtre adaptatif w suivant un algorithme donné.

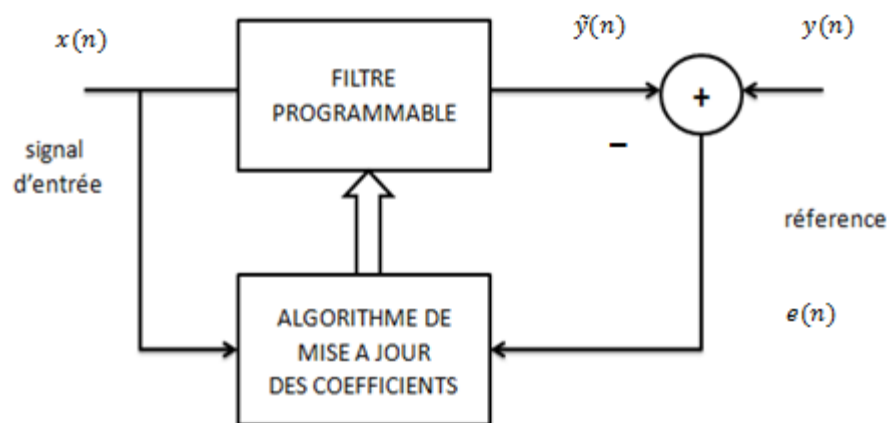


Figure 1.3. Principe du filtrage adaptatif [2].

1.3.4 Critères de comparaison des algorithmes adaptatifs

Les filtres adaptatifs peuvent être classés en fonction des choix qui sont faits sur les points suivants [6] :

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

1.3.5 Choix de l'algorithme

Le choix de l'algorithme se fera en fonction des critères suivants [6] :

- ❖ La rapidité de convergence qui sera le nombre d'itérations nécessaires pour converger vers la solution optimale de Wiener dans le cas stationnaire.
- ❖ La mesure de cette « proximité » entre cette solution optimale et la solution obtenue.
- ❖ La capacité de poursuite (tracking) des variations (non stationnarités) du processus. On examinera quels sont les algorithmes vraiment adaptatifs.
- ❖ La robustesse au bruit.
- ❖ La complexité (en nombre de MIPS).
- ❖ La structure (se prêtant plus ou moins à l'implémentation en VLSI).
- ❖ Les propriétés numériques (stabilité –précision) dans le cas d'une précision limitée sur les données et les coefficients (problèmes d'implémentation en virgule fixe).

1.4 Algorithmes du gradient stochastiques

1.4.1 Filtrage linéaire optimal

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.

➤ **Principe du filtre de Wiener**

La figure suivante schématise le principe d'un filtre dit de Wiener :

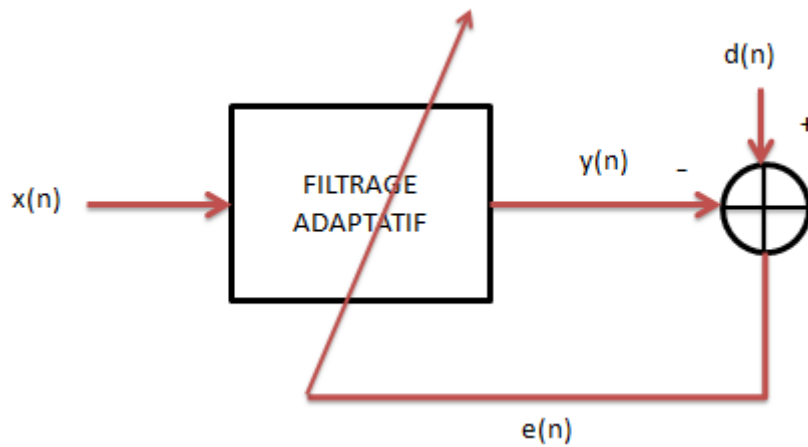


Figure 1.4. Filtre de Wiener [6].

Le principe du filtrage optimal, ainsi illustré sur la figure précédente et, de trouver le « meilleur » filtre c'est-à-dire celui permettant d'obtenir en sortie une réponse $y(n)$ la plus « proche » possible d'une réponse désirée $d(n)$ lorsque l'entrée est une certaine séquence $x(n)$ [6].

On note $e(n) = y(n) - d(n)$ l'erreur entre la réponse désirée $d(n)$ et la sortie $y(n)$. On note également $w(n)$ la réponse impulsionnelle du filtre.

Afin d'obtenir la configuration optimale du filtre, une méthode directe consiste à choisir une fonction appropriée pour l'estimation dite « fonction de coût $J(h)$ » (avec h étant un vecteur inconnu) qui permet d'obtenir les coefficients du filtre (w) de telle sorte à optimiser cette fonction dans un certain sens. Dans le cas d'un processus stochastique, on utilise une fonction de coût qui nous permet de minimiser l'erreur quadratique moyenne [9].

En effet, dans le filtrage de Wiener, on minimise l'erreur quadratique moyenne (EQM) [10] :

$$J = E \{e^2(n)\} \quad (1.16)$$

Ou

$$e(n) = d(n) - x^T(n)h \quad (1.17)$$

est le signal d'erreur, h est un filtre RIF de longueur L , $x(n)$ est le signal d'entrée, et $d(n)$ est le signal désiré.

Le gradient de J est:

$$\begin{aligned}
 g &= \frac{\partial J(h)}{\partial h} \\
 &= -2E\{x(n)e(n)\} \\
 &= -2E\{x(n)d(n) - x(n)x^T(n)h\} \\
 &= -2p + 2Rh.
 \end{aligned} \tag{1.18}$$

Avec p le vecteur d'intercorrélation, R la matrice d'autocorrélation,

On en déduit l'algorithme du gradient déterministe pour le filtrage de Wiener:

$$\begin{aligned}
 h(n+1) &= h(n) - \frac{1}{2}\mu g(n) \\
 &= h(n) + \mu[p - Rh(n)].
 \end{aligned} \tag{1.19}$$

Cet algorithme peut encore s'écrire en utilisant le signal d'erreur:

$$e(n) = d(n) - x^T(n)h(n) \tag{1.20}$$

$$h(n+1) = h(n) + \mu E\{x(n)e(n)\}. \tag{1.21}$$

Eventuellement, à l'infini, l'algorithme converge vers la solution optimale de Wiener-Hopf:

$$h(\infty) = h_{opt} = R^{-1}p \tag{1.22}$$

ou

$$p = Rh_{opt} \tag{1.23}$$

Le problème est qu'en pratique R et p ne sont pas connus.

1.4.2 Algorithme du gradient stochastique 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 [10].

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 [10]:

$$\hat{R}(n) = x(n)x^T(n) \tag{1.24}$$

$$\hat{p}(n) = x(n)d(n) \quad (1.25)$$

Ce sont simplement les estimées instantanées des corrélations.

En remplaçant $\hat{R}(n)$ et $\hat{p}(n)$ dans l'algorithme du gradient déterministe (première équation), on obtient:

$$\begin{aligned} h(n+1) &= h(n) + \mu[\hat{p}(n) - \hat{R}(n)h(n)] \\ &= h(n) + \mu x(n)[d(n) - x^T(n)h(n)] \\ &= h(n) + \mu x(n)e(n) \end{aligned} \quad (1.26)$$

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

Calcul de la sortie du filtre:

$$y(n) = h^T(n)x(n) \quad (1.27)$$

Calcul du signal d'erreur:

$$e(n) = d(n) - y(n) \quad (1.28)$$

Mise à jour du filtre:

$$h(n+1) = h(n) + \mu x(n)e(n) \quad (1.29)$$

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, où L est le nombre de coefficients du filtre [11].

➤ Convergence de l'algorithme LMS

L'analyse de la convergence du LMS se fait en utilisant les deux critères suivants [11]:

- Convergence en moyenne du filtre $h(n)$, c'est à dire:

$$\lim_{n \rightarrow \infty} E\{h(n)\} = h_{opt} \quad (1.30)$$

- Convergence du critère $J(n)$ (en moyenne quadratique), c'est à dire:

$$\lim_{n \rightarrow \infty} J(n) = J(\infty) = \text{constante} \quad (1.31)$$

1.4.3 Algorithme du gradient stochastique normalisé NLMS

L'algorithme NLMS (Normalized Least Mean Square) consiste à normaliser le pas d'adaptation μ dans l'algorithme LMS par rapport à l'énergie du signal d'entrée pour réduire au minimum l'effet de variation de la puissance du signal d'entrée et de rendre

ainsi la convergence plus ou moins uniforme en passant d'une étape d'adaptation à une autre [3] .

Dans l'équation qui va suivre, le pas d'adaptation μ est alors remplacé par un pas d'adaptation μ_n défini à chaque itération par [3]:

$$\mu_n = \frac{\mu}{x^T(n)x(n)} \quad (1.32)$$

La mise à jour des coefficients du filtre adaptatif par l'algorithme NLMS est alors donnée par :

$$w(n + 1) = w(n) + \frac{\mu}{x^T(n)x(n)\beta} e(n)x(n) \quad (1.33)$$

ou β est un facteur permettant de suivre plus ou moins rapidement les variations d'énergie dans un signal d'entrée $\{x\}$.

La convergence de cet algorithme est garantie pour un pas d'adaptation $0 < \mu < 2$.

L'intérêt de l'algorithme NLMS par rapport au LMS est de rendre l'algorithme indépendant de la variance du signal d'entrée. Cependant, la distribution des valeurs propres λ_i de la matrice d'autocorrélation R du signal d'entrée n'est en rien modifiée. Ceci implique précisément la même dépendance, dans les deux cas, de la convergence vis-à-vis de la statistique du signal d'entrée [3].

Pour les signaux de parole, l'algorithme NLMS apporte une amélioration significative sur le taux de convergence par rapport au LMS grâce à la normalisation du pas d'adaptation [8].

1.5 Applications

Ce qui différencie essentiellement les applications provient de la façon de définir la réponse désirée $y(n)$. On peut distinguer quatre grandes classes d'applications:

- ❖ l'identification des systèmes,
- ❖ la prédiction,

- ❖ la modélisation inverse,
- ❖ l'annulation d'interférences,

1.6 Conclusion

Dans ce chapitre, nous avons présenté une vue d'ensemble de filtres adaptatifs, soulignant les applications et les algorithmes de base dont on a d'ores et déjà prouvé l'utilité en pratique.

Malgré les nombreuses contributions dans le domaine, les efforts de recherche pour améliorer les filtres adaptatifs continuent à une allure forte et il est probable que de nouvelles applications de ces filtres soient développées dans l'avenir. Pour suivre les progrès de ces avancés, le lecteur est pressé de consulter des journaux comme les Transactions IEEE sur le Traitement de Signal aussi bien que les actes de conférences annuelles et des ateliers dans le traitement de signal et les domaines liés.

Chapitre 2 Rehaussement du signal de parole par filtrage adaptatif

2.1 Introduction

Dans ce chapitre nous développons en détails les structures d'annulations du bruit adaptatives. En effet, en fonction du nombre de microphones utilisés, on peut distinguer plusieurs méthodes de réduction du bruit et de rehaussement du signal de parole. L'annulation adaptative du bruit (ANC) est la plus connue des méthodes de débruitage bi-capteurs.

2.2 La parole et le bruit

2.2.1 Définition de la parole

La parole apparaît physiquement comme une variation de la pression de l'air causée et émise par le système articulatoire. La phonétique acoustique étudie ce signal en le transformant dans un premier temps en un signal électrique grâce à un transducteur approprié : le microphone (lui-même associé à un préamplificateur). De nos jours, le signal électrique résultant est le plus souvent numérisé. Il peut alors être soumis à un ensemble de traitements statistiques qui visent à en mettre en évidence les traits acoustiques : sa fréquence fondamentale, son énergie, et son spectre. Chaque trait acoustique est lui-même intimement lié à une grandeur perceptuelle : pitch, intensité, et timbre [1][12][13].

2.2.2 Mécanisme de phonation

La parole est un signal complexe résultant de mouvements coordonnés de plusieurs muscles du corps humain [14]. On retrouve le larynx qui est un organe situé au niveau de la gorge, et les cordes vocales qui sont deux lèvres symétriques placées en travers

du larynx. Plusieurs types de sons peuvent être produits en fonction de l'action et du placement des muscles du corps humain et on peut d'ores et déjà décomposer le signal de parole en deux catégories : les sons voisés et les sons non-voisés.

La production des sons non-voisés résulte simplement d'un souffle turbulent filtré par le conduit vocal, alors que la production des sons voisés résulte elle d'une vibration périodique des cordes vocales filtrée par le conduit vocal. La fréquence de la vibration produisant les sons voisés est appelée « pitch » et varie en fonction du locuteur et du sexe de la personne. En effet, pour un homme, le *pitch* peut varier de 70 à 250 Hz tandis que pour une femme le *pitch* peut varier de 150 à 400 Hz [1].

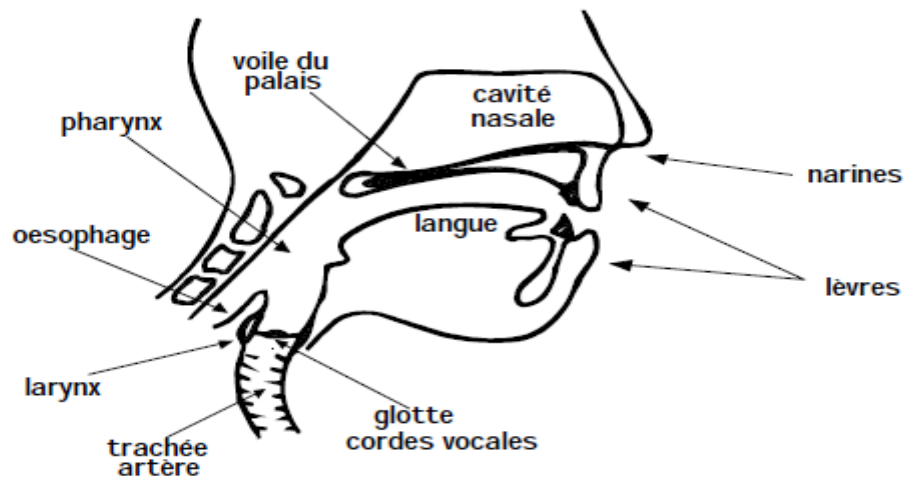


Figure 2.1. Représentation de l'appareil phonatoire humain [13].

2.2.3 Sons voisés et non-voisés

On constate que les caractéristiques du signal de parole varient suivant la région où l'on se place. Il est intéressant de constater que les sons voisés et non-voisés ont des propriétés statistiques différentes. En effet les sons voisés ont un caractère quasi-périodique, sont à forte énergie et sont fortement corrélés, alors que les sons non-voisés sont au contraire non périodiques, à faible énergie et peu corrélés. Par contre, le signal de parole est dans son ensemble considéré comme quasi-stationnaire, c'est-à-dire qu'il est stationnaire sur une courte période (de l'ordre de 20 à 40 ms) [1].

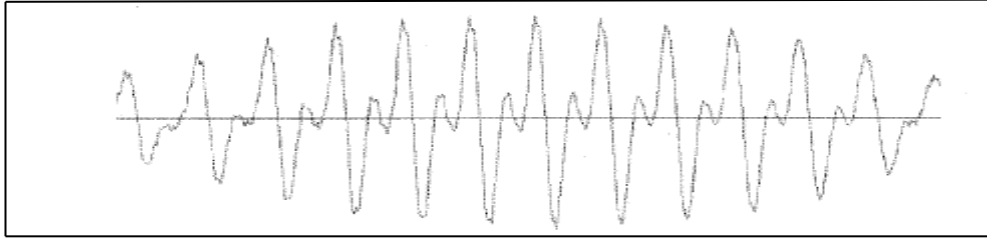


Figure 2.2. Exemple de son voisé [e] [1].

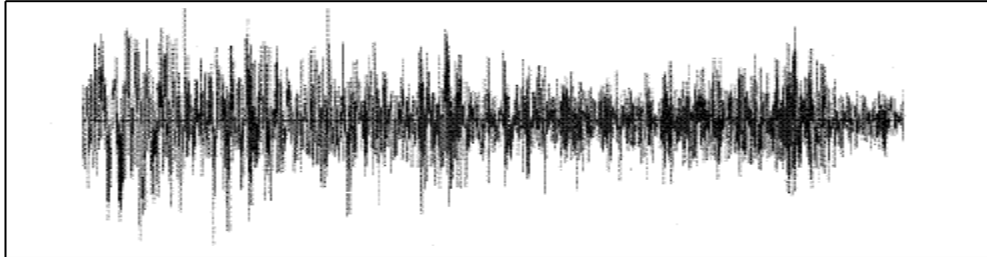


Figure 2.3. Exemple de son non-voisé [s] [1].

2.2.4 Modélisation autorégressive (AR) du signal de parole

Différentes techniques de rehaussement de la parole sont basées sur le modèle autorégressif AR et on peut d'ores et déjà cité comme exemple le filtrage de Wiener et le filtrage de Kalman. Afin de comprendre la relation entre ces différentes techniques et le modèle AR, il est nécessaire de savoir que la transformée de Fourier de l'autocorrélation $R_{ss}(n)$ est en fait égale à la densité spectrale de puissance (DSP) $P_{ss}(k)$ [1].

2.2.5 Bruit acoustique

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 ci-dessous, extrait de [12], représente les différentes classes auxquelles un bruit peut appartenir.

Propriétés	Types
Structure	Continu/Impulsif/Périodique
Type d'interaction	Additif/Multiplicatif/Convolutif
Comportement temporel	Stationnaire/Non-stationnaire
Bande de fréquence	étroit/large
Dépendance	Corrélé/ Décorrélé
Propriétés statistiques	Dépendant/Indépendant
Propriétés spatiales	Cohérent/Incohérent

Tableau 1.1. Les différentes classes de bruit [12].

2.2.6 Bruit blanc

Le bruit blanc doit son appellation à l'analogie avec la lumière blanche. Cette lumière blanche est due à la présence de photons de toutes les valeurs d'énergie, et est donc composée de toutes les couleurs. De même, un bruit blanc possède un spectre de fréquence continu et d'amplitude moyenne constante indépendante des fréquences.

$$\forall f, x(f) = x_0 \quad (2.1)$$

Le bruit blanc reste un modèle théorique et un tel bruit ne peut pas exister naturellement car il aurait sinon une puissance infinie :

$$P = \int_0^{\infty} x^2(f)df \quad (2.2)$$

A cause des limites de la perception, la définition du bruit blanc est restreinte à la bande de fréquence comprise entre 20Hz et 20000Hz [14].

2.2.7 Bruit coloré

Les bruits doivent aussi leur dénomination à l'analogie avec la lumière. Ces bruits n'ont pas un spectre d'amplitude constant. Au sens strict du terme, le terme bruit coloré est affecté à un bruit dont le spectre est caractérisé par une enveloppe très lisse avec une bande large, d'amplitude plus importante. Ainsi, la couleur correspondante

au spectre de la lumière associée est affectée au bruit. Par exemple, les bruits roses, bleus, violets, gris, . . . ont ainsi été définis [14].

2.2.8 Silences

Ce sont des intervalles où le signal utile est absent. En pratique il s'agit de bruits d'origines diverses, d'énergie négligeable devant celle du signal utile.

2.3 Principe de la réduction de bruit

Dans beaucoup de situations, le bruit est de nature aléatoire et ses composantes fréquentielles sont présentes à travers toute la bande passante du signal de parole. Ce signal peut être considéré comme quasi-stationnaire sur des périodes de l'ordre de la dizaine de milliseconde (typiquement de 20 à 40 ms).

Dans de nombreux environnements, le bruit est additif et le signal de parole bruité $x(n)$ peut alors s'écrire comme suit [1]:

$$x(n) = s(n) + n(n) \quad (2.3)$$

Où $s(n)$ représente le signal de parole propre et $n(n)$ le bruit additif. Les méthodes de réduction de bruit utilisent souvent une approche spectrale, de manière à utiliser l'information fréquentielle du signal pour filtrer le bruit gênant. Le principe en est illustré à la figure suivante [1] :

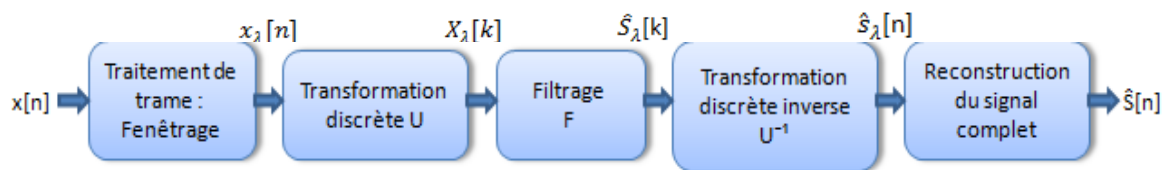


Figure 2.4. Schéma explicatif du principe de la réduction de bruit dans le domaine fréquentiel [1].

2.3.1 Estimation du bruit

On a pu voir dans ce qui précède que les techniques classiques de réduction du bruit s'appuient sur la DSP du signal de parole et sur celle du bruit, notés respectivement $P_{ss}(k)$ et $P_{nn}(k)$. Par le biais de la modélisation autorégressive AR, nous avons pu déterminer $P_{ss}(k)$ mais il reste à déterminer $P_{nn}(k)$.

Les techniques classiques d'estimation du bruit opèrent lors des temps de silence du signal. Afin de discriminer les temps de silence des temps d'activité vocale, les techniques d'estimation du bruit utilisent généralement un détecteur d'activité vocale, ou VAD. Cet outil permet de détecter la présence ou non de la voix dans un signal à un temps donné.

Un VAD opère généralement sur des trames de l'ordre de la dizaine de milliseconde et a pour principe d'extraire la puissance d'un signal d'entrée et de la comparer à un seuil déterminé à partir du bruit ambiant. La présence de parole est reconnue si les valeurs mesurées sont au-dessus du seuil d'activité de parole. Plusieurs méthodes ont été proposées pour détecter la voix. Ces méthodes se basent soit sur la distance LPC, soit sur les niveaux d'énergie ou encore la distance cepstrale dont le principe est un seuillage des mesures de distances cepstrale pour discriminer les temps de parole des temps de silence.

Les techniques d'estimation du bruit permettent de déterminer $P_{nn}[k]$ pendant les temps de pause dans la parole ou seul le bruit perturbant intervient. Le principe est illustré à la figure suivante [1] :

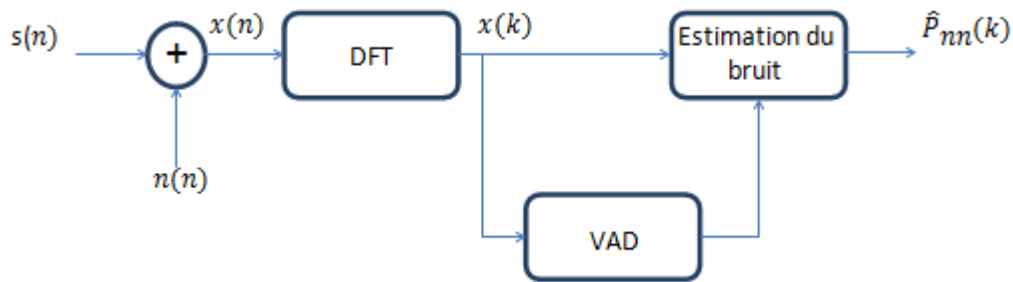


Figure 2.5. Principe de l'estimation du bruit [1].

La méthode d'estimation du bruit présentée ici est une méthode classique qui se base sur les statistiques minimums du spectre du bruit.

Des approches permettant de poursuivre les non-stationnarités du bruit en continu, *i.e.* indifféremment pendant l'inactivité ou l'activité vocale, ont été développées.

2.3.2 Les types de mélanges

Pour le type de mélange, on trouve dans la littérature plusieurs types. Une première catégorie, largement étudiée, concerne les mélanges "linéaires", c'est-à-dire le cas où les observations sont des mélanges linéaires des signaux sources. Cette catégorie est découpée en trois sous-catégories [17]:

a Mélanges linéaires instantanés

Ce sont des mélanges dans lesquels les valeurs des observations à un instant (lorsqu'il s'agit de signaux temporels monodimensionnels) ou à une position (lorsqu'il s'agit de signaux spatiaux bidimensionnels), ou à une longueur d'onde (lorsqu'il s'agit de spectres), s'écrivent comme des combinaisons linéaires des valeurs des signaux sources au même instant ou à la même position, ou à la même longueur d'onde. Dans ce cas, la relation liant les sources et les observations, s'écrit, pour des signaux temporels, sous la forme suivante :

$$x(n) = A s(n) \quad (2.4)$$

Où A est une matrice constante inconnue dite "matrice de mélange". Cette dernière contient des coefficients constants dits "coefficients de mélange". On trouve pratiquement ce genre de mélange dans les applications où les temps de propagation, ou les décalages spatiaux, sont négligeables.

b Mélanges linéaires à atténuations et retards

Dans lesquels les temps de propagation, ou les décalages spatiaux, sont pris en considération.

c Mélanges linéaires convolutifs

Dans lesquels les capteurs fournissent des superpositions de versions filtrées des signaux sources reçus. Ce type de mélange est le plus général des trois présentés.

En effet, dans la plupart des applications, le mélange de sources est un mélange convolutif. Dans ce cas, le problème consiste à trouver un filtre matriciel, opérant sur la suite des observations, tel que les différentes sources estimées en sortie du filtre séparateur soient aussi indépendantes que possible. Les premiers résultats ont été présentés dans le cas de deux sources et deux capteurs modélisé par deux filtres à réponse impulsionnelle finie (RIF). Les coefficients du filtre ont été estimés par une

généralisation du critère, qui était utilisé dans le cas du mélange instantané. On peut modéliser un mélange convolutif de la façon suivante [18] :

$$p_j(t) = \sum_{i=1}^N h_{ji}(t) * s_i(t) = \sum_{i=1}^N \sum_{k=1}^M h_{ji}(t-k) s_i(t) \quad (2.5)$$

Avec :

$h_{ji}(t)$ Représente la réponse impulsionnelle de la source i sur le capteur j

* Est l'opérateur de convolution

M Est l'ordre du filtre de mélange dont les coefficients sont h_{ji}

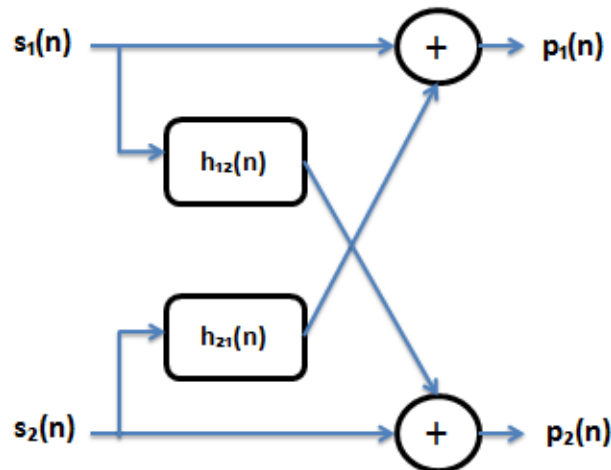


Figure 2.6. Principe du mélange avec deux sources et deux capteurs.

La seconde catégorie de mélange, concerne les mélanges non-linéaires, qui sont plus généraux. La complexité de cette dernière catégorie de mélanges limite son étude à des configurations particulières, comme par exemple le cas « post-non-linéaire » ou bien le cas d'un mélange linéaire-quadratique, ou encore le cas d'un mélange quadratique [17].

Les méthodes étudiées au cours de nos travaux ne concernent que les mélanges linéaires convolutifs.

2.4 Méthodes mono/bi/multi-capteurs pour la réduction du bruit et le rehaussement du signal de parole

Afin de mieux comprendre l'objectif général du sujet, il est nécessaire de connaître l'évolution des techniques de rehaussement de la parole. Ainsi, il est possible de différencier deux catégories de techniques : Les modèles multi-capteurs et les modèles mono-capteurs. Dans le cas des modèles multi-capteurs, les techniques utilisent soit un réseau de microphones soit deux microphones seulement. Dans le cas des modèles

mono-capteurs, un seul microphone est utilisé. Le présent sujet de recherche se situe dans les modèles bi-capteurs.

2.4.1 Méthodes d'annulation du bruit mono-capteurs

Un bon nombre d'algorithmes appliqués à des applications temps-réel comme la téléphonie se base sur l'estimation du bruit lors des temps de silence, qu'ils détectent via un détecteur d'activité vocale, ou DAV. Ces techniques réalisent ensuite un traitement de réduction de bruit sur ce bruit estimé. D'autres méthodes se basent sur les transformées de signaux à temps discret comme la transformée en cosinus discret (DCT) ou encore la transformée en ondelettes. Ces différentes approches permettent de profiter des caractéristiques spectrales des signaux bruités pour discriminer le bruit du signal propre. Les transformées discrètes se différencient principalement par leurs caractéristiques de stationnarité et de linéarité [1]. La transformée la plus répandue et utilisée en traitement de la parole est la Transformée de Fourier Discrète, ou DFT, qui est une transformée de signaux linéaire et stationnaire, au même titre que la Transformée en Cosinus Discret (DCT). La DFT permet de passer du domaine temporel au domaine fréquentiel. C'est la transformée la plus utilisée car elle est facilement implantable et le système auditif humain se rapproche d'une analyse de Fourier. Ceci étant, beaucoup d'algorithmes de réduction de bruit s'appuient sur cette transformée puisqu'elle permet une implémentation temps-réel.

a Réduction de bruit basée sur le modèle autorégressif

Les techniques qui utilisent le modèle autorégressif AR pour le rehaussement de la parole estiment les coefficients a_k à partir du signal bruité observé $x(n)$. Les deux approches généralement utilisées sont d'une part le filtrage de Wiener et d'autre part le filtrage de Kalman. Ces deux approches sont décrites ici [1]:

➤ Le filtrage de Wiener

Le filtre linéaire de Wiener tente de trouver une estimation au sens du minimum de l'erreur quadratique moyenne (MMSE) qu'on a d'ores et déjà détaillé dans le premier chapitre. Cette estimation minimise au mieux l'erreur quadratique moyenne (MSE) d'un signal $s(n)$ à partir d'un signal bruité $x[n]$. Le filtre peut s'écrire par la relation suivante :

$$F_w(k) = \frac{P_{ss}(k)}{P_{ss}(k) + P_{nn}(k)} \quad (2.6)$$

Ou $P_{ss}(k)$ et $P_{nn}(k)$ représentent de manière respective la densité spectrale de puissance, ou DSP, du signal de parole $s[n]$ et du bruit $n(n)$.

➤ **Le filtrage de Kalman**

Il a été vu que les coefficients a_k trouvés lors de l'analyse LPC permettaient d'une part de modéliser le signal de parole lors du modèle AR, mais étaient également très utiles au sein du filtrage de Wiener afin de déterminer la DSP du signal de parole $P_{ss}(k)$. Malheureusement, le filtrage de Wiener souffre d'un certain nombre de problèmes dont :

- ❖ L'inefficacité de l'algorithme en présence de signaux non stationnaires.
- ❖ Le nombre trop important de calculs mathématiques dont il dépend.

En 1960 filtre de Kalman est proposé afin de pallier ces problèmes [19]. De la même manière que le filtrage de Wiener, le filtre de Kalman résout de façon optimale l'estimation linéaire au sens du MMSE d'un signal de parole $s(n)$ à partir d'une observation $x(n)$. La méthode consiste à alterner entre la prédiction et la correction des coefficients de prédiction afin d'estimer au mieux le signal $s(n)$.

b Approche par soustraction spectrale

Une approche de réduction de bruit par soustraction spectrale permet d'estimer et de réduire le bruit qui corrompt les signaux de parole. Cette technique suppose que le bruit est un procédé stationnaire dans un temps court et qu'il est non-corrélé au signal de parole. On suppose que l'oreille humaine est insensible à la phase du signal et l'on considère que la phase du signal de parole est la même que celle du signal bruité [20]. La méthode générale de soustraction spectrale utilise la DFT pour travailler dans le domaine fréquentiel. Une règle de soustraction spectrale est réalisée afin de séparer le bruit du signal de parole. Cette règle peut s'exprimer par la relation suivante :

$$S(k) = F(k)X(k) \quad (2.7)$$

La figure suivante illustre le principe de soustraction spectrale mis en place par Boll :

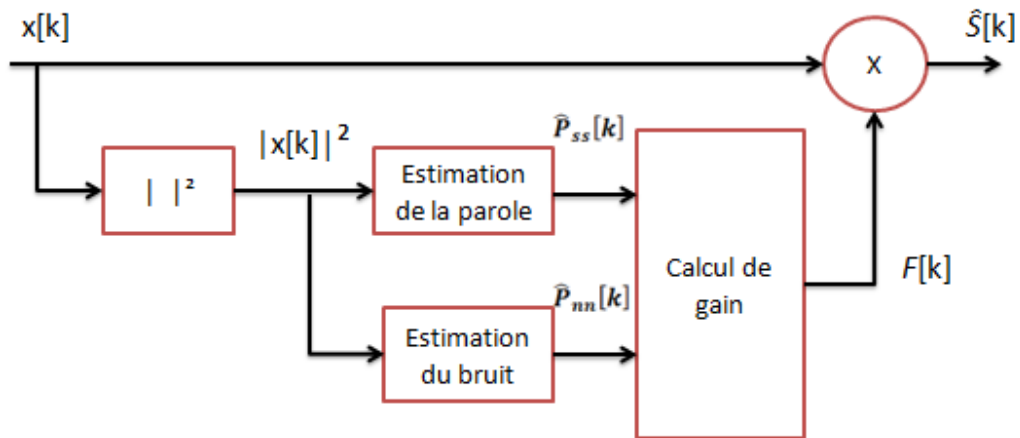


Figure 2.7. Principe de l'approche par soustraction spectrale [1].

2.4.2 Méthodes de débruitage bi-capteurs

Les réseaux de microphones sont utilisés dans le but d'exploiter les caractéristiques de l'espace pour cibler la source sonore que l'on veut traiter. Ils sont généralement utilisés pour procéder à de la formation de voies, qui correspond à un filtrage spatial opérant sur les sorties du réseau, de manière à rehausser le signal sonore corrompu par du bruit et des interférences directionnelles. On peut différencier deux catégories de formation de voies :

Les formations de voies conventionnelles et les formations de voies adaptatives. Les premières sont des techniques simples et procèdent généralement par la méthode de l'antenne retard-somme. Les sorties de chaque source sont pondérées et retardées (déphasées) de manière différente, afin de combiner les signaux provenant de ces sources. Seule l'information spatiale de l'emplacement des sources ainsi que la direction du signal à traiter sont utilisées pour rehausser le signal sonore. Les formations de voies adaptatives combinent cette information avec des propriétés du signal à traiter reçu par le réseau, et ce de manière à augmenter la rejection du bruit provenant des autres directions.

a Séparation aveugle de source

Dans le domaine audio, la séparation aveugle de source est surnommée « Cocktail Party ». C'est une discipline plus ou moins jeune, les premiers travaux sont attribués à Héroult, Jutten et Ans au milieu des années 80.

La séparation de sources consiste à restaurer L signaux sources inconnus s_j à partir de k observations x_i .

Ces observations sont le résultat de mélanges des signaux sources et proviennent généralement des capteurs, tels que les microphones, des antennes radiofréquences, ... Le mélange entre les différents signaux sources est dû à la propagation simultanée de ces derniers jusqu'aux capteurs.

La relation liant les sources et observations est donnée par [17] :

$$x = F(s) \quad (2.8)$$

Où $s = [s_1 \dots s_L]^T$ et $x = [x_1 \dots x_k]^T$ sont deux vecteurs colonnes représentant respectivement l'ensemble des signaux sources et l'ensemble des signaux d'observations. Et $F(s)$ dénote l'opérateur de mélange. L'originalité de la SAS réside dans le fait qu'on se place dans un contexte « AVEUGLE », c'est-à-dire que l'opérateur de mélange ainsi que les signaux sources sont inconnus.

L'objectif de cette technique est de fabriquer un système capable de fournir des sorties y_j égales aux signaux sources s_j . Ainsi, la reconstitution des signaux sources est réalisée grâce au modèle :

$$y = G(x) \quad (2.9)$$

Où $y = [y_1 \dots y_L]^T$ est un vecteur colonne représentant l'ensemble de signaux de sortie et $G(.)$ est l'opérateur inverse de $F(.)$. Il apparaît ici deux problèmes qui peuvent être différents. En effet, on peut considérer le problème de l'estimation du modèle de mélange, que l'on appelle problème d'identification Aveugle de Mélanges (IAM) (Blind Mixture Identification BMI). Ou bien, on peut considérer le problème appelé restitution des sources [21] consistant à proposer une estimation de sources sans pour autant proposer une estimation de la structure de mélange. Bien entendu, ces deux problèmes coïncident lorsque le mélange est inversible.

La configuration générale de la séparation aveugle de sources est donnée dans la figure suivante :

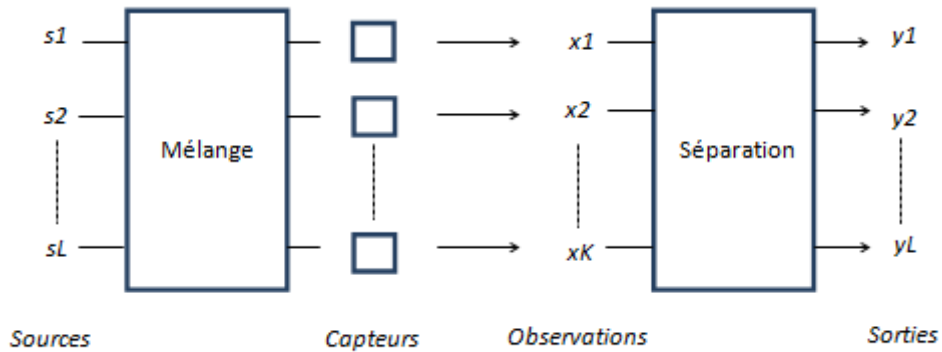


Figure 2.8. Configuration générale de la séparation aveugle de source [17].

La SAS est une « discipline générique » qui permet de nombreuses applications telles que la biologie, et le génie biomédical, la chimie, et l’astrophysique d’une manière générale.

La SAS possède plusieurs degrés de difficultés. Plusieurs paramètres sont à prendre en considération, plus particulièrement, les deux points théoriques qui sont le type de mélange et le nombre K par rapport au nombre L de signaux sources.

Concernant le nombre K d’observations par rapport au nombre L de signaux sources, on conçoit que le cas « (sur)détermine », c’est-à-dire $K=L$ ($K>L$), est plus simple à résoudre que le cas « sous-déterminé », c’est-à-dire $K<L$. Ce dernier cas ne peut être résolu qu’à l’aide d’importante informations a priori [17].

Dans le cas de mélange linéaire convolutif, les principaux outils nécessaires à la séparation aveugle de sources sont principalement des méthodes utilisant la décorrélation, et les statistiques d’ordres supérieurs [13].

b Annulation adaptative du bruit

Le principe d’annulation adaptative du bruit remonte aux contributions de Howells et Applebaum vers la fin des années 50. Le développement de l’algorithme LMS (Least Mean Square) par Widrow et Hoff a permis l’élaboration de cette technique sous sa forme actuelle. Nous présentons ici le principe de base de cette structure.

➤ Principe de base de l’ANC

Considérons une observation bruitée $d(n)$ d’un signal $s(n)$ c’est-à-dire :

$$d(n) = s(n) + b_0(n) \quad (2.10)$$

ou $b_0(n)$ est une contribution décorrélée du signal utile $s(n)$. $d(n)$ représente le signal d'observation sur la voie primaire. Supposons que l'on dispose au niveau d'un deuxième capteur d'un bruit $b(n)$ fortement corrélé à $b_0(n)$ mais décorrélée de $s(n)$. $b(n)$ fournit un signal secondaire d'observation sur la voie de référence au système ANC. La relation entre $b_0(n)$ et $b(n)$ est supposé linéaire, c'est-à-dire :

$$b_0(n) = h(n) * b(n) \quad (2.11)$$

ou $(*)$ représente le produit de convolution et $h(n)$ représente la réponse impulsionnelle du canal qui sépare la voie primaire de la voie secondaire, avec $H(f)$ comme fonction de transfert de ce canal.

Le principe de la technique d'annulation adaptative de bruit (Adaptive Noise Cancellation ou ANC) est de filtrer la référence $b(n)$ de manière adéquate afin d'obtenir la meilleure estimation de $b_0(n)$ (au sens de l'erreur quadratique moyenne) ce qui apporte une réduction de bruit en sortie [22]. Une illustration graphique de ce principe est présentée sur la figure ci-dessous :

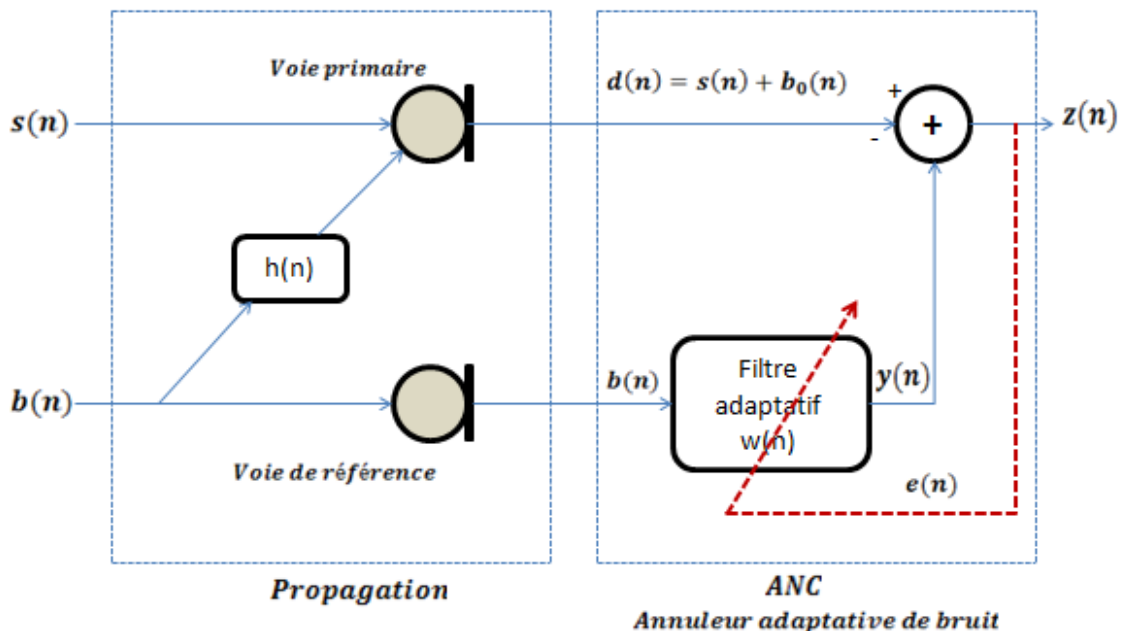


Figure 2.9. Structure d'annulation adaptative du bruit.

En tenant compte de la relation linéaire entre les bruits dans les deux observations, le signal de sortie peut s'écrire sous la forme :

$$z(n) = s(n) + (h(n) - w(n)) * b(n) \quad (2.12)$$

La solution d'un débruitage idéal $z(n) = s(n)$ est donc équivalente à $w_{opt}(n) = h(n)$. Une manière de déterminer le filtre $w_{opt}(n)$ sans connaissance ni du signal de parole $s(n)$ ni du filtre $h(n)$ consiste à considérer la solution de la minimisation de l'énergie du signal estimé $z(n)$. En effet, compte tenu de la décorrélation entre $b_0(n)$ et $s(n)$, on a d'après l'équation (2.12) :

$$E [z^2(n)] = E [s^2(n)] + E [((h(n) - w(n)) * b(n))^2] \quad (2.13)$$

La valeur minimale de $E [z^2(n)]$ est alors atteinte pour $w_{opt}(n) = h(n)$. La solution de ce problème correspond au filtre de Wiener.

2.4.3 Méthodes de débruitage multi-capteurs

Dans ce cas précis, on suppose que l'on dispose d'une multitude de microphones. Cette hypothèse autorise la duplication des techniques mono-voies et bi-voies précédemment présentés mais elle permet surtout d'envisager la combinaison des filtrages temporel et spatial par les méthodes de traitement d'antenne linéaire. On cite la méthode de réduction de bruit par le réseau d'antennes de Frost, et la structure de Griffiths et Jim.

2.5 Conclusion

Au cours de ce chapitre, nous avons introduit la notion de rehaussement de la parole et celle de la réduction du bruit, ainsi nous avons illustrés les différentes techniques utilisées à cet effet, des techniques qu'on peut différencier par le nombre de microphones utilisés.

On a ainsi présenté la structure d'annulation adaptative du bruit comme appartenant aux méthodes dites bi-capteurs. Ces dernières nous donnent des résultats optimaux.

Chapitre 3 Proposition d'un nouvel algorithme de décorrelation fréquentiel symétrique

3.1 Introduction

La technique de séparation de sources a suscité un intérêt croissant depuis son apparition au milieu des années 80. Cette technique est très vite devenue un domaine attractif du traitement du signal. Le problème consiste à retrouver des signaux utiles (par exemple : les signaux de parole ou des signaux émis par des téléphones portables), aussi appelés sources, à partir de mélanges de nature instantanée ou convolutif, appelés observations.

Ce chapitre est consacré à la présentation de notre nouvel algorithme de séparation de source aveugle de type récursive par une implémentation fréquentiel.

3.2 Séparation aveugle de source (SAS)

Dans un contexte aveugle, la séparation de sources est faite sans aucune connaissance a priori sur les sources et sur le processus de mélange (contexte d'observation), cette situation est alors appelée séparation aveugle de sources (SAS) (en anglais Blind Source Séparation (BSS)).

Pour résoudre ce problème, une solution possible consiste à supposer que les sources sont mutuellement indépendantes statistiquement.

On retrouve la SAS dans plusieurs applications tel que : les radiocommunications, traitement des images, rehaussement de la parole, débruitage, etc...

3.3 Solution de la séparation de sources

Pour résoudre ce problème, deux structures de séparation de sources bien connue sont classiquement utilisées (Directe et Récursive) [3]:

3.3.1 Structure directe (Forward)

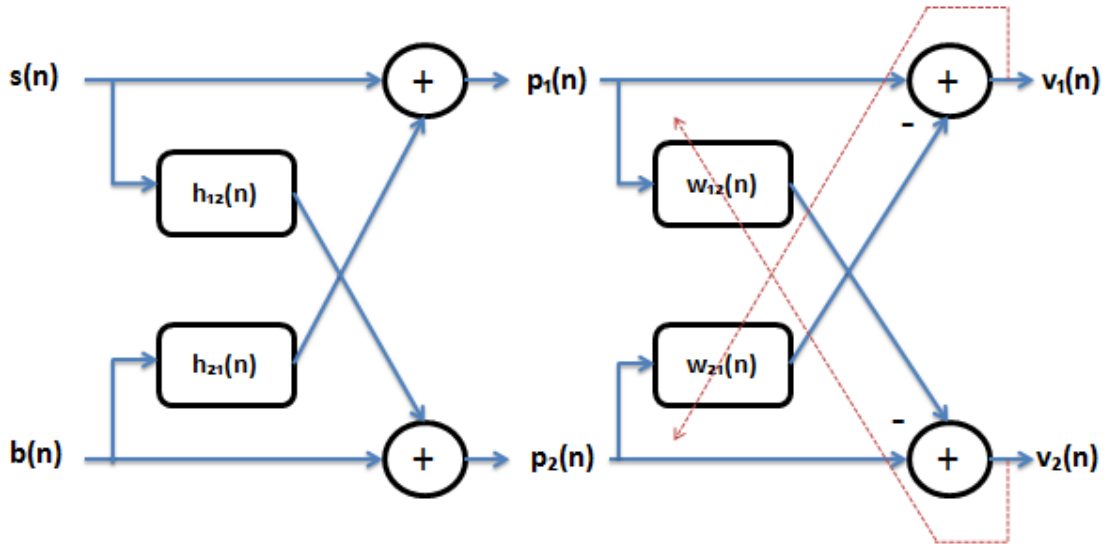


Figure 3.1. Structure de séparation de sources directe(Forward).

Lorsque chaque sortie du système de séparation est une simple combinaison linéaire des observations on parle de la structure directe connue sous le terme anglo-saxon Forward.

Le schéma de cette dernière est illustré par la figure qui suit :

Pour ce modèle, les signaux de sorties $v_1(n)$ et $v_2(n)$ sont données par les formules suivantes [24]:

$$v_1(n) = p_1(n) - p_2(n) * w_{21}(n) \tag{3.1}$$

$$v_2(n) = p_2(n) - p_1(n) * w_{12}(n) \tag{3.2}$$

Ces relations s'écrivent en fonction des sources $(s(n)$ et $b(n))$:

$$v_1(n) = b(n) * [h_{21}(n) - w_{21}(n)] + s(n) * [\delta(n) - h_{12}(n) * w_{21}(n)] \quad (3.3)$$

$$v_2(n) = s(n) * [h_{12}(n) - w_{12}(n)] + b(n) * [\delta(n) - h_{12}(n) * w_{21}(n)] \quad (3.4)$$

Si $w_{21}(n) = h_{21}(n)$ et $w_{12}(n) = h_{12}(n)$

Alors

$$v_1(n) = s(n) * [\delta(n) - h_{12}(n) * w_{21}(n)] \quad (3.5)$$

$$v_2(n) = b(n) * [\delta(n) - h_{21}(n) * w_{12}(n)] \quad (3.6)$$

La mise à jour des coefficients des deux filtres adaptatifs $w_{21}(n)$ et $w_{12}(n)$ est donnée par les expressions suivantes [23] :

$$w_{12}(n) = w_{12}(n-1) + \mu_{12} \frac{v_2(n)p_1(n)}{p_1^T(n)p_1(n)} \quad (3.7)$$

$$w_{21}(n) = w_{21}(n-1) + \mu_{21} \frac{v_1(n)p_2(n)}{p_2^T(n)p_2(n)} \quad (3.8)$$

Ou

$$p_1(n) = [v_1(n), v_1(n-1), \dots, v_1(n-L+1)]^T \quad (3.9)$$

$$p_2(n) = [v_2(n), v_2(n-1), \dots, v_2(n-L+1)]^T \quad (3.10)$$

3.3.2 Définition de la Backward

Malgré que la structure directe permet l'utilisation des filtres non causaux de manière naturelle vue qu'elle ne fait pas intervenir les différentes sorties du système dans le calcul des sorties des filtres de séparation (simplicité d'implémentation), reste que ces dernières sont distordues d'où le besoin d'utiliser des post-filtres.

La Backward est une structure mieux réfléchi car elle élimine le besoin d'utiliser des post-filtres cependant elle mène à obtenir de bons résultats du côté rapidité de convergence, cout en mémoire, et complexité algorithmique.

3.3.3 Structure recursive (Backward)

Dans l'algorithme de decorrélation temporel symétrique adaptative backward (BTD-SAD), chaque sortie du système de séparation est une combinaison linéaire d'une observation $p_i(n)$ et de l'autre sortie $u_j(n)$, cette structure est représentée par la figure qui suit :

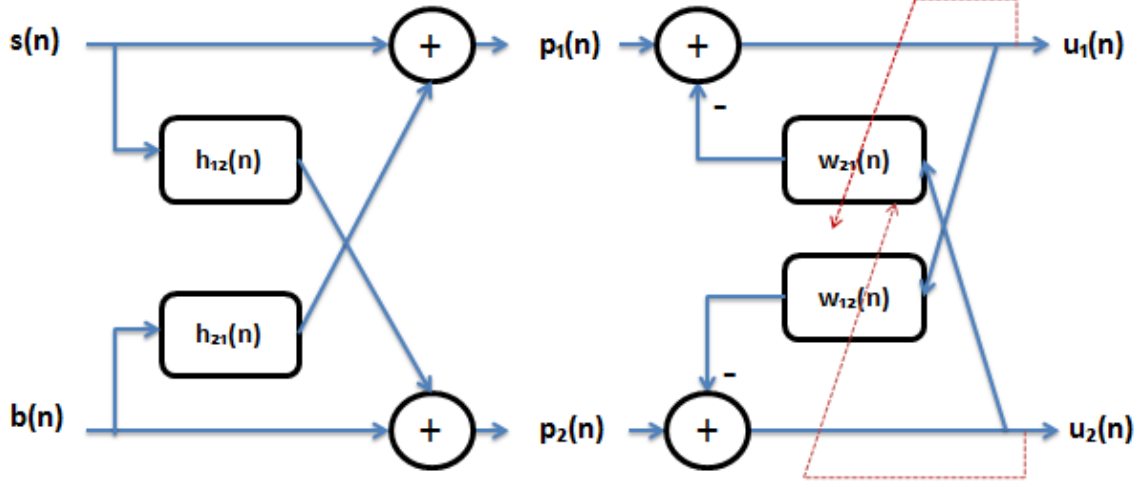


Figure 3.2. Structure de séparation de sources réursive (**Backward**).

Les signaux de sortie de mélange convolutif ($p_1(n)$ et $p_2(n)$), qui seront utilisés comme des entrées de la structure de séparation de sources Réursive de la figure (3.2), sont données par :

$$p_1(n) = s(n) + h_{21}(n) * b(n) \quad (3.11)$$

$$p_2(n) = b(n) + h_{12}(n) * s(n) \quad (3.12)$$

La structure Réursive conduit aux relations observations/sorties suivantes :

$$u_1(n) = p_1(n) - u_2(n) * w_{21}(n) \quad (3.13)$$

$$u_2(n) = p_2(n) - u_1(n) * w_{12}(n) \quad (3.14)$$

En remplaçant les expressions de $p_1(n)$ et $u_2(n)$ dans l'expression de $u_1(n)$, nous obtenons :

$$u_1(n) = s(n) + h_{21}(n) * b(n) - [p_2(n) - u_1(n) * w_{12}(n)] * w_{21}(n) \quad (3.15)$$

En remplaçant $p_2(n)$ par son expression dans $u_1(n)$, on obtient :

$$u_1(n) = \frac{1}{\delta(n) - w_{12}(n) * w_{21}(n)} \left(s(n) * (\delta(n) - h_{12}(n) * w_{21}(n)) + b(n) * (h_{21}(n) - w_{21}(n)) \right) \quad (3.16)$$

Et de même pour $u_2(n)$, en remplaçant les expressions de $p_2(n)$ et $u_1(n)$ ensuite de $p_1(n)$ dans l'expression de $u_2(n)$, nous obtenons :

$$u_2(n) = \frac{1}{\delta(n) - w_{12}(n) * w_{21}(n)} \left(b(n) * (\delta(n) - h_{21}(n) * w_{12}(n)) + s(n) * (h_{12}(n) - w_{12}(n)) \right) \quad (3.17)$$

Pour l'obtention du signal de parole sur la sortie $u_1(n)$ et le bruit sur la sortie $u_2(n)$, une hypothèse d'optimalité est utilisée pour les deux filtres adaptatifs

$$(w_{21}^{opt} = h_{21}(n) \text{ et } w_{12}^{opt} = h_{12}(n))$$

Dans ce cas, les sorties $u_1(n)$ et $u_2(n)$ deviennent :

$$u_1(n) = s(n) \quad (3.18)$$

$$u_2(n) = b(n) \quad (3.19)$$

Nous notons que les coefficients des deux filtres de séparation $w_{21}(n)$ et $w_{12}(n)$ sont adapté à partir de l'algorithme BTD-SAD, et les solutions optimales de ses filtres sont obtenues en utilisant le critère du minimum de l'erreur quadratique moyenne (MMSE ou bien EQMM)

L'adaptation des deux filtres $w_{21}(n)$ et $w_{12}(n)$ est donnée par les équations suivantes :

$$w_{12}(n) = w_{12}(n-1) + \mu_{12} \frac{u_2(n)p_1(n)}{p_1^T(n)p_1(n)} \quad (3.20)$$

$$w_{21}(n) = w_{21}(n-1) + \mu_{21} \frac{u_1(n)p_2(n)}{p_2^T(n)p_2(n)} \quad (3.21)$$

Ou

$$p_1(n) = [u_1(n), u_1(n-1), \dots, u_1(n-L+1)]^T \quad (3.22)$$

$$p_2(n) = [u_2(n), u_2(n-1), \dots, u_2(n-L+1)]^T \quad (3.23)$$

μ_{12} et μ_{21} sont respectivement les pas d'adaptations des deux filtres adaptatifs $w_{12}(n)$ et $w_{21}(n)$.

3.4 Recouvrement et addition (OLA)

La technique OLA consiste à conserver les échantillons filtrés partiellement (filtre tronqué) de la trame, puis à les combiner avec ceux issus des trames précédentes ou suivantes (selon que l'on considère le début ou la fin de la trame) afin d'obtenir des échantillons filtrés avec un filtre complet. Cette technique ne souffre pas des problèmes de "clics" que l'on rencontre avec la technique OLS car la transition entre les filtres des trames successives est assurée du fait de l'interpolation implicite d'une trame à l'autre.

Comme pour la technique OLS, l'implémentation peut se faire dans le domaine fréquentiel ou dans le domaine temporel. Au même titre que l'OLS, à l'origine l'OLA s'implémente dans le domaine fréquentiel (application de convolution rapide par FFT). Cependant son utilisation pour la réduction de bruit nécessite d'élargir le cadre classique [23].

3.5 Proposition d'un nouvel algorithme de séparation de source aveugle (BFD-SAD)

Dans ce paragraphe nous allons vous présenter notre nouvel algorithme fréquentiel pour la séparation de source aveugle de type régressive (Backward).

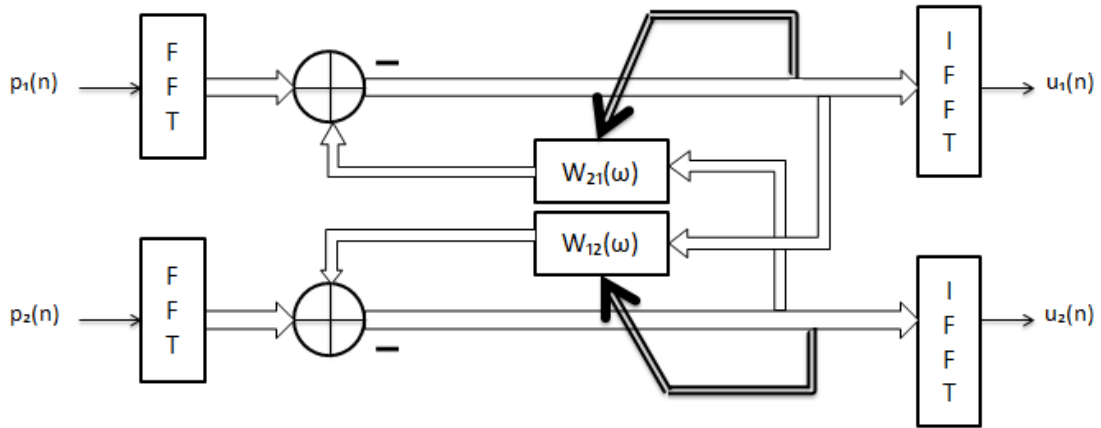


Figure 3.3. Nouvel algorithme fréquentiel proposé BFD-SAD.

La figure (3.3) nous montre le passage du domaine temporel au domaine fréquentiel ; Soit la transformation des deux observations bruité $p_1(n)$ et $p_2(n)$ par la transformée de Fourier rapide et par la suite l'obtention des observations fréquentielles $P_1(w, k)$ et $P_2(w, k)$.

Les résultats obtenus après la transformé de Fourier rapide sont caractérisés par les deux erreurs de filtrages $U_1(w, k)$ et $U_2(w, k)$ qui sont données par les relations suivantes :

$$U_1(w, k) = P_1(w, k) - W_{21}(w, k)P_2(w, k) \quad (3.24)$$

$$U_2(w, k) = P_2(w, k) - W_{12}(w, k)P_1(w, k) \quad (3.25)$$

Comme nous l'avons déjà cité au début de ce chapitre, cette technique est implémentée avec la structure Backward .Le nouvel algorithme fréquentiel proposé est basé sur l'implémentation fréquentielle des filtres adaptatifs $w_{12}(n)$ et $w_{21}(n)$ qui sont mis à jour par un algorithme adaptatif (BFD-SAD) sur une base trame par trame.

Pour chaque trame k , nous diffusons l'équation suivante :

$$W_{12}(w, k) = W_{12}(w, k - 1) + \mu_{12}(w, k) \frac{U_2(w)P_1^H(w)}{LS_{P_1}(w, k)} \quad (3.26)$$

$$W_{21}(w, k) = W_{21}(w, k - 1) + \mu_{21}(w, k) \frac{U_1(w)P_2^H(w)}{LS_{P_2}(w, k)} \quad (3.27)$$

Avec :

- $\mu_{12}(w, k)$ et $\mu_{21}(w, k)$ sont deux pas d'adaptations qui contrôlent la convergence des deux filtres $W_{12}(w, k)$ et $W_{21}(w, k)$, respectivement.
- L est la longueur de la transformée de Fourier rapide (FFT).
- 'H' est l'opérateur Hermitien.
- $S_{P_1}(w)$ et $S_{P_2}(w)$ sont les deux densités spectrales de puissance (DSPs) pour les deux observations $P_1(w, k)$ et $P_2(w, k)$, respectivement.

Dans ce travail nous proposons d'utiliser un lissage d'estimation récursive de ces deux DSPs comme suit :

$$S_{P_1}(w, k) = \alpha_1 S_{P_1}(w, k) + (1 - \alpha_1) |P_1(w, k)|^2 \quad (3.28)$$

$$S_{P_2}(w, k) = \alpha_2 S_{P_2}(w, k) + (1 - \alpha_2) |P_2(w, k)|^2 \quad (3.29)$$

Un détecteur d'activité vocale (DAV) contrôle notre algorithme de decorrélation fréquentiel BFD-SAD. Afin d'obtenir sur la voie 1 la sortie du signal de parole rehaussé $U_1(w, k)$ et sur la voie 2 la sortie des composants du bruit $U_2(w, k)$. ce système contrôle, dans le domaine fréquentiel, la convergence des deux filtres adaptatifs $W_{12}(w, k)$ et $W_{21}(w, k)$.

$$W_{12}(w, k) = W_{12}(w, k - 1) + \mu_{12}(w, k) \varphi(w, k) \frac{U_2(w)P_1^H(w)}{LS_{P_1}(w, k)} \quad (3.30)$$

$$W_{21}(w, k) = W_{21}(w, k - 1) + \mu_{21}(w, k) \theta(w, k) \frac{U_1(w)P_2^H(w)}{LS_{P_2}(w, k)} \quad (3.31)$$

Ou $\varphi(w, k)$ est égale a :

“1” en périodes de présence du signal de parole.

“0” pendant l’absence de signal de parole.

Le paramètre $\theta(w, k)$ est sélectionné par la relation suivante :

$$\theta(w, k) = 1 - \varphi(w, k)$$

Les relations précédentes montrent clairement que les filtres adaptatifs $W_{12}(w, k)$ et $W_{21}(w, k)$ sont mis à jour alternativement, pour permettre le débruitage du signal de parole que l’on obtient en sortie 1.

Une fois le filtrage réalisé, on applique une FFT inverse sur les trames du signal, de manière à revenir dans le domaine temporel.

Enfin, l’opération de recouvrement réalisée lors de la découpe du signal en trames, nécessite, lors de la reconstitution du signal après traitement de ce dernier, d’appliquer la méthode de overlap-add, ou OLA, afin de reconstruire parfaitement le signal débruité $\tilde{u}_1(n)$.

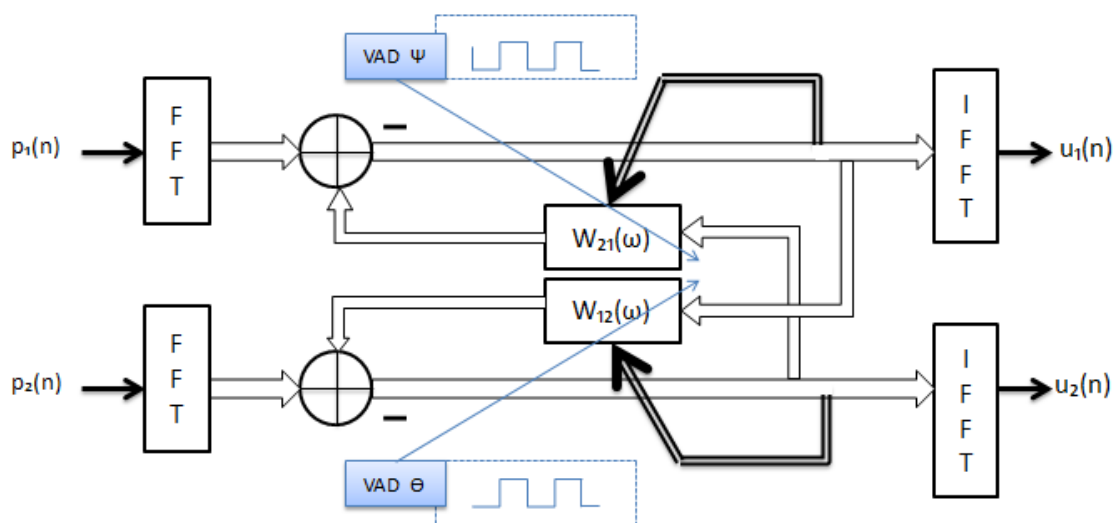


Figure 3.4. Algorithme proposé (BFD-SAD) contrôlé par un système de détection manuel d’activité vocal (DAV)

3.6 Conclusion

Au cours de ce chapitre, nous avons présenté le coté théorique de notre nouvel algorithme de décorrelation symétrique fréquentiel (BFD-SAD) pour la réduction de bruit et la séparation aveugle de sources.

Dans le chapitre qui suit, nous présentons les résultats de simulations de notre nouvel algorithme proposé (BFD-SAD).

Chapitre 4 Résultats de simulations

4.1 Introduction

Nous présentons dans ce chapitre une validation expérimentale de la structure de séparation de sources de type réursive, sous forme de figures et de tableaux démonstratifs.

Ce dernier chapitre est présenté de la manière suivante :

Pour commencer, nous définissons le cadre expérimentale de nos simulations à travers notamment l'utilisation du mélange convolutif. Les différents signaux sources utilisés (parole, bruit) seront également définis.

En second lieu, nous présentons les résultats complets des simulations du comportement de notre nouvelle structure Backward, ainsi qu'une évaluation des effets de certains paramètres décisifs (L : la taille du filtre, SNR : rapport signal a bruit) sur celle-ci.

Pour finir, une comparaison entre notre nouvel algorithme BFD-SAD et sa version temporel BTD-SAD.

4.2 Modèle expérimental du mélange

Nous mettons en œuvre dans cette partie une implémentation spécifique du modèle de mélange de la figure (2.6), du chapitre 2. Ce modèle est rappelé sur la figure ci-dessous :

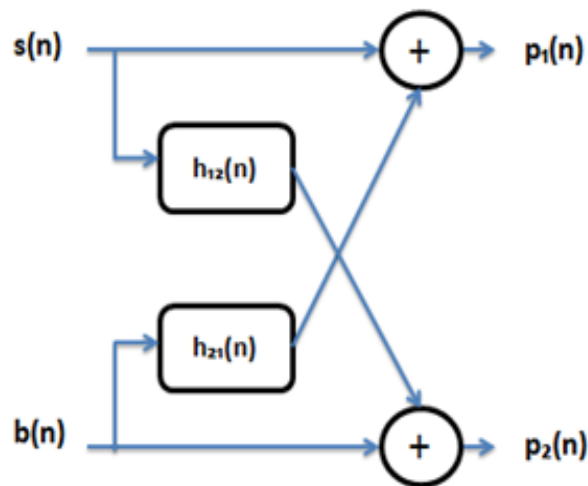


Figure 4.1. : Modèle de mélange convolutif.

Les réponses de couplage croisées sont h_{12} et h_{21} .

Notre idée vise à utiliser un modèle de mélange conforme à la physique du problème et capable de prendre convenablement en compte l'effet de la distance entre les deux microphones tout en restant conforme avec les fonctions de transfert supposées unitaires entre chaque source ponctuelle du signal et l'ensemble des chemins acoustiques directs.

4.3 Signaux de simulations utilisés

Pour générer les mélanges, nous avons utilisé un signal de parole échantillonné à 8kHz. Ainsi que deux sortes de bruits stationnaires.

Notre signal de parole dure 4 secondes, il est prononcé par un locuteur masculin « Un loup s'est jeté immédiatement sur la petite chèvre », qui est une phrase phonétiquement équilibrée, enregistrée dans une chambre ana-échoïque.

Les deux autres sources sont un bruit USASI et un bruit blanc :

-Le bruit USASI (United States of America Standard Institute) que nous avons sélectionné, correspond à un standard reconnu dans les tests acoustiques mais également en raison de sa densité spectrale de puissance, dont l'allure est relativement proche de celle de la parole réellement observé dans les véhicules.

-Le bruit blanc qui sert surtout à vérifier la stabilité de l'algorithme utilisé.

Vous allez trouver ci-dessous les représentations, dans le domaine temporel, de nos signaux sources.

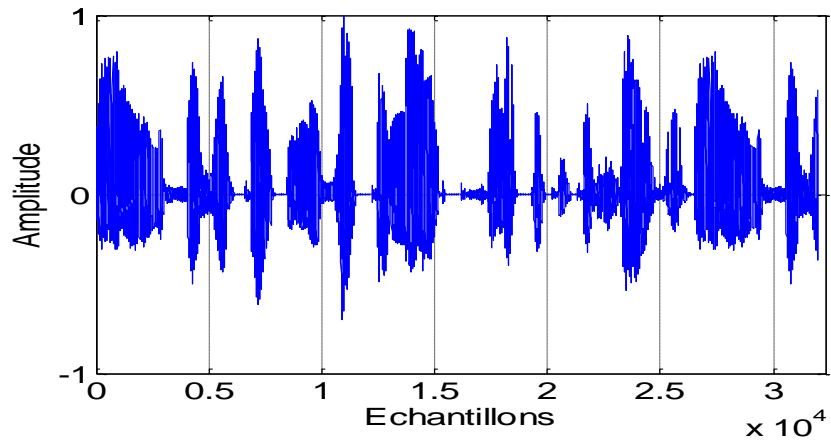


Figure 4.2.Signal de parole utilisé, $F_e = 8$ KHz.

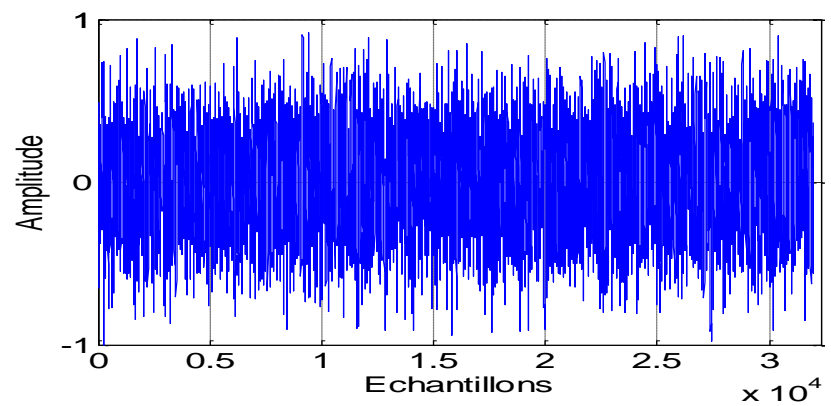


Figure 4.3.Signal du bruit USASI.

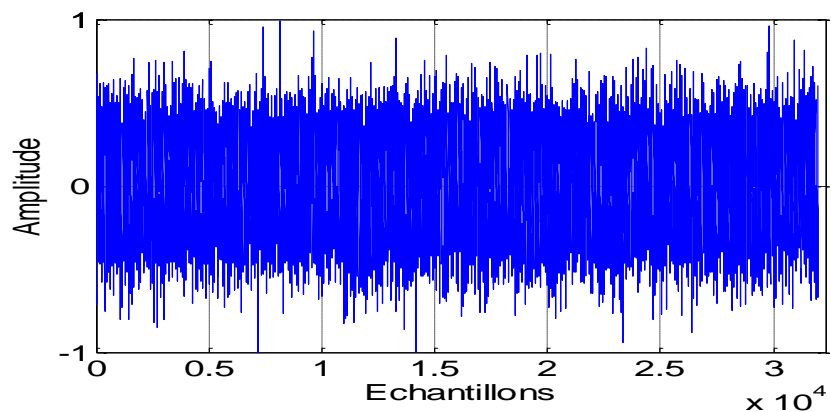


Figure 4.4.Signal du bruit blanc.

4.4 Simulations des réponses impulsionnelles

Les figures (4.5) et (4.6) représentent un exemple de réponses impulsionnelles acoustiques, générées avec une fonction exponentielle amortie selon [24]. La majeure partie de l'énergie contenue dans une réponse impulsionnelle correspond à un support d'environ $L=64$ échantillons à la fréquence d'échantillonnage de 8kHz. On rajoute à cela que la variance du bruit gaussien est égale à 0.5, ce qui correspond à une configuration de microphones relativement espacés.

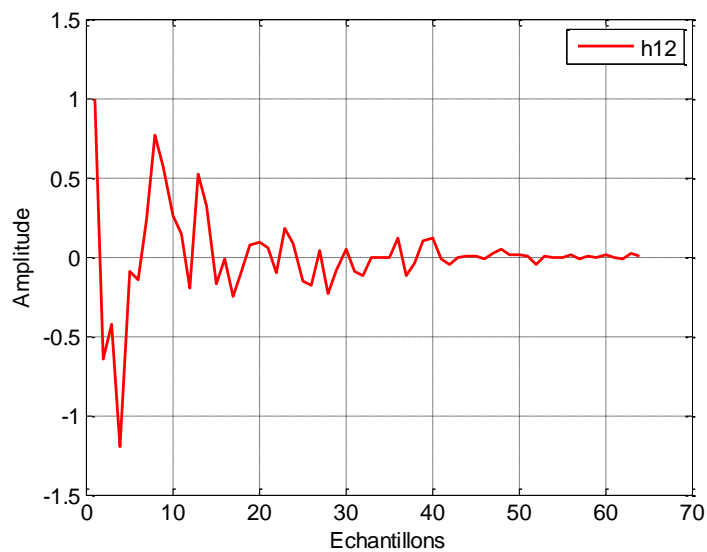


Figure 4.5. Exemple d'une réponse impulsionnelle simulée h_{12} .

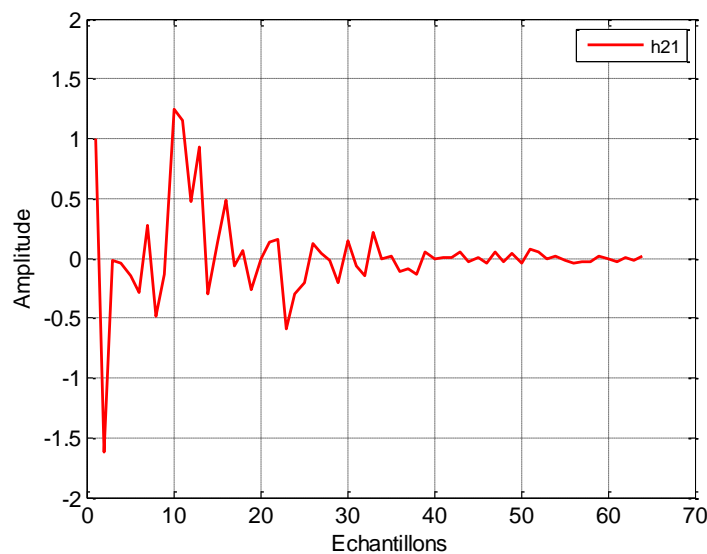


Figure 4.6. Exemple d'une réponse impulsionnelle simulée h_{21} .

4.5 Signaux de mélanges

Nous présentons dans cette partie les signaux de sortie de la structure de mélange convolutif. Les signaux source sont la parole et le bruit; Dans les deux figures ci-dessous, nous représentons nos deux signaux de mélange $p_1(n)$ et $p_2(n)$ avec un SNR égale à 3dB sur les deux voies et une longueur de filtre L, avec $L=32$.

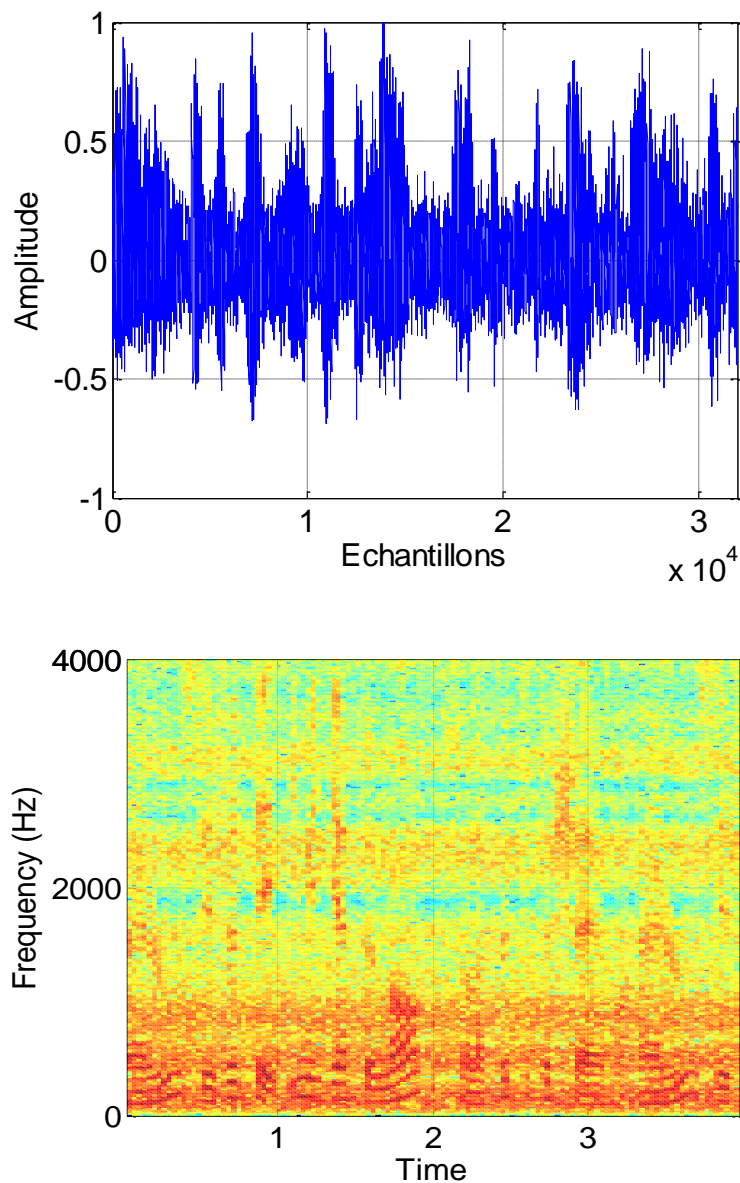


Figure 4.7.Signal de mélange $p_1(n)$ avec son spectrogramme.

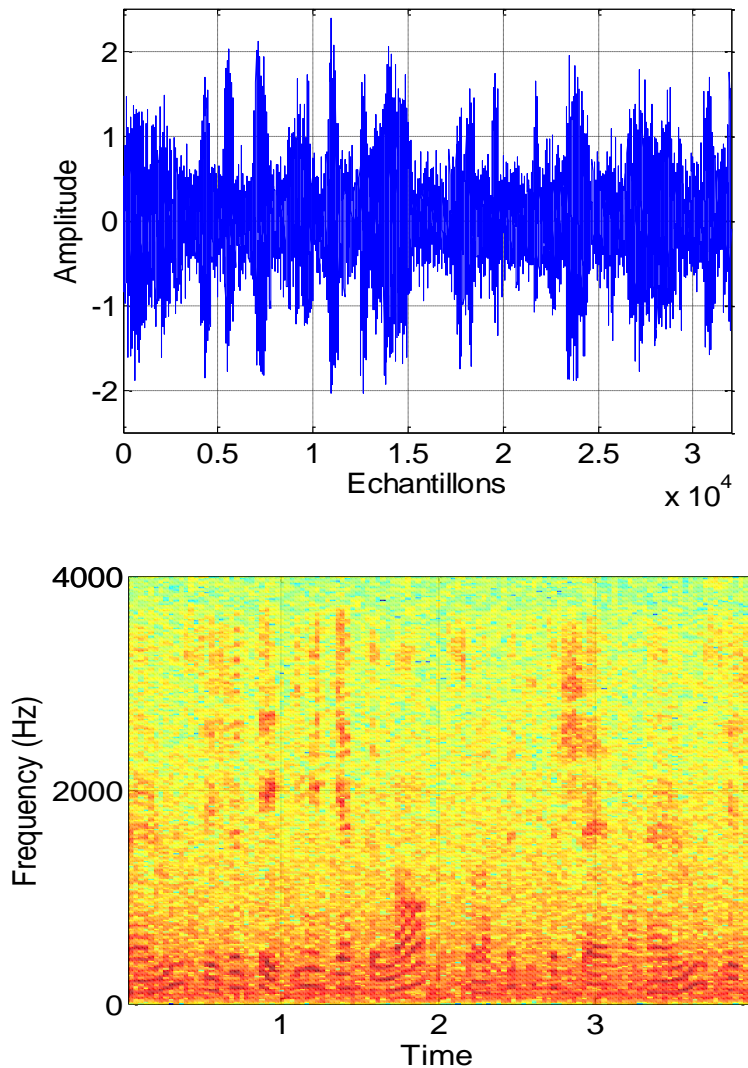


Figure 4.8.Signal de mélange $p_2(n)$ avec son spectrogramme.

➤ Définition du spectrogramme

Le spectrogramme représente dans une image à deux dimensions trois paramètres : le temps, la fréquence et la puissance sonore.

Pour y parvenir, on représente le temps sur un axe, le plus souvent horizontal, et la fréquence sur l'autre. Chaque point du diagramme correspond ainsi à un instant, d'une durée juste suffisante pour qu'on puisse analyser la fréquence du signal, et à une fréquence. La couleur du point représente la puissance selon une échelle conventionnelle. Les spectrogrammes sont utilisés en acoustique pour identifier des sons, comme des cris d'animaux et des sons d'instruments musicaux. Ils servent en phonétique à l'étude et à la comparaison des phonèmes.

4.6 Détection d'activité vocale

Dans ce cas de figure, la DSP du bruit est estimée à long terme et uniquement pendant l'inactivité vocale, c'est-à-dire qu'une DAV renseigne l'estimateur sur la présence ou non de parole pour la trame courante. Les périodes d'inactivité vocale doivent être suffisamment longues pour obtenir une estimée de la DSP du bruit possédant une faible variance ce qui ne pose pas de problème en pratique. Si la trame courante contient uniquement du bruit, alors l'estimée de la DSP du bruit est obtenue par un lissage exponentiel sinon, i.e. la trame courante contient de la parole, l'estimation de la DSP du bruit est figée.

Il a déjà été noté que les signaux en sortie de notre structure de séparation aveugle de sources de type récurive (Backward) sont obtenus à partir d'un système d'aiguillage (c'est-à-dire que, selon la configuration souhaitée par l'utilisateur, il est possible de maîtriser le fait que l'estimée d'un signal source soit disponible sur la sortie primaire ou secondaire de cette structure).

La structure de séparation de sources récurive utilise un mécanisme de détection d'activité vocale (DAV) manuelle ou automatique pour contrôler l'adaptation des filtres w_{12} et w_{21} . Cette adaptation est contrôlée de la manière suivante : le filtre w_{21} est adapté uniquement en période de bruit seul, alors que le filtre w_{12} est adapté uniquement pendant les périodes d'activité vocale (période de présence du signal de parole utile).

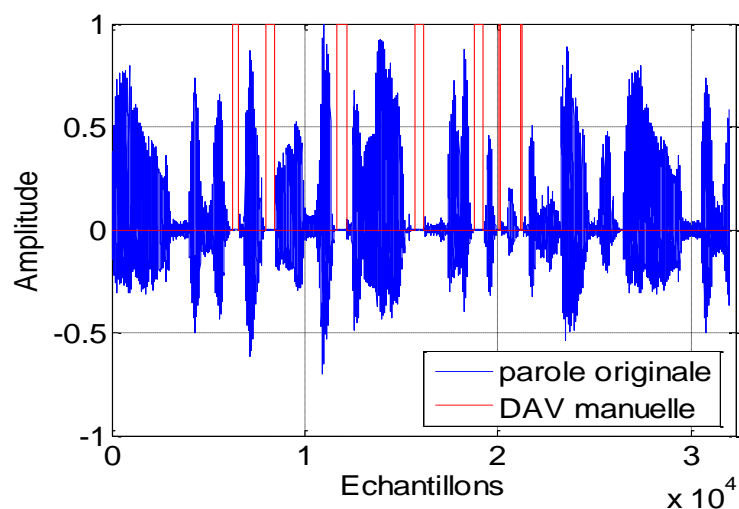


Figure 4.9. Signal de parole utilisé avec la segmentation (DAV) manuelle.

4.7 Critères objectifs de qualité

L'évaluation des performances des techniques de réduction du bruit représente une véritable problématique. En effet, ce sont des êtres humains qui sont amenés à juger de la qualité des traitements, jugement qui est donc essentiellement subjectif. Malheureusement l'évaluation de la qualité subjective est coûteuse en termes de temps et de ressources, il est par contre possible d'utiliser des mesures objectives qui sont beaucoup plus pratiques et moins coûteuses.

Les mesures objectives de qualité des signaux vocaux utilisées dans ce mémoire sont :

- Critère du rapport signal sur bruit (SNR).
- Critère de la distance cepstrale (CD).
- Critère de désajustement des coefficients (SM).

4.7.1 Le critère du rapport Signal a Bruit (SNR)

Le critère du rapport signal à bruit donne une information globale (pleine bande) du rapport entre la puissance du signal de parole rehaussé et celle du bruit, suivant [12] :

$$SNR = 10 \log_{10} \left\{ \frac{P_S(n)}{P_N(n)} \right\} \quad (4.1)$$

Où $P_S(n)$ et $P_N(n)$ sont les estimées courantes à court-terme de la puissance des composantes du signal parole et du bruit en cours d'étude.

- ♣ Le SNR (Signal to Noise Ratio) segmental segSNR est la mesure de la qualité objective la plus utilisée dans le domaine temporel. Il définit la moyenne des SNRs issus de plusieurs segments de courte durée (15 à 20 ms) [12]:

$$SNR_{seg} = \frac{10}{M} \sum_{m=0}^{M-1} \log_{10} \frac{\sum_{i=mN}^{mN+N-1} s^2(i)}{\sum_{i=mN}^{mN+N-1} (s(i) - \tilde{s}(i))^2} \quad (4.2)$$

Où $s(i), \tilde{s}(i)$, N et M sont respectivement le signal de référence, le signal débruité, la longueur d'un segment et le nombre total de segments [12].

Le SNR segmental souffre de deux limitations : d'abord si le signal de parole contient des segments de silence, ce qui est très probable, le $s(i)$ sera nul et n'importe quelle quantité de bruit entrainera un SNR en dB négatif pour ce segment; en effet le SNR total sera biaisé par cette quantité. Ce problème peut être résolu partiellement en choisissant un seuil d'énergie au-delà duquel le SNR segmental sera calculé. Ensuite, il faut nécessairement que les deux signaux comparés soient alignés temporellement car ce critère est très sensible aux déphasages [12].

4.7.2 Le critère de la Distance Cepstrale (CD)

La distance cepstrale fournit une mesure de similarité entre les log-spectres (cepstres) de deux signaux. Pour deux signaux temporels $s_1(n)$, et $s_2(n)$, la distance cepstrale, évaluée pour chaque trame k de N échantillons :

$$CD(k) = \sum_{i=1}^{N-1} [\text{IFFT}(\log_{10}(|S_1(f_i, k)|) - \log(|S_2(f_i, k)|))]^2 \quad (4.3)$$

Où f_i est une fréquence (discrétisée) de la FFT. La fréquence nulle f_0 , correspondant à la puissance des signaux, n'est pas prise en compte.

La raison principale de l'emploi de cette distance vient du comportement de l'oreille humaine, plus sensible aux variations logarithmiques.

4.7.3 Le critère du désajustement

Le critère de désajustement est un critère robuste dans l'évaluation des performances. Ce critère se calcule à partir de la distance euclidienne entre les coefficients des réponses impulsionnelles réelles et mesurées (estimées). Ce critère, noté SM pour « System Mismatch », est défini par l'expression suivante :

$$SM_{dB} = 10 * \log_{10} \left(\frac{\|\hat{w}_{ij} - h_{ij}\|^2}{\|h_{ij}\|^2} \right) \text{ Où } i \neq j \in \{1,2\} \quad (4.4)$$

Où h_{ij} représente les coefficients du filtre réel (h12 ou h21), et w_{ij} représente la valeur moyenne des coefficients estimés (coefficients des filtre adaptatif w_{12} et w_{21}).

4.8 Résultats de simulations du nouvel algorithme BFD-SAD

Dans cette partie, nous présentons les résultats de simulations de notre nouvel algorithme appliqué à la structure réursive (Backward). En effet, nous avons analysé les performances de l'algorithme adaptatif proposé (BFD-SAD) en évaluant les différents critères définis auparavant dans la section 4.7.

Les paramètres utilisés pour le réglage de notre algorithme sont les suivants :

- **Paramètres du mélange**

Signal de parole originale : signal de la figure (4.2)

Un bruit USASI : signal de la figure (4.3)

Longueur du filtre : $L=32$

SNRs d'entrées sur les deux voies sont de 3dB.

- **Paramètres du nouvel algorithme BFD-SAD**

Longueur de la trame : 256

La taille de la FFT : 512

Recouvrement : 50%

Méthode de reconstitution du signal : Overlap-Add

Type de fenêtre : Hanning

Le signal de parole originale, le signal de bruit USASI et les signaux de sortie de l'algorithme BFD-SAD sont représentés avec leurs spectrogrammes sur les figures (4.10, 4.11, 4.12), respectivement.

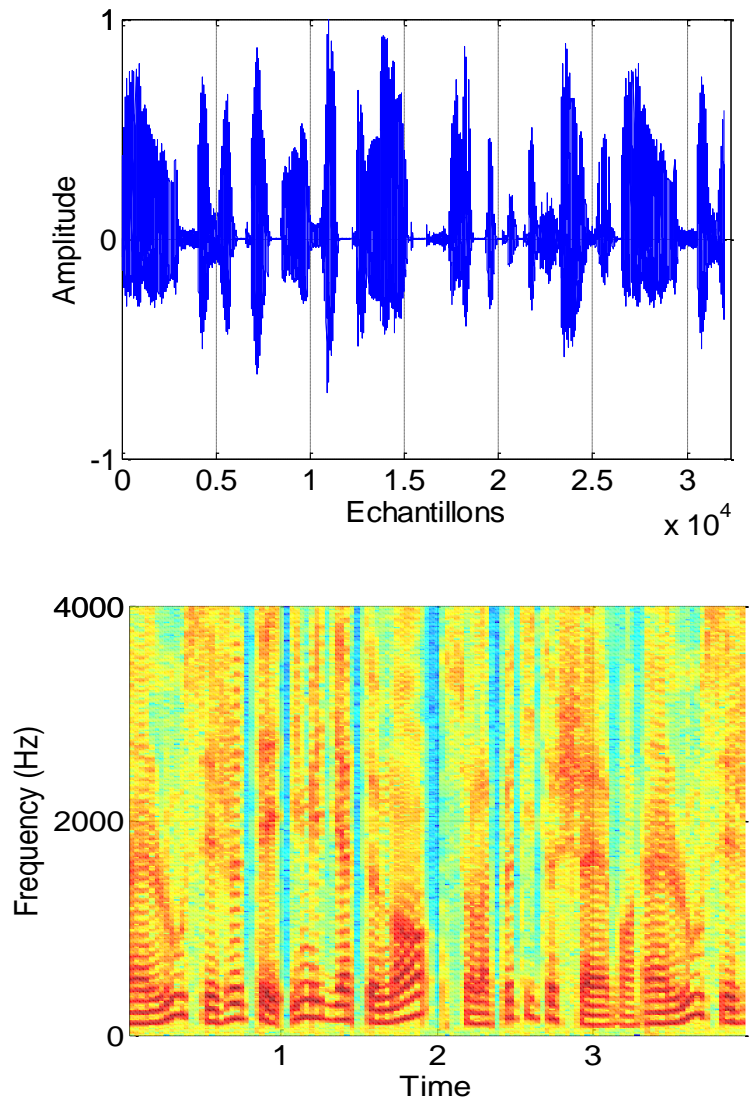
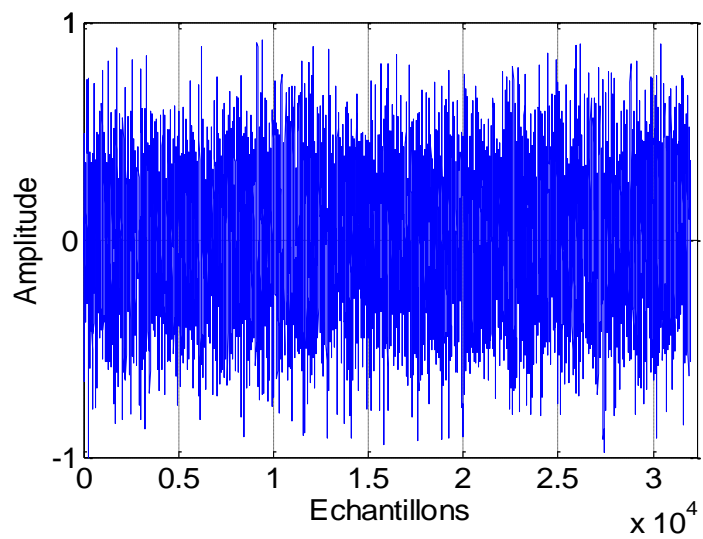


Figure 4.10. Signal de parole originale avec son spectrogramme.



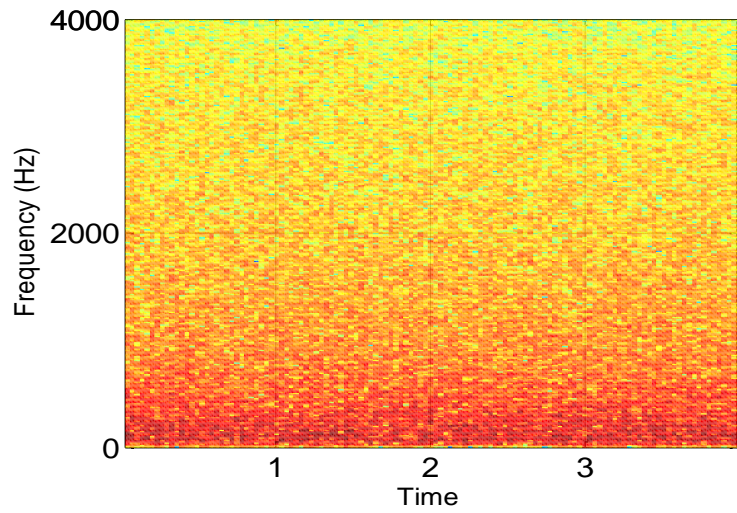


Figure 4.11.Signal de bruit USASI avec son spectrogramme.

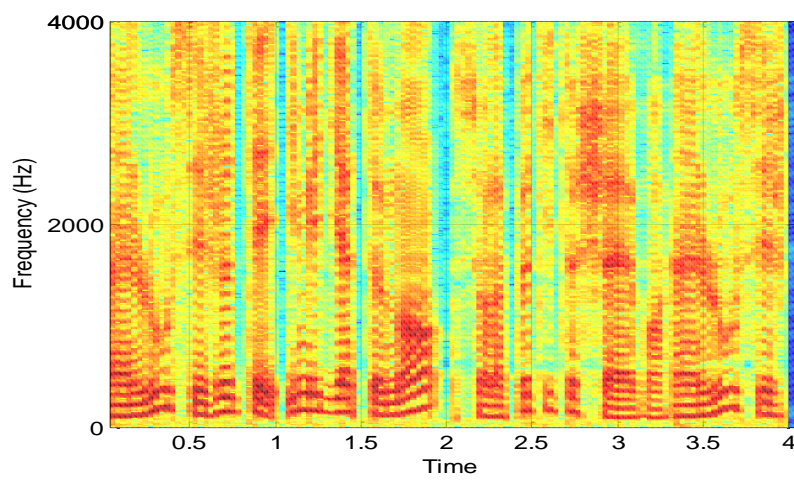
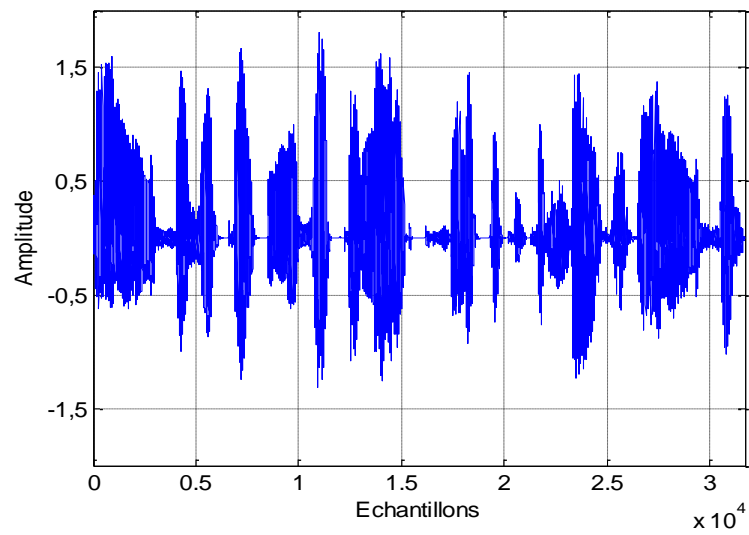


Figure 4.12.Signal de sortie $u_1(n)$ obtenue avec le nouvel algorithme proposé BFD-SAD.

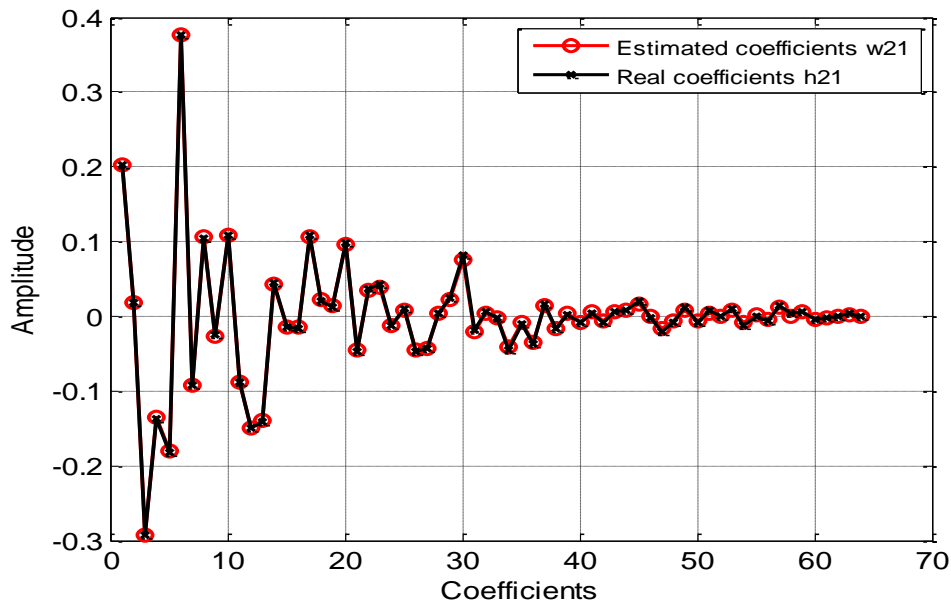


Figure 4.13. Comparaison entre la réponse impulsionnelle réelle $h_{21}(n)$ et estimée $w_{21}(n)$.

Nous notons que d'après la figure (4.12), le signal de parole en sortie $u_1(n)$ est visuellement nettoyé du bruit et très proche du signal de parole originale. Et ceci grâce au contrôle de l'adaptation du filtre $W_{21}(w, k)$ par une technique de détection d'activité vocale (DAV), ce filtre est adapté uniquement pendant les périodes de non activité vocale du mélange $p_2(n)$. Durant ces périodes d'inactivité vocale, le filtre adaptatif $W_{21}(w, k)$ n'a à son entrée que du bruit, ce qui lui permet d'identifier la réponse de couplage $h_{21}(n)$ puis supprimer le bruit du signal de sortie $u_1(n)$.

Nous présentons sur la figure (4.13), la comparaison entre la réponse impulsionnelle réelle $h_{21}(n)$ et identifiée $w_{21}(n)$, nous remarquons que les deux filtres sont presque égaux. Ainsi on est très proche de la solution optimale de Wiener.

4.9 Influence des paramètres L et SNR sur l'algorithme BFD-SAD

Pour étudier les performances du nouvel algorithme proposé BFD-SAD, nous avons réalisé plusieurs simulations pour voir l'effet de chaque paramètre sur les performances de ce dernier, c'est-à-dire l'étude du comportement de celui-ci en terme de vitesse de convergence.

Les paramètres de cette simulation sont :

Signal de parole source : signal de la figure (4.2)

Le bruit : signaux de figures (4.3) et (4.4)

L : taille du filtre adaptatif

SNRs : rapport signal sur bruit sur les deux voies d'entrées du mélange.

L'évaluation de l'effet de chaque paramètre cité précédemment sur les performances de notre structure, a donné les résultats suivants :

4.9.1 Effet de la taille du filtre adaptatif L

Pour voir l'effet de la taille du filtre L sur l'algorithme proposé BFD-SAD, nous avons varié celle-ci en prenant les valeurs suivantes : L=32, 64, et 128 pour des valeurs de SNRs égales à 3 dB.

Pour réaliser cette comparaison, nous avons tracé l'évaluation temporelle du désajustement (Système Mismatch) pour les trois tailles, et nous avons obtenu les résultats présentés sur les figures (4.14) et (4.15).

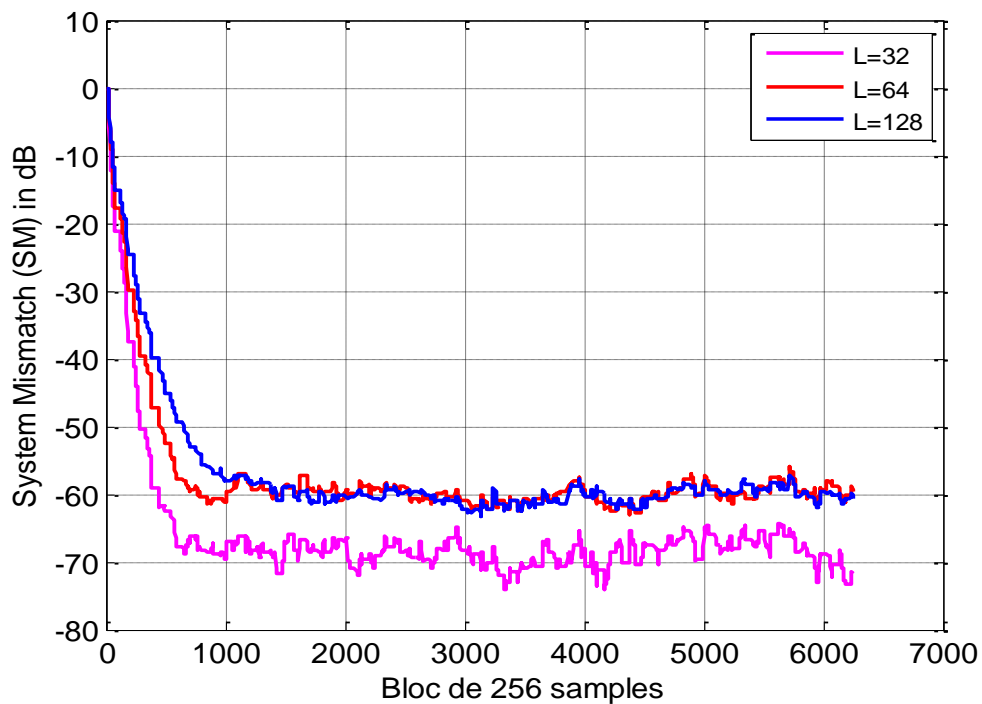


Figure 4.14. Effet de la taille du filtre sur le nouvel algorithme BFD-SAD en utilisant un bruit USASI dans le mélange.

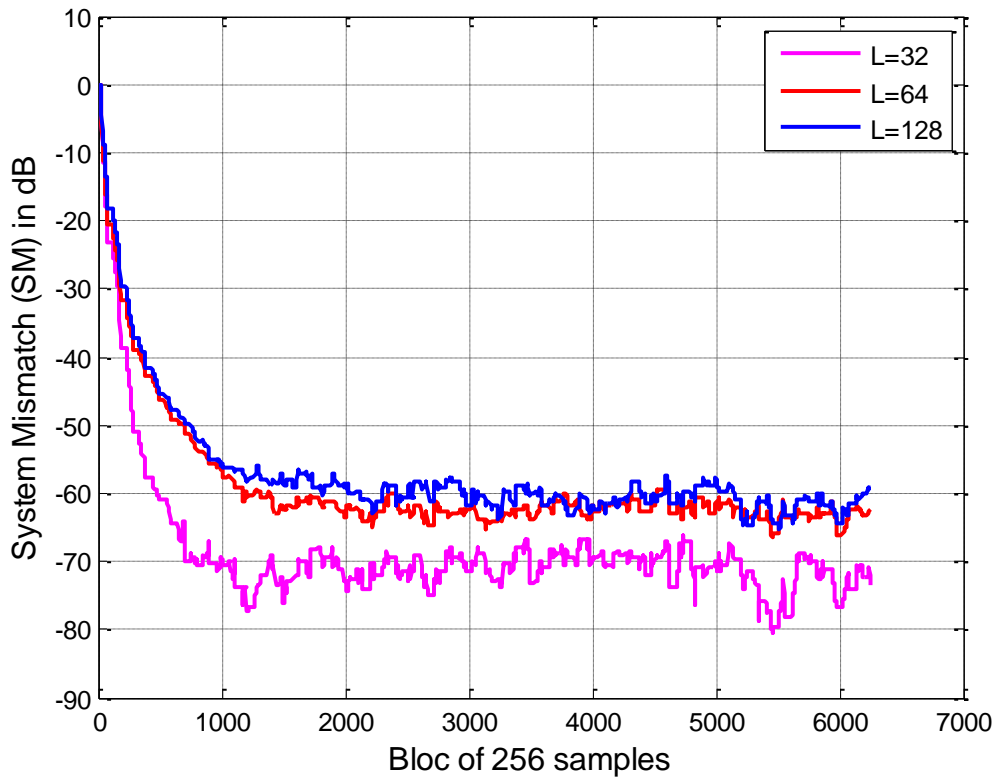


Figure 4.15. Effet de la taille du filtre sur le nouvel algorithme BFD-SAD en utilisant un bruit blanc dans le mélange.

D'après ces deux figures, nous montrons que notre nouvel algorithme proposé BFD-SAD converge mieux lorsque la taille du filtre L est faible, ce qui montre que la vitesse de convergence de cet algorithme est inversement proportionnelle à la taille du filtre adaptatif.

4.9.2 Effet du SNR du mélange

Dans cette simulation, nous avons étudié l'influence du rapport signal sur bruit de cet algorithme. Les paramètres de cette simulation sont :

La taille du filtre : $L=32$, nous avons varié la valeur du rapport signal sur bruit SNRs, en prenant trois valeurs $SNR=-3, 0, 3$. Les résultats de simulation du désajustement sont représentés sur les figures (4.16) et (4.17).

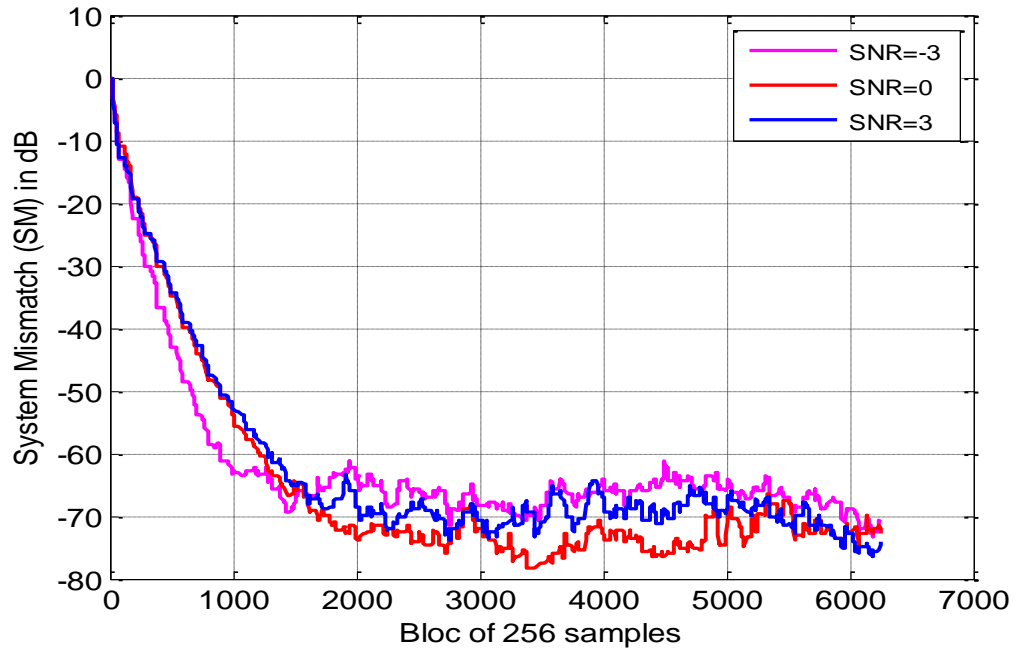


Figure 4.16. Effet du SNR d'entrée sur le nouvel algorithme BFD-SAD en utilisant un bruit USASI dans le mélange.

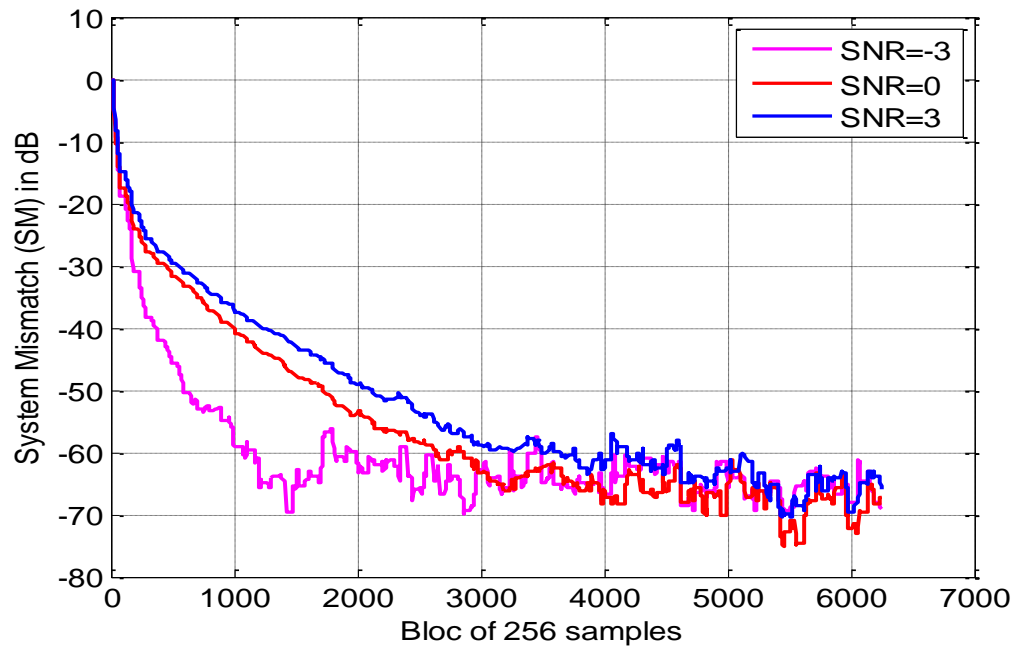


Figure 4.17. Effet du SNR d'entrée sur le nouvel algorithme BFD-SAD en utilisant un bruit blanc dans le mélange.

D'après les figures (4.16) et (4.17), nous avons remarqué que la vitesse de convergence de notre nouvel algorithme proposé BFD-SAD vers l'optimum est plus importante dans le cas où le rapport signal sur bruit SNR est faible (SNR=-3).

La figure (4.18), représente la comparaison du signal de parole de sortie (voie 1) et le signal de bruit de sortie (voie1) obtenue avec notre nouvel algorithme BFD-SAD.

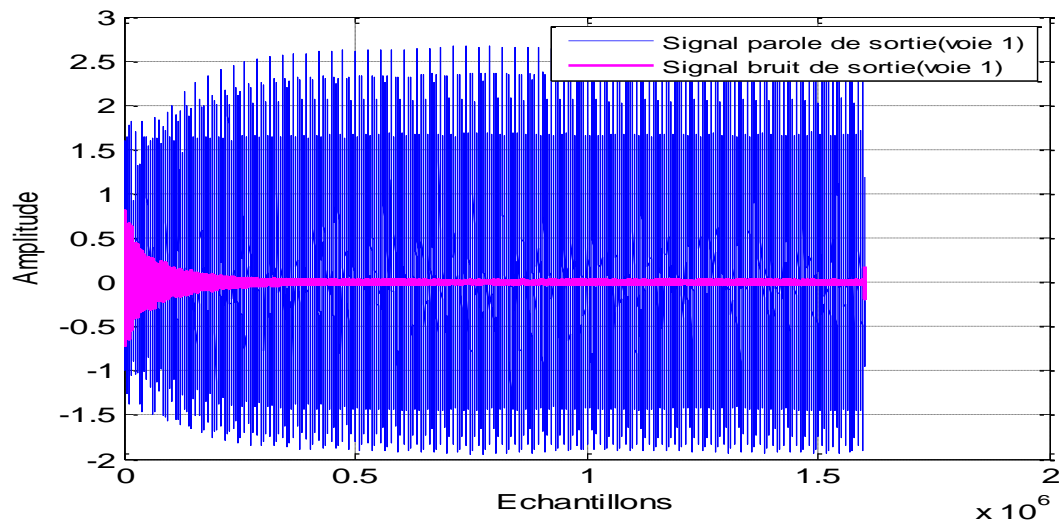


Figure 4.18. Comparaison entre le signal de parole de sortie (voie 1) en bleu, et le signal bruit (voie1) en magenta.

4.10 Comparaison entre les algorithmes BFD-SAD et BTD-SAD

Dans cette partie, nous faisons la comparaison entre les deux algorithmes BFD-SAD et BTD-SAD qui ont été bien illustrés dans le chapitre 3.

Les paramètres de simulation pour les deux algorithmes sont les suivants :

- **Paramètres du mélange**

Signal de parole originale : signal de la figure (4.2)

Un bruit USASI : signal de la figure (4.3)

Longueur du filtre : L=32.

SNRs d'entrées sur les deux voies sont de 3dB.

- **Paramètres du nouvel algorithme BFD-SAD**

Longueur de la trame : 256

La taille de la FFT : 512

Recouvrement : 50%

Méthode de reconstitution du signal : Overlap-Add

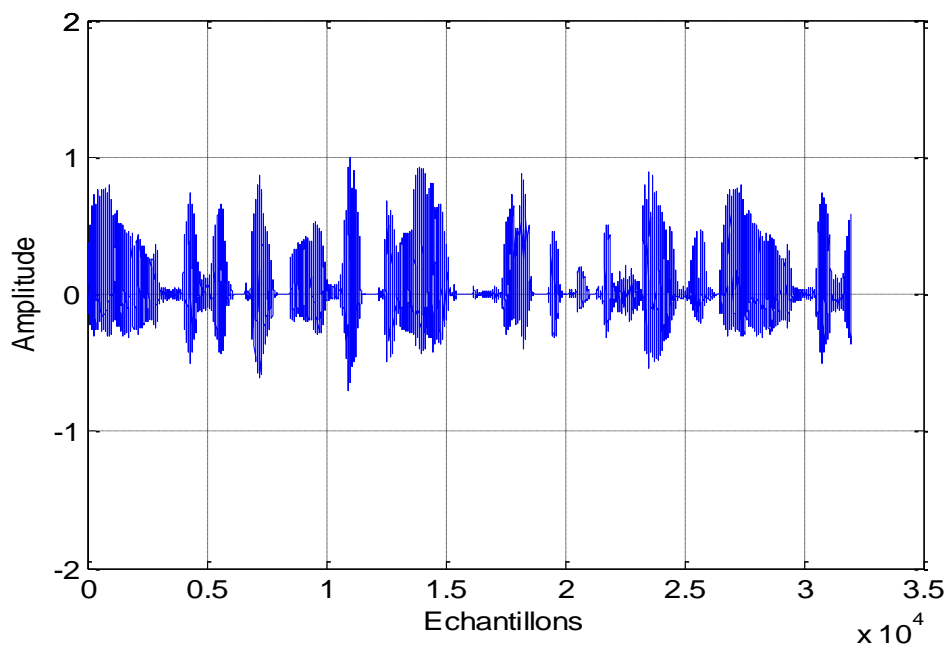
Type de fenêtre : Hanning

- **Paramètres du BTD-SAD**

Longueur de filtre : $L=32$.

Pas d'adaptation $\mu_{12} = \mu_{21} = 0.008$.

Afin de comparer les performances de la structure réursive (Backward) dans le domaine temporel (BTD-SAD) et fréquentiel (BFD-SAD), nous présentons sur les figures suivantes les signaux de sorties obtenus.



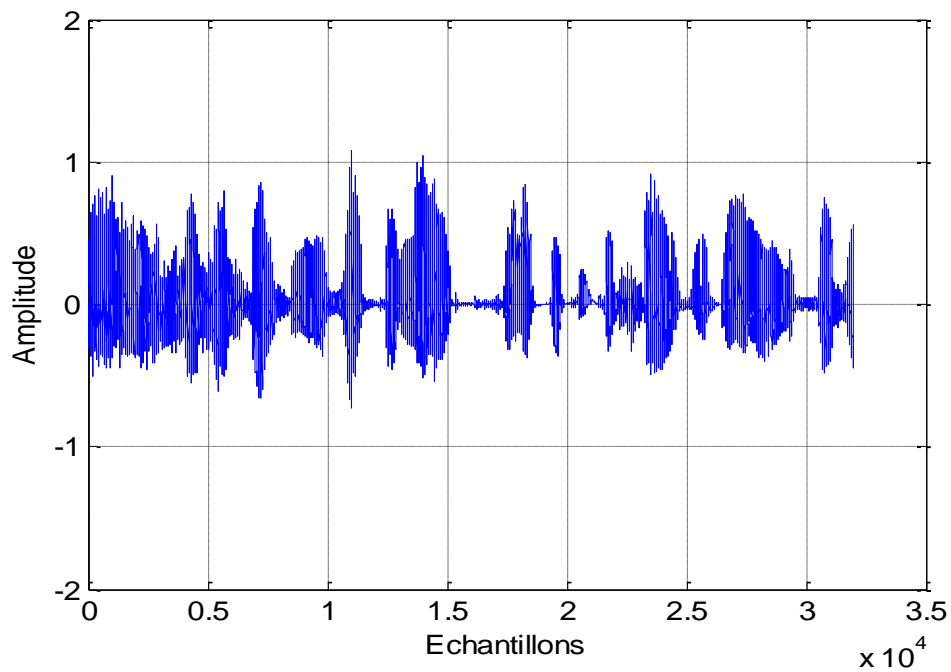
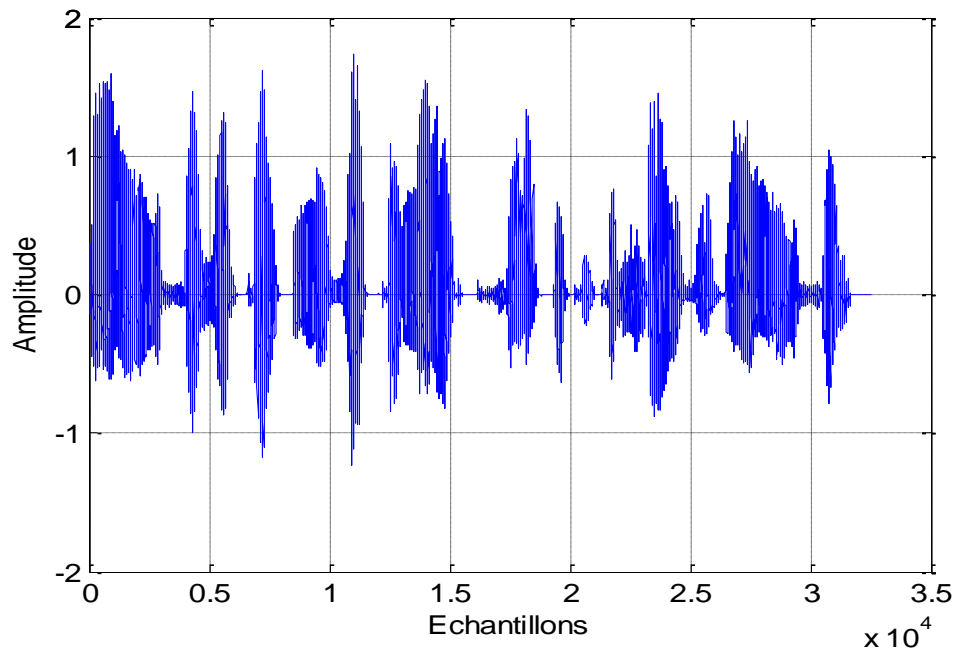


Figure 4.19. Signal de parole $s(n)$ originale (Haut), Signal de l'algorithme BFD-SAD (Milieu), Signal de l'algorithme BTD-SAD (Bas).

D'après la figure (4.19), nous pouvons conclure que les deux algorithmes BTD-SAD et le BFD-SAD proposé fonctionnent bien, c'est-à-dire que le signal de parole en sortie est presque totalement débruité et très proche du signal de parole originale.

Afin d'analyser les distorsions introduites par chacun des deux algorithmes dans le signal de parole de sortie, nous avons représenté sur la figure (4.20) la distance cepstrale entre le signal de parole original et le signal disponible en sortie de traitement pour chaque algorithme.

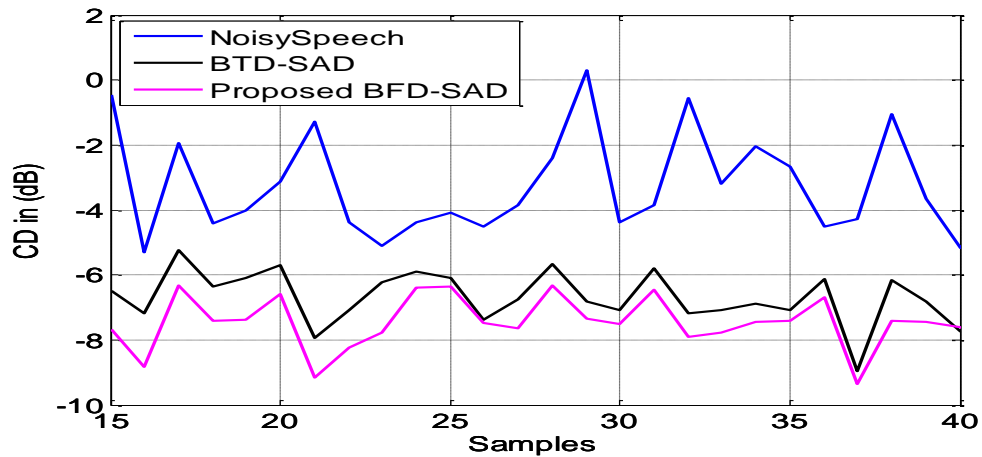


Figure 4.20. Evaluation du critère de la distance cepstrale des deux algorithmes BFD-SAD et BTD-SAD.

Nous remarquons que la distance cepstrale obtenue à partir du nouvel algorithme BFD-SAD est inférieure (supérieure en termes de qualité) à celle obtenue avec sa version temporelle. Ces résultats montrent bien le bon comportement de cet algorithme.

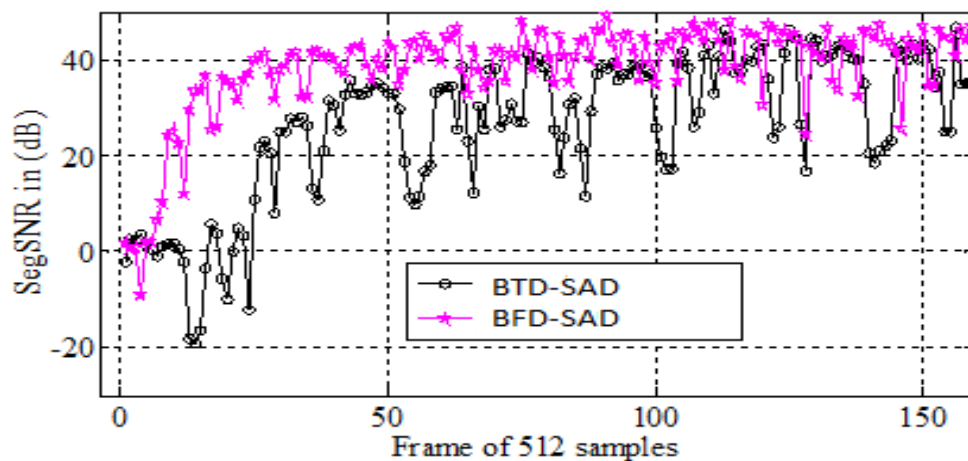


Figure 4.21. Evolution du critère du rapport signal sur bruit pour les deux algorithmes BFD-SAD [Magenta], et BTD-SAD [Noir].

Les résultats obtenus et présentés sur la figure précédente montrent la bonne performance de notre algorithme proposé par rapport à sa version temporel, en termes de rapport signal sur bruit (SNR). Les résultats que nous avons obtenus confirment l'efficacité de notre nouvel algorithme.

Pour compléter notre analyse, nous présentons sur la figure (4.22) l'évaluation temporelle du désajustement mesuré sur les coefficients du filtre $w_{21}(n)$ pour les deux algorithmes BFD-SAD et BTD-SAD.

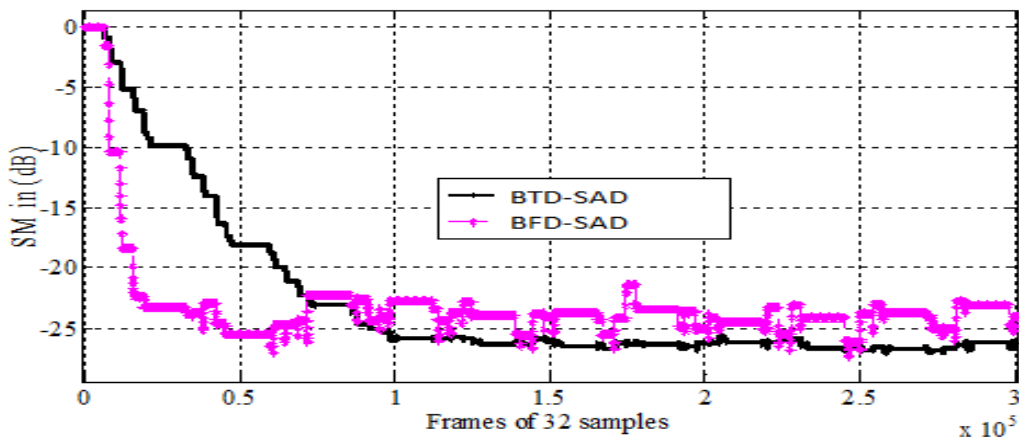


Figure 4.22. Comparaison du désajustement (SM) obtenu avec les deux algorithmes BFD-SAD [Magenta] et BTD-SAD [NOIR].

4.11 Critères subjectifs de qualité

Les mesures objectives sont largement utilisées pour juger de la qualité des traitements. Cependant, elles ne permettent pas de prendre en compte certains effets comme le bruit musical, le naturel du bruit ou la réverbération (due au lissage des estimateurs) qui sont pénalisants dans les tests subjectifs.

L'idéal est donc dans la mesure du possible de confirmer les résultats objectifs par des tests subjectifs. Les tests subjectifs consistent à faire écouter des séquences de signaux restaurés à un panel d'auditeurs dans des conditions bien définies. Ces auditeurs peuvent être choisis "naïfs", c'est-à-dire non habitués aux dégradations susceptibles d'affecter les signaux, ou au contraire "experts", c'est-à-dire habitués au traitement de la parole ou plus généralement du son. Il existe différents types de tests subjectifs,

parmi les plus courants, on peut citer les suivants, issus de la recommandation UIT-T P.800 [P800 1996] [23]:

– Le test ACR (Absolute Category Rating) consiste à faire écouter aux auditeurs chaque séquence restaurée sans référence. Ils donnent ensuite une note sanctionnant la qualité de cette séquence et par conséquent le traitement qui l’a généré.

– Le test DCR (Degradation Category Rating) consiste à faire écouter aux auditeurs la séquence à évaluer après un signal de référence. Ils doivent ensuite noter la dégradation de cette séquence par rapport à la référence.

– Le test CCR (Comparison Category Rating) consiste à faire écouter aux auditeurs des paires A-B de séquences. Ils doivent ensuite noter la séquence B par rapport à la séquence A.

Dans le cadre des travaux menés, le test CCR a été retenu car il est reconnu comme le plus discriminant pour évaluer les systèmes de réduction de bruit, et qu’il permet de comparer facilement deux traitements. Les consignes pour ce type de test sont de “prendre en compte la distorsion de la parole et le niveau de réduction de bruit” pour donner une note selon l’échelle suivante :

3	Bien meilleure
2	Meilleure
1	Légèrement meilleure
0	A peu près équivalente
-1	Un peu moins bonne
-2	2 Moins bonne
-3	Beaucoup moins bonne

Tableau 4.1. Les différentes classes de bruit.

Les notes d’opinions sont données en score CMOS (pour Comparative Mean Opinion Score). En utilisant une échelle de ce type, les auditeurs formulent implicitement deux jugements avec une seule réponse : “Quel est la séquence de meilleure qualité ?” et

“Quelle est la différence de qualité entre les deux séquences ?”. Ceci permet donc à l’issue du test de savoir quel est le traitement préféré et d’appréhender la distance de qualité qui les sépare.

Les résultats obtenus de la comparaison entre le BTD-SAD et le BFD-SAD sont listé dans le tableau suivant. Dans ce tableau nous avons interrogés 10 sujets âgés entre 20 et 30 et qui sont d’une bonne audition.

SUJETS	BFD-SAD par rapport à BTD-SAD
1	2
2	2
3	2
4	3
5	2
6	2
7	3
8	2
9	2
10	2

Tableau 4.2. Test CCR pour BFD-SAD et BTD-SAD.

On remarque que la majeure partie des sujets interrogés ont préféré notre nouvel algorithme BFD-SAD par rapport au BTD-SAD, ce qui confirme encore une fois les performances de ce dernier.

4.12 Conclusion

D'après les résultats de simulation de l'algorithme de décorrelation adaptative symétrique de type Backward (BEF-SAD), on peut conclure que celui-ci donne des bons résultats pour la séparation de sources et la réduction du bruit. Les résultats obtenus montrent bien que l'algorithme BFD-SAD converge vers la solution souhaitée d'une manière stable et rapide, avec un minimum de complexité.

Les critères objectifs retenus pour l'analyse des performances ont également été présentés, à savoir le SNRseg, la distance cepstrale, et le désajustement des coefficients des filtres adaptatifs.

On rajoute à cela que les critères subjectifs montrent tout aussi bien le bon comportement de notre algorithme proposé.

Conclusion générale

L'étude que nous avons présentée dans ce mémoire concerne le débruitage du signal de parole dans les systèmes de téléphonie mains libres a bord des véhicules, en utilisant la structure de séparation de sources récursive pour extraire le signal source à partir de la seule connaissance des signaux de mélange.

Pour cela, nous avons proposé un nouvel algorithme de séparation de source aveugle (BFD-SAD) en utilisant une méthode de débruitage bi-voies qui permet de filtrer le bruit d'un signal enregistré avec la mise en œuvre de deux microphones.

Ce dernier correspond à une version améliorée de l'algorithme de décorrélation adaptative symétrique temporel (BTD-SAD). Cette nouvelle structure utilise un système de détection d'activité vocale (DAV) manuelle qui permet de contrôler l'adaptation des filtres de séparation.

Les performances de notre algorithme sont prouvées dans les résultats de simulations représenté dans le chapitre quatre, ce qui montre en effet le bon comportement de cette structure (BFD-SAD) et sa convergence vers la solution souhaitée de manière stable.

Nous avons également remarqué que l'évolution de la vitesse de convergence de l'algorithme BFD-SAD est inversement proportionnelle à la taille du filtre et au rapport signal sur bruit (SNR) sur la voie de référence.

Une étude comparative avec les deux algorithmes de séparation de source BFD-SAD et BTD-SAD a été effectuée, nous avons remarqué que notre nouvel algorithme (BFD-SAD) a permis d'obtenir des performances supérieure en terme de distance cepstrale (DC), et de rapport signal sur bruit (SNR) en sortie de traitement.

Afin de conclure cette étude comparative nous avons réalisé des tests d'écoute informels des signaux obtenus en sortie de traitement pour les deux structures citées précédemment. Ce test a confirmé le bon comportement de notre algorithme BFD-SAD.

Bibliographie

- [1] André GIRARD, « Réduction de bruit de signaux de parole mono-capteurs basée sur la modélisation par EMD », Mémoire de maitrise, Université de SHERBROOKE Canada.
- [2] F. Cottet, «Aide-mémoire Traitement du signal » Livre, 2005 Dunod, Paris.
- [3] Meriem ZOULIKHA, « Méthodes améliorées pour la réduction du bruit acoustique et le rehaussement du signal de parole », Mémoire de Magister, USDB, Avril 2014.
- [4] G. Baudoin et J.F-. Bercher, « Eléments de traitement du signal », cours de traitement de signal version 0.89, Ecole supérieur d'ingénieurs en électrotechnique et électronique, Université Paris-Est, 1998.
- [5] O.Sentieys, « Introduction au traitement numérique du signal », ENSSAT – Université de Rennes 1, IRISA-équipe de recherche R²D², 2003.
- [6] J.-F.Bercher, P.Jardin, Notes de cours: Introduction au filtrage adaptatif, ESIEE Paris, 2003.
- [7] J. El Mhamdi, F.Regragui et M. Harnafi, « Traitement adaptatif appliqué au signal sismique », Bulletin de l'institut scientifique, Rabat, section Sciences de la Terre, no.30, 2008.
- [8] H. Alaeddine, « Application de la transformée en nombres entiers a la conception d'algorithmes à faible complexité pour l'annulation d'échos acoustiques », Thèse de Doctorat, Université de Bretagne Occidentale, 2010.
- [9] K. Meghriche, « Filtrage Adaptatif Utilisant les Statistiques d'Ordre Supérieur », Thèse de Doctorat d'Etat, Ecole Nationale Polytechnique, 2006.
- [10] J.Benesty . Notes de cours : Algorithme du gradient déterministe (steepest descent) .INRS-EMT Canada.
- [11] J.Benesty. Notes de cours : Algorithme du gradient stochastique (Least-mean-square-LMS).INRS-EMT Canada.

- [12] Asmaa Amehraye, « Débruitage perceptuel de la parole », Thèse de doctorat, Université de Bretagne Sud, 2009.
- [13] Thierry Dutoit(2000).Notes de cours : Introduction au Traitement Automatique de la parole .Faculté Polytechnique de Mons.
- [14] Pierre HANNA, « Modélisation statistique de sons bruités : étude de la densité spectrale, analyse, transformation musicale et synthèse », Thèse de Doctorat, Université Bordeaux I, 2003.
- [15] R. Boite, H. Boulard, T. Dutoit, J. Hancq et H.Leich, « Traitement de la Parole », Presses Polytechniques Universitaires Romandes, Lausanne, 2000.
- [16] N. Virag. "Single channel speech enhancement based on masking properties of the human auditory system".*IEEE Trans. Speech and Audio Processing*, vol. 7, pages 126–137, 1999.
- [17] Moussa Sofiane Karoui, « Méthodes de séparation aveugle de sources et application à la télédétection spatiale », Thèse de Doctorat, Université Toulouse III, 2012.
- [18] Moussa Akil, « Séparation de sources convolutives », Thèse de Doctorat, Université de Grenoble I, 2008.
- [19] R.Kalman," A new approach to linear filtering and prediction problems", *Journal of basic Engineering*, volume 82, n° 1, p. 35-45. (1960).
- [20] S.Boll,"Suppression of acoustic noise in speech using spectral subtraction". *IEEE Transactions on Acoustics, Speech and Signal Processing*, volume 27, n° 2, p. 113-120. (1979).
- [21] L. Miao, H. Qi, "End member Extraction from Highly Mixed Data Using Minimum Volume Constrained Nonnegative Matrix Factorization," *IEEE Transactions on Geosciences and Remote Sensing*, 45, pp. 765-777, 2007.
- [22] N. Charkani, Al Hassani, « Séparation auto-adaptative de sources pour des mélanges convolutifs. Application à la téléphonie mains-libres dans les voitures », Thèse de Doctorat, Institut Nationale Polytechnique de Grenoble, 1996.
- [23] Cyril PLAPOUS, « Traitement pour a réduction du bruit. Application à la communication parlée », Thèse de Doctorat, Université Rennes I, Décembre 2005.

- [24] M. Djendi, A. Gilloire, P.Scalart, Noise cancellation using two closely spaced microphones: experimental study with a specific model and two adaptive algorithms, IEEE Int.Conf.ICASSP, Toulouse, France, 14-19 May 2006, vol.3, pp.744-747.
- [25] M. Djendi, F.Khemies, A.Morsli, specom, Submitted to 17TH international conference on speech and computer specom 2015, 20-24 Septembre, 2015 Greece. Athens.
- [26] M.Djendi, S.Bensafia, M.Safi, Accepted by the international conference on engineering & MIS 2016, Agadir,Maroc, 22-24 septembre,2016.