

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE MOHAMED BOUDIAF M'SILA

FACULTE : Science et Technologies

DEPARTEMENT: Electronique

N° :



DOMAINE :

FILIERE :

OPTION : Systèmes des Télécommunications(STLC)

**Mémoire présenté pour l'obtention
du diplôme de Master Académique**

Par : BALAH Oualid

Thème :

**Séparation Aveugle de Signaux Audio Utilisant
Particle Swarm Optimisation (PSO)**

Évalué par le jury :

Président :	Dr. LALAOUI Lahouaoui	Université Mohamed Boudiaf M'sila
Rapporteur :	Dr. KHALFA Ali	Université Mohamed Boudiaf M'sila
Examineur :	Dr. ZERDOUMI Zohra	Université Mohamed Boudiaf M'sila

Année universitaire : 2020 / 2021

REMERCIEMENTS

Mes vifs remerciements sont en premier lieu adressés à mon encadreur, **Dr. KHALFA Ali**, de l'université Mohamed Boudiaf de M'sila, pour sa guidance et ses conseils afin d'accomplir ce travail.

Je tiens également à exprimer ma profonde gratitude aux membres de jury qui ont accepté de valoriser et de juger ce travail de recherche, notamment, **Dr. LALAOUI Lahouaoui** de l'Université Mohamed Boudiaf de M'sila, qui me fait l'honneur de présider le jury et **Dr. ZERDOUMI zohra**, de l'Université Mohamed Boudiaf de M'sila en tant que membre examinateur de jury.

Que tous ceux qui m'ont encouragé et aidé pour la finalisation de ce travail trouvent ici ma gratitude.

DEDICACES

A mes trois anges Ilyes, Younes et

Layan Nour

A toute ma famille

A tous mes amis et spécialement

Dr. Messoud Gareh,

Mr Djamel Kitiri,

Mr Ilyes Ferhi

A vous tous je vous dédie ce modeste
travail.

ملخص

يعد الفصل الأعمى للمصادر موضوع بحث هام جدا في مجال معالجة الإشارات. تعتمد هذه التقنية على استعادة الإشارات التي يراد فصلها من خليط مرصود مكون من عدة إشارات دون معلومات مسبقة معروفة حول هذه الإشارات الأصلية باستثناء استقلاليتها الثابتة. في هذه الدراسة، استخدمنا خوارزمية سرب الجسيمات (PSO)، وهي خوارزمية تطورية تستخدم على نطاق واسع، نظرا لسرعة تقاربها، في مسألة تحسين الفصل الأعمى للصوت.

الكلمات المفتاحية : الفصل الأعمى للمصدر (BSS)، الصوت، خوارزمية سرب الجسيمات (PSO).

Résumé

La Séparation Aveugle de Sources (BSS) est devenue un sujet de grand intérêt dans le domaine du traitement du signal. BSS se base sur la récupération des signaux inconnus de source, uniquement à partir des mélanges observés. En effet, aucune information sur les signaux sources originaux n'est connue sauf leur indépendance statique. Dans la présente étude nous avons utilisé l'algorithme d'optimisation par essaim de particules (PSO), un algorithme évolutionnaire largement utilisé pour sa rapidité de convergence, dans le problème d'optimisation de la séparation aveugle de sources audio.

Mots clés : Séparation Aveugle de Sources BSS; Audio, Optimisation par essaim de particules PSO.

Abstract

Blind Source Separation (BSS) became a subject of great interest in the signal processing field. BSS is based on recovering unknown source signals only from the observed mixtures. In fact, no information about the original signals sources is known except their static independence. In the present study we have used Particle Swarm Optimization algorithm (PSO), an evolutionary algorithm widely used, for its convergence speed, in the optimization problem of blind audio source separation.

Keywords: Blind Source Separation BSS, Audio, Particle Swarm Optimization PSO.

INTRODUCTION GENERALE

CHAPITRE I : Etat de l'Art en Séparation Aveugle de Sources

I.1 Introduction

- I.2 Description générale de la séparation aveugle de sources
- I.3 Bref historique sur la séparation aveugle de sources
- I.4 Applications de la séparation aveugle de source
- I.5 Les types de mélanges

I.6 Conclusion

CHAPITRE II : Signal Audio

II.1 Introduction

- II.2 Classification des signaux
- II.3. Le signal vocal
- II.4. Système phonatoire et modèle de production
 - II.4.1 Le niveau acoustique
 - II.4.2 Le niveau phonétique
 - II.4.3 Le niveau phonologique
 - II.4.4 Le niveau morphologique
- II.5 Application du signal parole

II.6 Conclusion

CHAPITRE III : Algorithmes Evolutionnaires

III.1 Introduction

III.2. Les algorithmes évolutionnaires (AE)

III.3. Classes des AE

III.4. Structure des AE

III.5. La méthode PSO

III.6. Mécanisme de confinement

III.7 Conclusion

CHAPITRE IV : Séparation Aveugle de Sources Audio utilisant PSO_Résultats de Simulations

IV.1 Introduction

IV.2 Les signaux sources de test

IV.3. Evaluation de la fonction objective

IV.4 Mesure de performances

IV.5 Résultats

IV.6 Interprétations

IV.7 Conclusion

CONCLUSION GENERALE

LISTE DES FIGURES

Figure II.1 Enregistrement acoustique d'un signal acoustique

Figure II.2 Spectrogramme à large bandes (en bas), à bande étroite (en haut) et évolution temporelle de la phrase anglaise 'Alices's adventures'

Figure II.3 (a) Son voisé

Figure II.3 (b) Son non voisé

Figure II.4 L'appareil phonatoire

Figure II.5 Section du larynx, vue de haut

Figure II.6 Représentation des voyelles dans plan F1-F2

Figure II.7 Le système auditif

Figure III.1 Algorithme de la PSO

Figure IV.1 Signal original et signal estimé par PSO_(Music_Herbalizer.wav)

Figure IV.2 Signal original et signal estimé par PSO_(Music_Piano.wav)

Figure IV.3 Signal original et signal estimé par PSO_(speech1.wav)

Figure IV.4 Signal original et signal estimé par PSO_(speech2.wav)

Figure IV.5 Signal original et signal estimé par PSO_(Music_Herbalizer.wav)

Figure IV.6 Signal original et signal estimé par PSO_(speech2.wav)

LISTE DES TABLEAUX

Tableau IV.1 Nos résultats avec PSO _ Résultats avec JADE [88]

INTRODUCTION

GENERALE

INTRODUCTION GENERALE

Le traitement de la parole, qui fait partie de la discipline scientifique du traitement de signal, a connu une révolution technologique dans plusieurs domaines d'applications. Une révolution liée principalement au développement des moyens et des techniques de télécommunications.

Le signal est le support de l'information et le moyen de communication entre l'homme et la machine. Il est issu de l'environnement physique (vibrations sonores, images visuelles...) et mesuré, dans la plupart des cas, par des capteurs dans le but de le traiter par la suite en fonction de la modification que l'on souhaite apporter.

Cet environnement externe impose aux signaux une transformation au cours de leur propagation telle qu'une superposition avec d'autres signaux de l'environnement. Les organes auditifs de l'être humain constitue un cas typique de tels systèmes du moment où l'être humain est capable de distinguer, dans le mélange qu'il reçoit à son oreille, une des sources de signal détecté.

C'est dans ce contexte que la séparation aveugle de sources (BSS) trouve son application. Elle consiste à estimer un ensemble constitué, dans la plupart des situations, par la somme de plusieurs contributions de signaux inconnus dits signaux sources à partir des signaux connus dits observations. Les observations sont des mélanges des signaux sources reçues par des capteurs.

On peut la comprendre à travers les exemples suivants :

- Dans une fête animée, on a tendance à entendre un mélange de voix, de bruits et de la musique. Si ces sons sont captés par un microphone, comment peut-on identifier une seule voix d'intérêt à partir du son mélangé ?

- Lors de la réception de signaux radio provenant d'un champ de bataille, comment peut-on capturer, séparer et distinguer une communication amie de celles des troupes ennemies ?
- Dans le domaine de la télédétection par l'analyse d'images, les informations de sol, celles de minerai et d'autres informations d'interférence sont mélangées dans le signal reçu. Comment déterminer les informations de sol et de minerai à partir de ce signal mélangé?

Ce travail vise en effet, à étudier la séparation aveugle de sources pour des mélanges instantanés; en particulier, des sources audio (parole et musique).

Ce mémoire est organisé en quatre chapitres :

Après une introduction générale qui résume le contexte de notre travail, le premier chapitre est un état de l'art de la séparation aveugle de sources.

Dans le deuxième chapitre sont exposées les principales notions du signal audio.

Le troisième chapitre aborde les algorithmes évolutionnaires, leurs principe de base et les fameuses techniques utilisées en littérature, notamment la méthode adoptée dans notre travail d'optimisation, la PSO (Particle Swarm Optimization).

Le quatrième chapitre résume nos résultats de simulations et interprétations.

Le mémoire est couronné par une conclusion générale.

CHAPITRE I

Etat de l'Art en Séparation Aveugle de Sources

I.1. Introduction

Dans les systèmes de traitement audio, qui visent à améliorer la qualité des signaux, de les compresser, ou d'en extraire des informations, une ou plusieurs entrées peuvent être considérées à la fois. L'environnement externe impose à ces signaux, une transformation au cours de leur propagation telle qu'une superposition avec d'autres signaux de l'environnement. La récupération des signaux originaux fait l'objet de la séparation aveugle de sources, qui est devenue un grand champ d'intérêt de beaucoup de chercheurs à travers le monde et dans divers domaines.

I.2. Description générale de la séparation aveugle de sources

La séparation aveugle de sources (en anglais BSS : Blind Separation Sources), est une puissante méthode en traitement du signal qui a été proposée à la fin des années 80. Dérivée à partir des réseaux de neurones artificiels, du traitement statistique du signal et de la théorie de l'information, la SAS est devenue un sujet important de la recherche et de développement dans de nombreux domaines, notamment les sciences biomédicales, les communications, le traitement de la parole, le traitement d'images, les sciences de la Terre, etc.

À l'heure actuelle, les travaux de recherche liés à la SAS ont acquis un grand intérêt dans le domaine du traitement du signal et sont donc devenus d'une grande valeur pratique.

La SAS, qui est un problème classique et complexe dans le traitement des signaux, consiste à extraire et à récupérer les signaux sources à partir de données statistiques à variables multiples. Le mot "Aveugle" signifie que le signal source est inconnu (non observé) et que, soit les caractéristiques du système hybride ne sont pas

connues à l'avance, soit qu'il n'y a qu'une petite quantité d'information a priori sur le signal source (comme la non Gaussienne, le cycle-stabilité et l'indépendance statistique).

Dans les applications scientifiques et techniques, de nombreux signaux observés peuvent être perçus comme un mélange d'une pluralité de signaux sources, c'est-à-dire que le signal mélange observé est une série de sorties de capteurs, chacun recevant différentes combinaisons de signaux sources. La principale tâche de la SAS est donc de récupérer le signal source qui nous intéresse à partir des données observées Figure I.1.

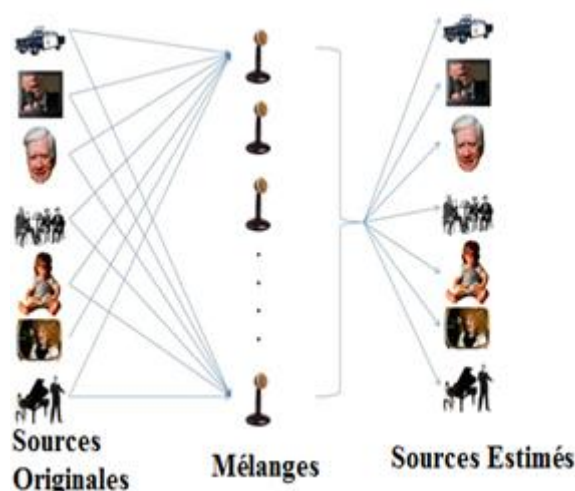


Figure I.1. Principe de la séparation aveugle de sources

I.3. Bref historique sur la séparation aveugle de sources

C.Jutten et J.Hérault étaient les premiers chercheurs qui ont menés des travaux dans ce sens en 1985, en présentant un modèle de réseau neuronal de rétroaction pour séparer deux signaux sources indépendants mélangés [1]. L'algorithme proposé traite le problème du mélange instantané sans bruit avec une matrice carrée et des signaux de réelles valeurs.

J.L Lacoume et P.Ruiz ont proposé un algorithme en 1988, qui minimise une fonction objective basée sur les cumulants croisés des sorties d'un filtre inverse, qui représentent les outils naturels de mesure d'indépendance [2].

J.F Cardoso et Souloumiac en 1989, et dans le but de rendre unitaire la matrice de mélange, ont proposé de considérer les propriétés algébriques des cumulants d'ordre quatre en utilisant l'information d'ordre deux puis d'utiliser des techniques de diagonalisation pour l'identifier. Ils ont montré que dans leur algorithme, les cumulants d'ordre quatre sont suffisants pour l'estimation du mélange. Une version non itérative de leur méthode a été aussi présentée dans [3].

P. Comon a développé dans [4] une approche qui élargit la théorie de l'analyse en composantes principales (PCA, Principale Components Analysis) à l'analyse en composantes indépendantes. Il a introduit la notion de fonction de contraste où il a utilisé comme contraste à maximiser la somme des cumulants d'ordre quatre.

M. Gaeta et J. L. Lacoume [5] proposent de résoudre le problème de séparation par une technique du maximum de vraisemblance approchée.

En 1993, B.Laheld et J.F Cardoso ont développé un ensemble d'algorithmes adaptatifs basés sur la notion de gradient matriciel relatif. Cardoso et Souloumic dans la même année présentent une solution algébrique à la maximisation de leur contraste basé sur les cumulants d'ordre 4. Ils ont aussi étendu l'algorithme de Jacobi afin de diagonaliser conjointement un ensemble de matrices [6].

En 1995, Delfosse et Loubaton [7] ont proposé une manière originale d'optimiser le contraste uni-source d'ordre 4 sur lequel repose leur méthode. Hyvärinän et Oja présentent en 1997 l'algorithme Fast-ICA dont la version initiale [8] exploite le principe de déflation et la mesure de néguentropie représentée par la valeur approximée à la valeur absolue du kurtosis de la source estimée.

En 2001, Comon trouve une solution analytique à la maximisation de son contraste défini au signe près. En fait, ce contraste a été présenté initialement par Moreau et al. Dans [9],[10],[11] où ils estiment que ce critère est un contraste sous réserve que les kurtosis des sources soient de même signe.

Fety [12], Tong et al. [13], [14] sont les premiers à exploiter les propriétés algébriques de deux matrices de covariance des observations, associées à un retard nul et non nul

respectivement. Comon et Lacoume démontrent plus tard le bien-fondé de ces méthodes.

Belouchrani et al. dans l'algorithme SOBI (Second Order Blind Identification) [15], [16] exploitent plusieurs matrices de covariance des observations associées à des retards non nuls. Ils montrent qu'après blanchiment des observations, une diagonalisation conjointe des matrices en question permet d'estimer le mélange à une matrice triviale près. Cependant cette approche nécessite que les spectres des sources soient distincts. Ils ont, quelques années plus tard, amélioré l'algorithme SOBI, et plus particulièrement l'étape de blanchiment. L'avantage de RobustSOBI [17] est d'être insensible asymptotiquement à un bruit temporellement décorrélé et de cohérence spatiale inconnue.

Durant cette période, Ziehe et Müller ont proposé TDSEP (Temporal Decorrelation Source Separation) [18].

Les méthodes exploitant les propriétés algébriques des cumulants d'ordre deux tirent profit de la présence de sources non stationnaires. L'idée consiste à diagonaliser conjointement un ensemble de matrices RTFS (Représentations Temps-Fréquence Spatiales) en utilisant la classe de Cohen [19]. Le défi est d'avoir les mêmes propriétés offertes par les matrices de RTFS que par les matrices de covariance des observations. Ainsi, Fevotte et al. présentent la méthode TFBSS (Time-Frequency Blind Source Separation) en proposant un critère de sélection automatique des bonnes matrices de RTFS [20], [21], [22].

Une autre idée consiste à supposer que les sources non stationnaires sont stationnaires par morceaux. Une diagonalisation conjointe des matrices de covariance estimées à partir de chaque bloc de données permet de ce fait de répondre au problème.

D'autres approches dans la littérature basées sur la représentation géométrique des mélanges ont été proposées. Citons par exemple les travaux de Puntonet et al. [23], [24] ainsi que ceux de Diamanteras [25].

D.T Pham et J.F Cardoso dans [26], [27] en 2000 et 2001 traitent la séparation de sources non stationnaires.

Un certain nombre d'auteurs parmi eux K.Kokkinakis en 2006 [28], ont utilisé des critères en combinant le domaine temporel et domaine fréquentiel avec la mise en œuvre de la rapidité des calculs.

Un des moyens les plus efficaces de la séparation aveugle de sources dans le cas des mélanges convolutifs utilise les relations statistiques de la densité de puissance du signal, au lieu de résoudre la séparation de sources dans chaque bande de fréquence pour le cas du mélange instantané. Kim [29] propose une version d'algorithme multidimensionnel pour l'estimation de la densité [30] en exploitant les dépendances en fréquence d'ordre supérieur.

Dans les travaux menés sur la séparation de sources audio, les propriétés des signaux audio notamment les propriétés de non-stationnarité et de parcimonie dans différents domaines ainsi que le caractère harmonique des signaux sont à exploiter afin de mettre en œuvre des méthodes de séparation plus efficaces tout en améliorant les performances de séparation et réduisant le coût de calcul.

La première méthode de séparation de sources audio a été réalisée en utilisant des algorithmes de décomposition modale [31] [32]. Cette approche permet de traiter le cas des mélanges instantanés et convolutifs et permet en particulier, de traiter le cas sous-déterminé (le cas où l'on a moins de capteurs que de sources).

Dans [33], la même méthode est utilisée pour un mélange convolutif sur-déterminée en lui conjuguant une décomposition en valeurs et vecteurs propres et un critère de parcimonie approprié.

D'autres méthodes, utilisées avec différentes transformées dans le domaine temps-fréquence (DTFQ : Distributions Temps-Fréquence Quadratiques et TFCT : Transformée de Fourier à Court Terme), supposent que les sources sont disjointes

dans le domaine temps-fréquence, autrement dit seulement une source parmi l'ensemble est présente dans chaque point temps fréquence.

Les premiers travaux sur la séparation de sources de mélanges convolutifs ont été réalisés dans le domaine temporel [34]. L'approche ainsi introduite pour résoudre le problème de séparation de sources audio dans le cas sous-déterminé nécessite une pré-estimation du canal de mélange.

Dans [35], nous trouvons un algorithme simple s'intéressant aux situations physiques réelles comme le cas d'un environnement acoustique.

L. Parra et ses collaborateurs dans [36] ont appliqué concrètement la séparation aveugle de sources pour identifier les signaux acoustiques (non-stationnaires) simultanément enregistrés dans un environnement multi-trajets dans un système de mélange convolutif.

Dans [37], le travail s'intéresse au problème de la séparation de mélanges convolutifs des sources audio dans une situation réelle où le principe de l'analyse en composantes indépendantes (ICA) est adapté dans le domaine fréquentiel.

Un algorithme en ligne dédié au traitement de la parole [38], permet de séparer des sources enregistrées dans un environnement réel où des personnes parlent simultanément en se déplaçant dans une salle. Le problème de séparation et de déréverbération d'un mélange dynamique des sources a été présenté aussi dans [39] où les caractéristiques acoustiques des sources sont prises en considération.

En 2007, Scott C. Douglas [40] a proposé son approche Spatio-TemporalFast-ICA, une extension de l'algorithme Fast-ICA.

Le travail rapporté dans [41] montre que la séparation de sources peut être également appliquée au rehaussement de la parole.

K. Torkkola [42], inspiré par les méthodes de déconvolution aveugle, a modélisé le procédé de séparation comme un filtre RIF.

T. W. Lee et al. [43] ont modélisé la procédure de séparation comme un filtre IIR en supposant que l'environnement de l'enregistrement est à phase minimale, ce qui n'est pas toujours valable.

Mukai et al. [39], [44] ont étudié le problème de séparation de mélanges convolutifs en tenant compte des propriétés acoustiques des signaux. Ils traitent le problème en fréquence en appliquant une transformée de Fourier à court terme.

Le problème formulé par P. Smaragdis [45], [46] repose sur l'estimation de la matrice de séparation pour chaque trame de fréquence qui peut avoir une solution au sens du maximum de vraisemblance.

H. Bousbia et al. [47] ont utilisé la non-stationnarité des signaux dans le domaine temps-fréquence toujours pour le problème de séparation de mélanges convolutifs.

En bande étroite peu de travaux ont été faits [48]. En large bande, W. Jun [49] propose une approche bi-spectrale. V. Capdevielle et al. ([50], [51] et [52]) ont proposé une méthode basée sur la continuité du spectre entre les différentes bandes de fréquences.

H. Attias et C. E. Schreiner [53] ont présenté une approche très intéressante en combinant ou non les informations à la fois temporelles et fréquentielles, mais cette méthode est limitée aux signaux sur-gaussiens, ce qui n'est généralement pas applicable aux signaux audio et de parole.

H. C.Wu et J. Principe [54] puis L. Parra et C. Spence [55] présentent également des approches orientées temps-fréquence basées sur des critères Infomax.

Une multitude d'algorithmes et d'approches sont récemment mis en œuvre traitant le problème de séparation de sources. Citons par exemple A. Ghazdali et al, qui ont présenté en 2017 une nouvelle approche BSS en considérant le cas des modèles de mélanges instantanés linéaires bruités de sources indépendantes / dépendantes. Dans [56] ils ont combiné deux méthodes [57] [58].

I.4. Applications de la séparation aveugle de sources

La séparation de sources est un champ de recherche très actif en raison de son intérêt théorique et de ses diverses applications pratiques. Nous citons dans cette liste d'applications non exhaustive :

I.4.1. Séparation de signal parole

Étant donné que la situation en temps réel des positions des microphones par rapport aux sources peut être assez arbitraire, le processus de mélange n'est pas connu et doit donc être estimé à l'aveugle. Dans cette situation, les algorithmes BSS sont importants.

De plus, les algorithmes BSS convolutifs ont tendance à être de préférence appliqués dans le domaine de la musique. Douglas a séparé deux voix en utilisant des techniques BSS convolutives [59], tandis que Vincent dans [60] a appliqué la technologie ICA pour obtenir les sons de différents instruments dans une bande sonore.

I.4.2. Communication et traitement du signal

BSS a une large application dans les communications numériques sans fil [61]. Divers chercheurs ont proposé d'utiliser un système de réception de réseau d'antennes et des techniques de filtrage spatial ainsi qu'un algorithme BSS pour réaliser la séparation du signal multi-utilisateur, de sorte qu'une pluralité de signaux d'utilisateur peuvent être transmis sur le même canal en même temps, améliorant ainsi considérablement la capacité du canal. Dans un système à accès multiple par répartition en code CDMA (Code Division Multiple Access), plusieurs utilisateurs partagent la même bande passante de canal et transmettent des signaux simultanément. Les utilisateurs peuvent distinguer leurs propres signaux par les

codes d'étalement orthogonale plus longue et, par conséquent, l'utilisation de l'ICA peut aider à améliorer le SNR (rapport signal/bruit) du côté du récepteur.

En 2006, l'institut de recherche japonais RIKEN et Chaos ware Company ont développé une nouvelle technique pour séparer précisément une onde électromagnétique hybride de téléphone, ainsi améliorant considérablement la vitesse de transmission et la clarté des communications.

I.4.3. Traitement et reconnaissance d'image

Actuellement, la technique BSS est largement utilisée dans le domaine du traitement d'image, avec la technologie de séparation aveugle utilisée dans l'extraction de caractéristiques d'image, la reconnaissance faciale, la détection d'objets en mouvement, les filigranes d'images numériques, le débruitage d'image, la séparation d'images et la restauration d'images [62] [63].

Par exemple, dans une application de restauration et de reconstruction d'images, les méthodes BSS disponibles peuvent aider à récupérer le contenu original des images à partir d'images contaminées et à éliminer les effets négatifs, tels que le bougé de l'appareil photo.

I.4.4. Traitement de l'information géologique spatiale

Dans le traitement des informations spatiales géo-scientifiques, les techniques BSS ont été appliquées à de nombreux aspects du traitement des données de télédétection, notamment la fusion d'images de télédétection, la décomposition mixte de pixels et la classification et la reconnaissance d'images de télédétection [64] [65].

Dans la prévision des minéraux, les méthodes BSS ont été appliquées avec succès à l'analyse des données géophysiques et géochimiques. Les techniques BSS peuvent également être utilisées dans l'exploration sismique. L'utilisation de techniques BSS dans le dynamitage explosif du sol peut aider à distinguer différentes caractéristiques lithologiques des signaux réfléchis ou diffractés reçus [66].

I.4.5. Traitement du signal biomédical

Il s'agit d'un domaine d'application très encourageant pour les techniques BSS, non seulement parce que c'est un domaine de croissance rapide et de grande importance, mais aussi parce que certains types de données d'imagerie cérébrale semblent être assez bien décrits par le modèle BSS.

En particulier pour l'électrocardiogramme (ECG), l'EEG, l'électromyogramme (EMG), le MEG, ainsi que les signaux IRM (Imagerie par Résonance Magnétique fonctionnelle), qui sont des enregistrements de champs électriques et magnétiques de signaux provenant de courants neuronaux dans le cerveau [67]. Par exemple, dans le traitement du signal ECG fœtal, le signal ECG mesuré par une femme enceinte comprend en fait les signaux des femmes enceintes et du bébé à naître, tandis que les paramètres du milieu de transmission sont inconnus, ce qui est une application typique de la technique BSS. En utilisant des méthodes BSS, nous pouvons distinguer et séparer le signal ECG d'une femme enceinte de celui du bébé à naître, fournissant ainsi des données précises pour un diagnostic ultérieur. De plus, l'ICA a été appliqué avec succès au traitement du signal IRM du cerveau pour obtenir des informations sur les caractéristiques cachées, fournissant des données utiles pour la recherche scientifique sur le comportement humain [68].

L'algorithme NMF a été appliqué à la génétique et à la découverte de médicaments. Les résultats expérimentaux ont montré que les résultats des séquences de molécules d'ADN génomique analysées par le NMF sont plus fiables. De même, le NMF peut également aider à sélectionner les ingrédients pharmaceutiques les plus efficaces avec le moins d'effets négatifs [69].

I.5. Les types de mélanges

Il existe plusieurs types de mélanges divisés en deux classes : les mélanges linéaires et les mélanges non-linéaires.

On dit que les mélanges sont linéaires instantanés quand les observations s'écrivent comme des combinaisons linéaires des sources. Dans ce cas, le nombre de sources est supposé égal au nombre d'observations. Supposons qu'il existe N sources statistiquement indépendantes passant par un système de mélange linéaire instantané et reçues par M capteurs. Ensuite, chaque signal observé est une combinaison linéaire de ces sources. Le modèle d'un mélange linéaire est donné sur la figure I.2.

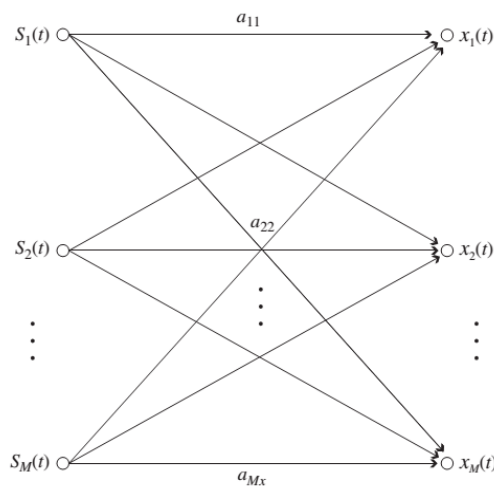


Figure I.2. Modèle du mélange linéaire

L'équation suivante régit la fonction de mélange instantané linéaire invariant dans le temps :

$$x_j(t) = \sum_{i=1}^N a_{ji} s_i(t) \quad (\text{I.1})$$

avec a_{ji} ($i \in \{1, 2, \dots, N\}, j \in \{1, 2, \dots, M\}$) sont les paramètres du mélange, $s_i(t)$ ($i \in \{1, 2, \dots, N\}$) sont les signaux de source et $x_i(t)$ ($i \in \{1, 2, \dots, M\}$) sont les signaux observés.

Il est plus facile d'utiliser la forme vectorielle pour exprimer le processus de mélange :

$$s(t) = [s_1(t) \dots \dots \dots s_N(t)]^T, x(t) = [x_1(t) \dots \dots \dots x_M(t)]^T.$$

Donc l'équation (I.1) peut être écrite comme suit :

$$x(t) = A \cdot s(t) \tag{I.2}$$

Où $A \in R^{M \times N}$ est la matrice de mélange et $s \in R^{M \times N}$ est le vecteur source dont les éléments sont les coefficients de la matrice A .

Comme le canal de transmission et le réseau de capteurs peuvent inclure du bruit additif qui est à prendre en compte, l'équation (1.2) devient :

$$x(t) = A \cdot s(t) + n(t) \tag{I.3}$$

Où

$n(t) = [n_1(t) \dots \dots \dots n_M(t)]^T$ représente le vecteur de bruits additifs.

Si le nombre d'observations est inférieur au nombre de sources, on dit que le mélange est sous déterminé, et dans ce cas, il est impossible de résoudre le problème sans hypothèses supplémentaires [70].

Dans le cas où le nombre d'observations est supérieur aux nombres de sources, on dit que le mélange est sur-déterminé. Dans ce cas, Bermon et al. utilisent l'ACI pour résoudre le problème de séparation aveugle de sources à condition de supposer qu'il y ait au plus une source gaussienne [71].

I.5.1 Mélanges linéaires convolutifs

On parle de mélanges convolutifs lorsque les sources s'écrivent comme la somme de sources convoluées par un filtre de propagation, c.à.d. les signaux reçus au niveau

des capteurs dépendent à la fois des sources et de leurs versions retardées. Les sources peuvent contribuer au mélange avec plusieurs délais différents. Cela peut se produire dans plusieurs domaines d'applications tels que les télécommunications où les modèles de canal incluent souvent la propagation par trajets multiples puisque la réverbération en traitement audio peut introduire des retards [72]. Le modèle de mélange convolutif est plus compliqué que le modèle de mélange instantané linéaire. Cependant, il a été prouvé dans [73] que les mélanges convolutifs sont séparables, c'est-à-dire que dans un système de mélange-séparation convolutif l'indépendance des sorties assure la séparation des sources. Mais dans ce cas, l'indétermination n'est pas aussi triviale que dans le cas instantané puisque l'indétermination de la permutation existe toujours. Cette indétermination peut être inacceptable, car elle peut fortement altérer les sources. Cependant, comme l'a proposé Simon [74], après la séparation des sources, l'effet de chaque source sur chaque capteur peut être estimé.

Le modèle des mélanges convolutifs des sources est plus proche d'un modèle de mélanges réel.

Supposons qu'il existe des sources statistiquement indépendantes $s_i(t), i = 1, 2, \dots, N$ reçues par M capteurs après le processus de mélange par convolution (figure I.3).

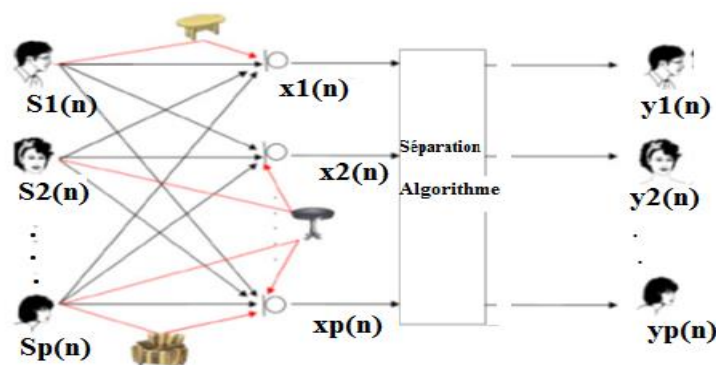


Figure I.3. Modèle des mélanges convolutifs

Les signaux mélangés sont notés par $x_i(t)$, $i = 1, 2, \dots, M$, et le modèle de mélange par convolution peut être exprimé par :

$$x_j(t) = \sum_{i=1}^N a_{ji}(t) * s_i(t) = \sum_{i=1}^N \sum_{\tau=0}^{L-1} a_{ji}(\tau) s_i(t - \tau) \quad (\text{I.4})$$

où $*$ désigne l'opération de convolution et $a_{ji}(\tau)$ est la réponse impulsionnelle du capteur j sur la source i .

Étant donné que chaque canal peut être désigné par un filtre à réponse impulsionnelle finie (FIR) du nième ordre, le système de mélange par convolution peut être représenté par la matrice FIR via l'équation (I.5).

$$x = As \quad (\text{I.5})$$

où A est une matrice FIR de la forme :

$$A = \begin{bmatrix} a_{11}^T & \dots & a_{1n}^T \\ \vdots & \ddots & \vdots \\ a_{m1}^T & \dots & a_{mn}^T \end{bmatrix} \quad (\text{I.6})$$

où a_{ji} est une colonne de dimension L qui désigne l'ordre du filtre FIR.

La forme vectorielle de l'équation (I.5) est :

$$x(t) = \sum_{\tau=0}^{L-1} A(\tau) s(t - \tau) \quad (\text{I.7})$$

où $s(t) = [s_1(t) \dots \dots \dots s_N(t)]^T$, $x(t) = [x_1(t) \dots \dots \dots x_M(t)]^T$, et A la matrice de mélange.

Pour $L=1$, le mélange convolutif devient un mélange instantané.

II.5.2. Mélanges linéaires instantanés

On parle de mélange linéaire instantané (LI) lorsque les valeurs des observations à un instant s'écrivent comme des combinaisons linéaires des valeurs des sources au même instant.

En particulier, dans le cas où les sources sont des signaux monodimensionnels notés $s_j(t)$, les observations $x_i(t)$ sont de la forme :

$$x_i(t) = \sum_{j=1}^N a_{ij} s_j(t) \quad i = 1, 2, \dots, P \quad (\text{I.8})$$

Où a_{ij} est un réel constant qui correspond à l'amplitude de la contribution de la source j sur le capteur i . Ce modèle est bien adapté à divers signaux de télécommunications ou signaux biomédicaux par exemple. Il n'est pas par contre représentatif des réalités acoustiques dans le cadre de mélanges réels. Cependant, sous réserve qu'aucun effet de post-traitement ne soit appliqué, les mélanges LI modélisent généralement bien les mélanges de sources de parole ou de musique, réalisés sous table de mixage, en studio.

I.6 Conclusion

Nous avons exposé les éléments théoriques de la séparation aveugle de sources. Nous avons aussi introduit les notions des signaux audio. Dans les chapitres suivants nous présenterons notre travail qui porte sur l'optimisation avec PSO dans le but de faire la séparation aveugle de sources audio (parole et musique).

CHAPITRE II

Signal Audio

II.1. Introduction

La parole peut être décrite comme le résultat de l'action volontaire et coordonnée d'un certain nombre de muscles. Cette action se déroule sous le contrôle du système nerveux central qui reçoit en permanence des informations par rétroaction auditive et par les sensations Kinesthésiques. Le signal parole est de nature très redondante; ce qui explique sa résistance aux différentes conditions de transmission. C'est cette qualité qui assure la pérennité de l'utilisation du signal parole dans les télécommunications. Néanmoins, dans plusieurs applications, la principale préoccupation est d'extraire des messages acoustiques les paramètres significatifs caractérisant le signal vocal afin de réduire la redondance de l'information. Ce volume réduit de données rend le traitement envisagé, tels que la reconnaissance de la parole ou l'identification de locuteur, plus facile [75].

II.2. Classification des signaux

Nous pouvons classer les signaux en deux classes principales :

Signaux déterministes dont l'évolution peut être représentée grâce à une fonction mathématique. Un signal déterministe peut être périodique ou non périodique.

Signaux aléatoires qui peuvent être caractérisés mathématiquement mais leurs évolution ne peut pas être devinée à un instant donné. Un signal aléatoire peut être stationnaire ou non stationnaire.

II.3. Le signal vocal

Normalement, la voix est transmise à l'oreille humaine au moyen d'une onde acoustique voyageant dans l'air à la vitesse du son. Les voix sont des choses importantes pour les humains. Ils sont le moyen par lequel nous communiquons beaucoup avec le monde extérieur : nos idées, bien sûr, mais aussi nos émotions et notre personnalité. La voix est même symbole du locuteur, inventé de manière éternelle dans le tissu de la parole. En ce sens, chacun de nos énoncés de la langue

parlée porte non seulement son propre message mais aussi, par l'accent, le cri de la voix et la qualité habituelle de la voix, il est en même temps une déclaration audible de notre appartenance à des groupes sociaux régionaux particuliers, de notre identité physique et psychologique individuelle, et de notre humeur momentanée. Les voix sont également l'un des médias par lesquels nous reconnaissons (avec succès, la plupart du temps) d'autres humains qui sont importants pour nous - les membres de notre famille, les personnalités des médias, nos amis et nos ennemis.

II.4. Système phonatoire et modèle de production

L'information portée par le signal de parole peut être analysée de bien des façons. On en distingue généralement plusieurs niveaux de description non exclusifs : acoustique, phonétique, phonologique, morphologique, syntaxique, sémantique et pragmatique [76].

II.4.1 Le niveau acoustique

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 au 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 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. L'opération de numérisation, schématisée à la figure I.1, requiert: un filtrage de garde, un échantillonnage, et une quantification.

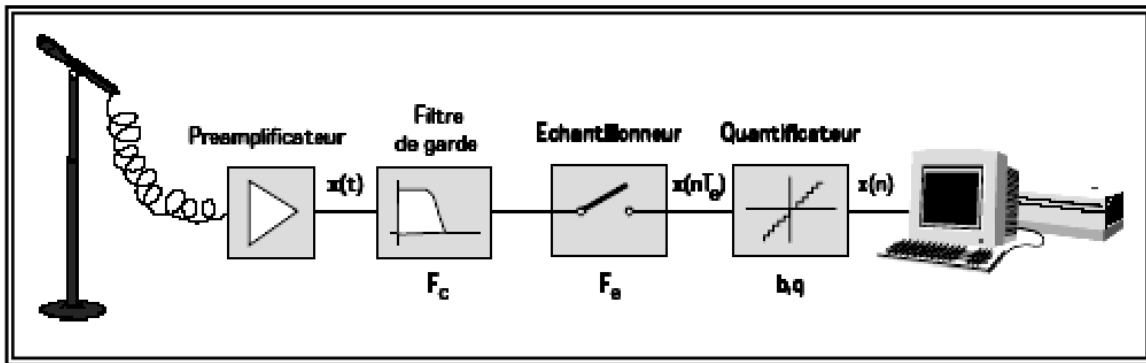


Figure II.1 Enregistrement acoustique d'un signal acoustique

II.4.1.1 Spectrogramme [77]

Il est souvent intéressant de représenter l'évolution temporelle du spectre à court terme d'un signal, sous la forme d'un spectrogramme. L'amplitude du spectre y apparaît sous la forme de niveaux de gris dans un diagramme à deux dimensions temps fréquence. On parle de spectrogramme à large bande ou à bande étroite selon la durée de la fenêtre de pondération. Les spectrogrammes à bande large sont obtenus avec des fenêtres de pondération de faible durée (typiquement 10 ms); ils mettent en évidence l'enveloppe spectrale du signal, et permettent par conséquent de visualiser l'évolution temporelle des formants. Les périodes voisées y apparaissent sous la forme de bandes verticales plus sombres. Les spectrogrammes à bande étroite sont moins utilisés. Ils mettent plutôt la structure fine du spectre en évidence: les harmoniques du signal dans les zones voisées y apparaissent sous la forme de bandes horizontales.

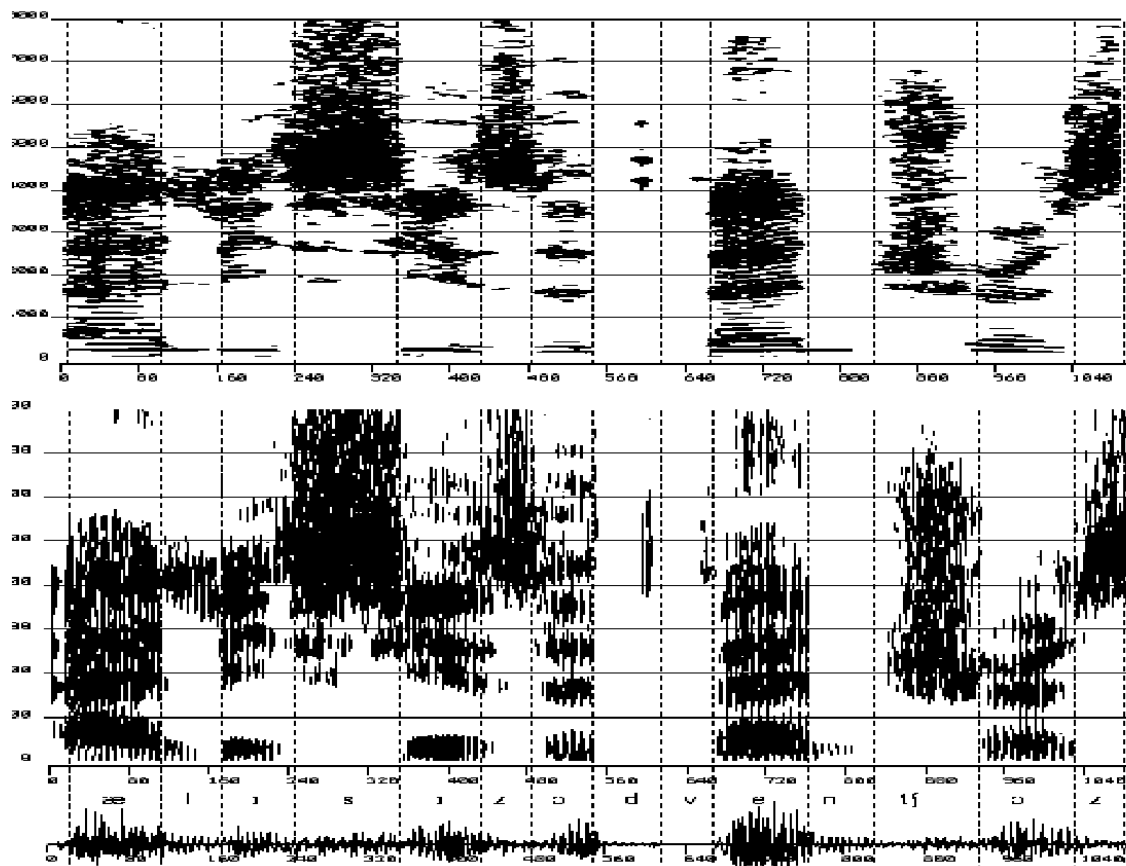


Figure II. 2 : Spectrogramme à large bandes (en bas), à bande étroite (en haut) et évolution temporelle de la phrase anglaise 'Alices's adventures'

II.4.1.2. Sons voisés et Sons non - voisés

Les cordes vocales vibrent à une certaine fréquence, appelée fréquence fondamentale ou pitch. L'excitation est alors quasi périodique et son spectre de puissance a une structure harmonique. La fréquence fondamentale varie typiquement entre 75 et 400 Hz [78]. Pour des sons non -voisés, les cordes vocales ne vibrent pas et le spectre de l'excitation ressemble à un spectre de bruit blanc. La figure 1.3 (a) donne une représentation d'un son voisé. La figure 1.3 (b) donne une représentation d'un son non voisé. Temps (échantillons ,1échantillons =0.125ms).

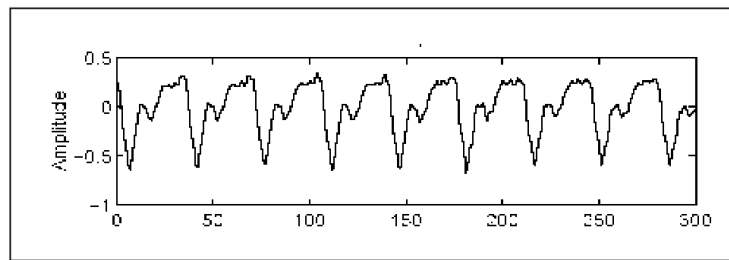


Figure II. 3 (a) : Son voisé

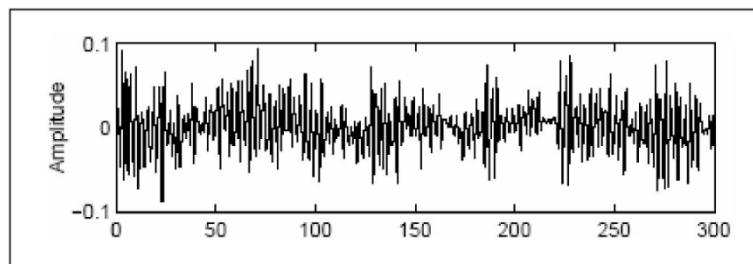


Figure II. 3 (b) : Son nonvoisé

II.4.2. Le niveau phonétique [79]

Au contraire des acousticiens, ce n'est pas tant le signal qui intéresse les phonéticiens que la façon dont il est produit par le système articulatoire, présenté à la figure 1.4, et perçu par le système auditif.

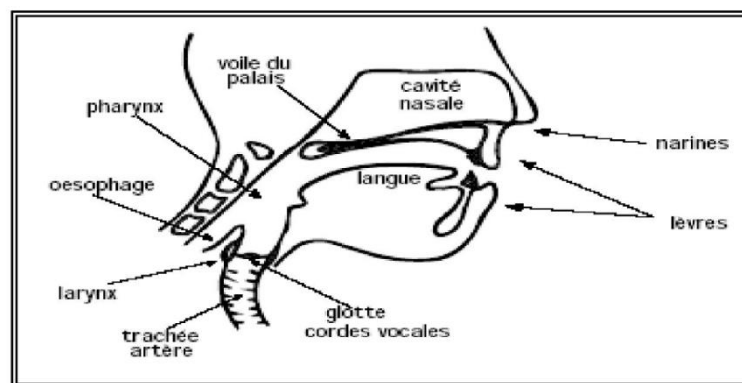


Figure II.4 : L'appareil phonatoire

II.4.2.1 Phonation

La parole peut être décrite comme le résultat de l'action volontaire et coordonnée d'un certain nombre de muscles. Cette action se déroule sous le

contrôle du système nerveux central qui reçoit en permanence des informations par rétroaction auditive et par les sensations kinesthésiques. Au sommet de celle-ci se trouve le larynx ou Le larynx est un ensemble de muscles et de cartilages mobiles qui entourent une cavité située à la partie supérieure de la trachée la figure 1.6. Les cordes vocales sont en fait deux lèvres symétriques placées en travers du larynx.

Ces lèvres peuvent fermer complètement le larynx et, en s'écartant progressivement, déterminer une ouverture triangulaire appelée glotte. L'air y passe librement pendant la respiration et la voix chuchotée, ainsi que pendant la phonation des sons non - voisés (ou sourds). Les sons voisés (ou sonores) résultent au contraire d'une vibration périodique des cordes vocales. Le larynx est d'abord complètement fermé, ce qui accroît la pression en amont des cordes vocales, et les force à s'ouvrir, ce qui fait tomber la pression, et permet aux cordes vocales de se refermer ; des impulsions périodiques de pression sont ainsi appliquées au conduit vocal, composé des cavités pharyngienne et buccale pour la plupart des sons. Lorsque la luvette est en position basse, la cavité nasale vient s'y ajouter en dérivation. Notons pour terminer le rôle prépondérant de la langue dans le processus phonatoire. Sa hauteur détermine la hauteur du pharynx : plus la langue est basse, plus le pharynx est court. Elle détermine aussi le lieu d'articulation, région de rétrécissement maximal du canal buccal, ainsi que l'aperture, écartement des organes au point d'articulation.

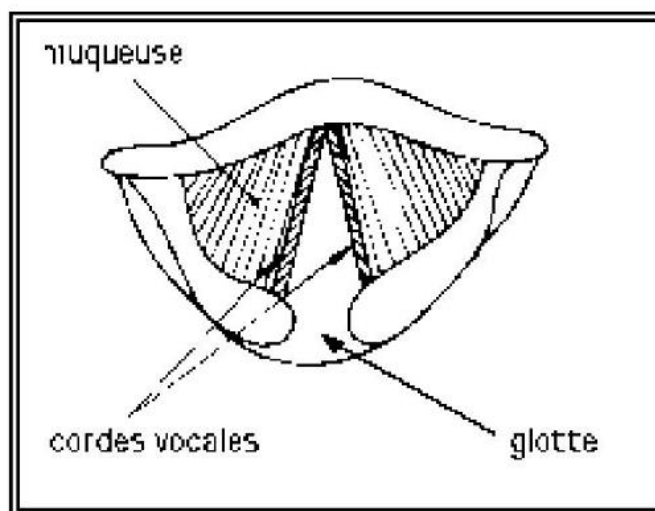


Figure II.5 : Section du larynx, vu de haut

ce qui met en parallèle les cavités nasales et buccale. Notons que, dans un contexte plus général que celui de la seule langue française, d'autres critères peuvent être nécessaires pour différencier les voyelles, comme leur labialisation, leur durée, leur tension, leur stabilité, leur glottalisation, voire même la direction du mouvement de l'air.

Ainsi, par exemple, on peut approximativement représenter les voyelles dans le plan des deux premiers formants (Figure(1.6)). On observe en pratique un certain recouvrement dans les zones for-mantiques correspondant à chaque voyelle (un affichage en trois dimensions figurant les trois premiers formants permettrait une meilleure séparation).

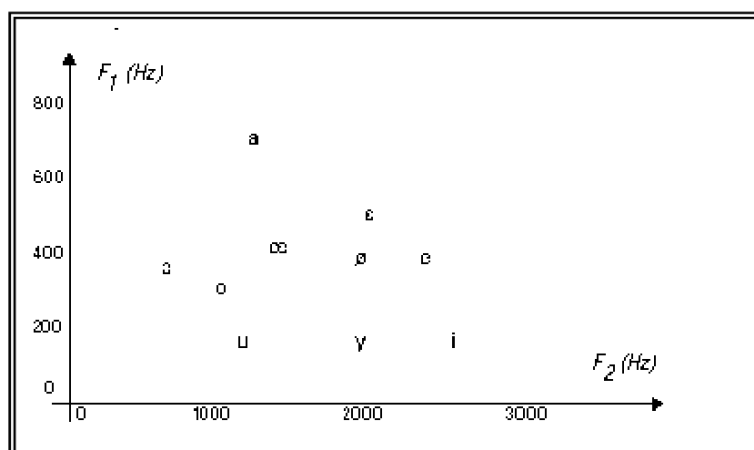


Figure II.6 : Représentation des voyelles dans plan F1-F2

II.4.2.3 Audition perception [79]

Dans le cadre du traitement de la parole, une bonne connaissance des mécanismes de l'audition et des propriétés perceptuelles de l'oreille est aussi importante qu'une maîtrise des mécanismes de production. En effet, tout ce qui peut être mesuré acoustiquement ou observé par la phonétique articulatoire n'est pas nécessairement perçu. Par ailleurs, nous avons déjà souligné, le rôle fondamental que joue l'audition dans le processus même de production de la parole.

Les ondes sonores sont recueillies par l'appareil auditif, ce qui provoque les sensations auditives. Ces ondes de pression sont analysées dans l'oreille interne qui envoie au cerveau l'influx nerveux qui en résulte ; le phénomène physique induit ainsi un phénomène psychique grâce à un mécanisme physiologique complexe. L'appareil auditif comprend l'oreille externe, l'oreille moyenne, et l'oreille interne (figure 1.7). Le conduit auditif relie le pavillon au tympan : c'est un tube acoustique de section uniforme fermé à une extrémité, son premier mode de résonance est situé vers 3000 Hz, ce qui accroît la sensibilité du système auditif dans cette gamme de fréquences.

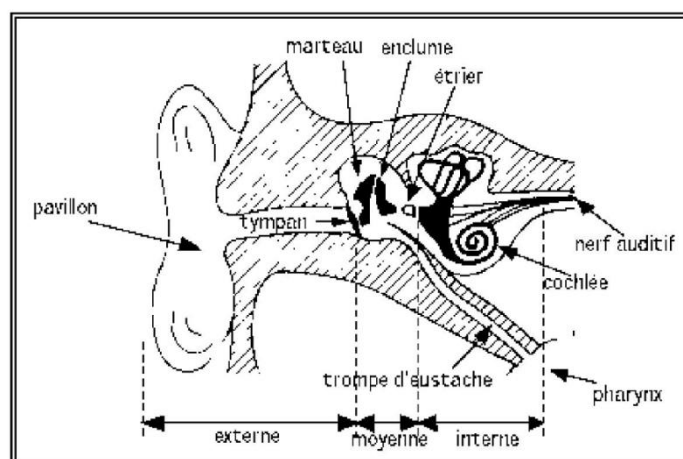


Figure II.7 : Le système auditif

II.4.3 Le niveau phonologique [79]

La phonologie (parfois appelée phonétique fonctionnelle) est l'interface nécessaire entre la phonétique et les descriptions linguistiques de niveau plus élevé. Dans les sections précédentes, nous avons décrit la parole comme si elle n'était porteuse d'aucune signification. Les sons de parole ont d'ailleurs été présentés indépendamment les uns des autres. La phonologie introduit la notion d'unité abstraite du discours (par opposition aux sons observés, perçus, ou articulés) :

Le phonème est la plus petite unité phonique fonctionnelle, c'-à- d distinctive. Il n'est pas défini sur un plan acoustique, articulatoire, ou perceptuel, mais bien sur le plan fonctionnel. Ainsi, les phonèmes n'ont pas d'existence indépendante: ils constituent un ensemble structuré dans lequel chaque élément est intentionnellement différent de tous les autres, la différence étant à chaque fois porteuse de sens. La liste des phonèmes pour la plupart des langues européennes a été établie dès la fin du 19^e siècle sur la base de l'étude de paires minimales, composées de paires de mots différant par un seul son, lequel suffit à changer leur sens (ex :[nlmàç_èè--] dans 'bon-pont').

II.4.4 Le niveau morphologique [80]

La suite des phonèmes prononcés correspond à des mots, choisis dans le lexique des mots de la langue. Si l'on sait que le Petit Robert compte à peu près 50.000 entrées et que seules les formes canoniques y sont répertoriées (masc. sing. des noms et adjectifs, infinitifs des verbes), on peut estimer la richesse lexicale d'une langue comme le français à plusieurs centaines de milliers de mots. Lorsqu'on étudie les formes écrites et phonétiques d'une langue, il est frappant de constater que les mots qui la composent, bien que très nombreux, sont eux-mêmes constitués d'unité plus petites (comme dans image, images, imagine, imagination, imagerie, ...etc). La morphologie est la branche de la linguistique qui étudie comment les formes lexicales sont obtenues à partir d'un ensemble réduit d'unités porteuses de sens, appelées morphèmes.

II.5. Application du signal parole [81] [82]

Les applications du signal parole sont très nombreuses. On peut citer entre autre :

Les aides aux personnes handicapées: A l'aide d'un clavier spécialement adapté et/ou d'un logiciel d'assemblage de phrases, un handicapé peut s'exprimer par la voix de son synthétiseur.

Outils d'enseignement assisté par ordinateurs : Les synthétiseurs peuvent répéter deux fois la même chose. Ils peuvent donc servir à valider des théories relatives à la production, à la perception ou à la compréhension de la parole

Applications industrielles : L'utilisateur d'une voix de synthèse dans un centre de contrôle de site industriel permet d'attirer l'attention du personnel de surveillance sur un problème urgent. De même, l'intégration d'un synthétiseur dans une cabine de pilotage, d'un avion permet d'éviter au pilote d'être dépassé par la quantité d'information visuelle qu'il a analysée.

Communication homme machine : Le développement de synthétiseur de haute qualité permettra à l'homme de communiquer avec la machine de manière plus naturelle.

Application grand public non téléphonique : Le marché du jouet est déjà touché par la synthèse vocale. Des ordinateurs pour enfants possédant une sortie vocale qui augmente l'attrait chez les jeunes enfants.

Télématique vocale : Dans ce domaine, on peut citer :

- Les serveurs vocaux d'informations ;
- Les serveurs de lecture vocale de fax ou de messages électroniques ;
- Automatisation de services de, prise de commande (vente par correspondance) ;
- Automatisation de services de renseignement (annuaire, standard d'entreprise, etc...).
- Programmation sans clavier : On peut s'adresser à une machine et lui dicter un programme par un combiné téléphonique, à travers un micro servant de terminal et ceci grâce à un système de reconnaissance. Contrôle de qualité, saisie de données. Dans de nombreux environnement de travail, la possibilité de décharger le travailleur, grâce à une interface vocale, apporte un gain incontestable de liberté et de rapidité de mouvement.

II.6. Conclusion

Ce chapitre a présenté un récapitulatif sur les notions du système phonatoire, le système d'audition et des caractéristiques du signal parole. Différentes applications de la parole ont été également citées. Le prochain chapitre sera réservé aux algorithmes évolutionnaires.

CHAPITRE III

Algorithmes Evolutionnaires

III.1 Introduction

De point de vue informatique, les algorithmes évolutionnaires (AE) font partie de la famille des algorithmes d'optimisation stochastiques dont la mise en place est complexe et le coût de calcul est important. Ils sont particulièrement destinés à traiter des problèmes qui n'ont pas de solutions classiques pour la résolution de problèmes où les algorithmes classiques d'optimisation, d'apprentissage ou de conception automatique sont incapables de produire des résultats satisfaisants.

La fonction f à optimiser, appelée aussi performance, est définie sur un espace de recherche D . L'algorithme fait évoluer une population, un sous-ensemble de l'espace de recherche. L'idée fondamentale est que la sélection favorise les individus qui optimisent la performance et que les variations font apparaître, dans la population sélectionnée, des individus qui peuvent être meilleurs vis à vis de la performance. Dans cette évolution, les générations successives de la population restent à taille constante et l'aspect stochastique ne dépend que de la génération précédente.

L'objectif de ce chapitre est donc de présenter les notions relatives aux algorithmes évolutionnaires et spécifiquement à la méthode d'optimisation utilisée dans notre travail, la PSO.

III.2. Les algorithmes évolutionnaires

Cette théorie explique comment, depuis l'apparition de la vie, les espèces ont su évoluer de générations en générations dans le sens d'une meilleure adaptation des individus à leur environnement, en favorisant la survie et la reproduction des individus les plus adaptés.

Les AE ne s'agissent pas de trouver une solution exacte, mais ils s'agissent de trouver des solutions satisfaisant au mieux différents critères, souvent contradictoires. S'ils

ne permettent pas de trouver à coup sûr la solution optimale de l'espace de recherche, du moins on peut constater que les solutions fournies sont généralement meilleures que celles obtenues par des méthodes plus classiques, pour un même temps de calcul.

III.3. Classes des AE

On distingue quatre grandes familles d'AE qui ont isolément été développées par des grandes écoles [83],[84].

III.3.1. La programmation évolutionnaire (PE)

Elle était initialement mise au point pour la découverte d'automates à états finis pour l'approximation de séries temporelles. Plus tard, la PE a été généralisée à des espaces de recherche très variés. Ce modèle évolutionniste accentue l'utilisation de la mutation et n'utilise pas dans sa version originale la recombinaison des individus par croisement. Il est souvent appliqué à la résolution de problèmes d'optimisation à variables réelles.

L'opérateur de mutation consiste à ajouter une perturbation Gaussienne à chaque composante du vecteur à variables réelles constituant l'individu. Cette perturbation est basée sur la performance de l'individu. L'idée consiste à faire subir des mutations importantes aux mauvais individus et inversement des mutations faibles aux bons individus. L'opérateur de sélection est de type probabiliste utilisant la méthode du tournoi basée sur une compétition entre individus choisis aléatoirement.

III.3.2. Les stratégies d'évolution (SE)

Elles ont été développées par I. Rechenberg et H.P. Schwefel, 1965 à Berlin. La technique des SE a été mise au point par ces deux jeunes élèves ingénieurs travaillant sur des problèmes d'optimisation de tuyères. Ce modèle utilise le principe de mutation sur les réels du modèle de la PE. Leur principe a été affiné de sorte que la fonction de perturbation Gaussienne est contrôlée par l'ensemble de la population

courante. Si la proportion de mutation réussie est élevée, l'espace de recherche exploré est restreint autour d'un optimum local, il faut donc diversifier la population en augmentant le taux de mutation. Ces approches utilisent un opérateur de sélection de type déterministe : les solutions dont la fitness est mauvaise sont éliminées de la population. En outre, dans le modèle originel, les populations des parents et de leurs descendants sont généralement de taille différente.

III.3.3. Les algorithmes génétiques (AG)

L'algorithme génétique (AG) est un algorithme de recherche basé sur les mécanismes de la sélection naturelle et de la génétique. Il combine une stratégie de survie des plus forts avec un échange d'information aléatoire mais structuré. L'algorithme génétique de base comporte trois opérations simples qui ne sont pas plus compliquées que des opérations algébriques : Sélection, Reproduction et Mutation.

III.3.4. La programmation génétique (PG)

Proposée par J. Koza en 1988 qui était auparavant un sous-groupe des AG. La spécificité de la PG est son espace de recherche, un espace de programmes le plus souvent représentés sous forme d'arbres.

Ces quatre classes d'algorithmes ne se différencient principalement que sur le détail d'implantation des opérateurs et des procédures de sélection et remplacement dans la population.

III.4. Structure des AE

Un AE est un algorithme itératif qui possède une structure bien définie. Il procède par différentes étapes en vue de résoudre le problème d'optimisation.

1. Initialisation de population

C'est une première phase de l'algorithme, dans laquelle la population initiale est construite de manière aléatoire ou à travers des résultats issus d'autres techniques d'optimisation.

2. Evaluation de la population

Elle consiste à calculer la valeur de la fonction de coût pour chaque individu.

3. La sélection des individus

Elle consiste à choisir les individus les plus adaptés de la population courante afin de former la nouvelle génération.

4. La reproduction des individus

C'est la phase dans laquelle une nouvelle population est construite à partir des individus sélectionnés, via des opérateurs de reproduction (tels que le croisement et la mutation pour les AG). En notant que cette étape est souvent stochastique.

5. Arrêt de l'AE

Il s'agit d'un test de l'efficacité de l'algorithme, à travers une valeur de la fonction objectif à atteindre, le nombre d'itérations ou le temps d'exécution. La solution courante est prise quand ce test est vérifié ; sinon, l'algorithme passe à l'itération suivante à partir de l'étape d'évaluation.

III.5. Optimisation par la méthode PSO

L'optimisation par essaim de particules est un algorithme évolutionnaire qui utilise une population de solutions candidates pour développer une solution optimale au problème en question. Cet algorithme a été proposé par Russel Eberhart (ingénieur en électricité) et James Kennedy (socio-psychologue) en 1995 [85]. Il s'inspire à l'origine du monde du vivant, plus précisément du comportement social des animaux évoluant en essaim, tels que les bancs de poissons et les vols groupés d'oiseaux. L'essaim de particules correspond à une population d'agents simples, appelés particules. Chaque particule dispose d'un groupe de particule dites informatrices, appelé son voisinage. Chaque particule est considérée comme une solution du problème, où elle possède une position et une vitesse, $x(t)$ et $v(t)$ respectivement.

De plus, chaque particule possède une mémoire lui permettant de se souvenir de sa meilleure performance et de la meilleure performance atteinte par les particules informatrices.

La particule change son état sous l'influence de trois composantes :

1. Une composante *d'inertie* qui exprime son propre inertiel : la particule tend à suivre sa direction courante de déplacement.

2. Une composante *cognitive* qui exprime la position personnelle la plus optimale: la particule tend à se diriger vers le meilleur site par lequel elle est déjà passée (personal best, pbest).

3. Une composante *sociale* qui exprime la position la plus optimale de l'essaim : la particule tend à se fier à l'expérience de ses congénères et, ainsi, à se diriger vers le meilleur site déjà atteint par ses voisins (global best, gbest).

Ces composantes déterminent une nouvelle position de chaque particule. À chaque itération, la fonction fitness de chaque particule est évaluée avec la fonction de coût.

Les équations de mise à jour de chaque particule sont données par:

$$v_i(t+1) = wv_i(t) + c_1r_1(t)(pbest_i(t) - x_i(t)) + c_2r_2(t)(gbest_i(t) - x_i(t)) \quad (\text{III.1})$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (\text{III.2})$$

Avec : pour une itération donnée :

- w représente le coefficient d'inertie, il exprime le poids inertiel attaché à la position atteinte à l'itération précédente de la particule.
- c_1 et c_2 représentent les constantes d'accélération.
- r_1 et r_2 représentent des nombres aléatoires choisis dans $[0,1]$.

Les trois composantes mentionnées ci-dessus sont représentées dans l'équation (III.1) par les termes suivants :

1. wv correspond à la composante d'inertie du déplacement, où le paramètre w contrôle l'influence de la direction de déplacement sur le déplacement futur ;

2. $c_1 r_1 (p_{best} - x)$ correspond à la composante cognitive du déplacement, où le paramètre c_1 contrôle le comportement cognitif de la particule ;

3. $c_2 r_2 (g_{best} - x)$ correspond à la composante sociale du déplacement, où le paramètre c_2 contrôle l'aptitude sociale de la particule.



Figure III.1. Algorithme de la PSO

L'algorithme PSO présenté sur la figure III.1 peut se résumer en :

Etape 1: Initialiser une population de particules et de vitesses, uniformément réparties dans l'espace de recherche, et fixer le critère d'arrêt.

Etape 2: Evaluer la fonction objectif, pour chaque particule.

Etape 3: Mettre à jour la meilleure position pour chaque particule et la meilleure position globale dans la population.

Etape 4: Mettre à jour la position et la vitesse utilisant les équations (III. 1) et (III. 2).

Etape 5: A chaque fois qu'une des particules, suite à l'algorithme est déplacée vers la limite ou à l'extérieur de l'espace de recherche prédéfini, le mécanisme de confinement des particules est appliqué et ramène de nouveau la particule à l'intérieur de l'espace de recherche.

Etape 6: Vérifier le critère d'arrêt. Si le critère d'arrêt n'est pas satisfait, aller vers l'étape 2; Sinon, le programme se termine, et la solution optimale est produite

L'étape 5 étant facultative au cas où nous nous intéressons à améliorer la méthode d'optimisation PSO ainsi utilisée.

III.6. Mécanisme de confinement [86]

Il est possible que le déplacement d'une particule soit trop rapide et la conduise à sortir de l'espace de recherche, nous pouvons introduire un nouveau paramètre V_{max} , qui permet de limiter la vitesse sur chaque dimension et ainsi de contrôler l'explosion du système. Notons que cela ne restreint pas les valeurs de x_i à l'intervalle $[V_{min}, V_{max}]$, mais limite seulement la distance maximale qu'une particule va parcourir au cours d'une itération. Cette méthode permet de contrôler la divergence de l'algorithme et de réaliser ainsi un compromis efficace entre intensification et diversification.

De plus, une stratégie de confinement des particules peut être introduite. Une telle stratégie permet de ramener une particule sortie de l'espace de recherche D à l'intérieur de celui-ci à nouveau.

Dans ce cadre, plusieurs méthodes peuvent être employées :

- La particule est laissée à l'extérieur de l'espace de recherche, mais nous n'évaluons pas sa fonction objectif. Ainsi, elle ne pourra pas attirer les autres particules en dehors de l'espace de recherche D ;
- La particule est arrêtée à la frontière et les composantes associées à sa vitesse sont annulées ;
- La particule rebondit sur la frontière. La particule est stoppée sur la frontière, mais les composantes correspondantes de la vitesse sont multipliées par un coefficient tiré aléatoirement dans l'intervalle $[-1,0]$.

III.7. Conclusion

Dans le chapitre prochain nous allons exposer les résultats de simulations de la technique de séparation aveugle de sources audio (musique et parole) avec la méthode d'optimisation PSO.

CHAPITRE IV

**Séparation Aveugle de Source
Audio avec PSO_ Résultats de
Simulations**

IV.1. Introduction

Nous avons appliqué l'algorithme PSO dans l'optimisation du problème de séparation aveugle de sources audio. Pour analyser l'efficacité de la technique PSO sur la séparation des signaux de parole et de musique, nous avons réalisé des expériences et nous avons par la suite évalué les performances par des facteurs de test utilisés en littérature.

IV.2. Les signaux source de test

Dans la première simulation, nous avons pris comme signaux source test deux signaux musicaux.

Dans la deuxième simulation, nous avons pris deux signaux vocaux.

Enfin, dans la troisième simulation, nous avons pris un mélange; un signal musical et un signal vocal.

IV.3. Evaluation de la fonction objective

La fonction de fitness adoptée est basée sur le kurtosis et l'information mutuelle. Kurtosis joue un rôle important sur le BSS, utilisé pour mesurer la non gaussianité du signal et trier les composants indépendants. Les informations mutuelles mesurent la dépendance entre les composants. Pour parvenir à l'indépendance des composants, l'information mutuelle doit être minimale. L'information mutuelle est définie comme suit :

$$I(y_1 y_2, \dots, y_n) = \sum_{i=1}^n H(y_i) - H(y) \quad (\text{IV.1})$$

Où $H(y_i) = -E \log p_{y_i}(y_i)$ et $H(y) = -E \log p_y(y)$ sont l'entropie des signaux de source audio estimés, E est l'espérance mathématique et p_y la densité de y .

Le kurtosis des signaux de source audio estimés peut être calculé à l'aide de la formule suivante :

$$\text{kurtosis}(y) = \sum_{i=1}^n |E(y_i^4) - 3E^2(y_i^2)| \quad (\text{IV.2})$$

La fonction fitness peut être exprimée par :

$$J(\mathbf{y}) = I(y_1 y_2, \dots, y_n) + \frac{1}{\sum_{i=1}^n |E(y_i^4) - 3E^2(y_i^2)|} \quad (\text{IV.3})$$

Lorsque $J(\mathbf{y})$ est maximisée, les signaux de source audio estimés sont mutuellement indépendants.

IV.4. Mesure de performances [87]

Les performances de la technique simulée sont évaluées à l'aide de trois indices :

1. **Source to distortion ratio (SDR)**
2. **Valeur absolue des coefficients de corrélation (AVCC)**
3. **Erreur quadratique moyenne (MSE)**

Ils sont, respectivement définis comme suit :

$$\text{SDR}(y_i, s_i) = 10 \log \left[\frac{\sum_{i=1}^M [y_i(t)]^2}{\text{MSE}} \right] \quad (\text{IV.4})$$

$$\text{AVCC} = \left| \frac{\sum_{i=1}^M [y_i(t) s_i(t)]}{\sqrt{\sum_{i=1}^M [y_i(t)]^2 \sum_{i=1}^M [s_i(t)]^2}} \right| \quad (\text{IV.5})$$

$$\text{MSE} = \sum_{i=1}^M [y_i(t) - s_i(t)]^2 / M \quad (\text{IV.6})$$

Où S_i représentent les sources originales, y_i sont leurs signaux estimés avec la méthode adoptée et M est le nombre d'échantillons.

IV.5. Résultats

Simulation1: Les signaux originaux (Music_Herbalizer.wav) et (Music_Pianos.wav) et leurs signaux estimés par PSO sont montrés dans les figure IV.1, figure IV.2.

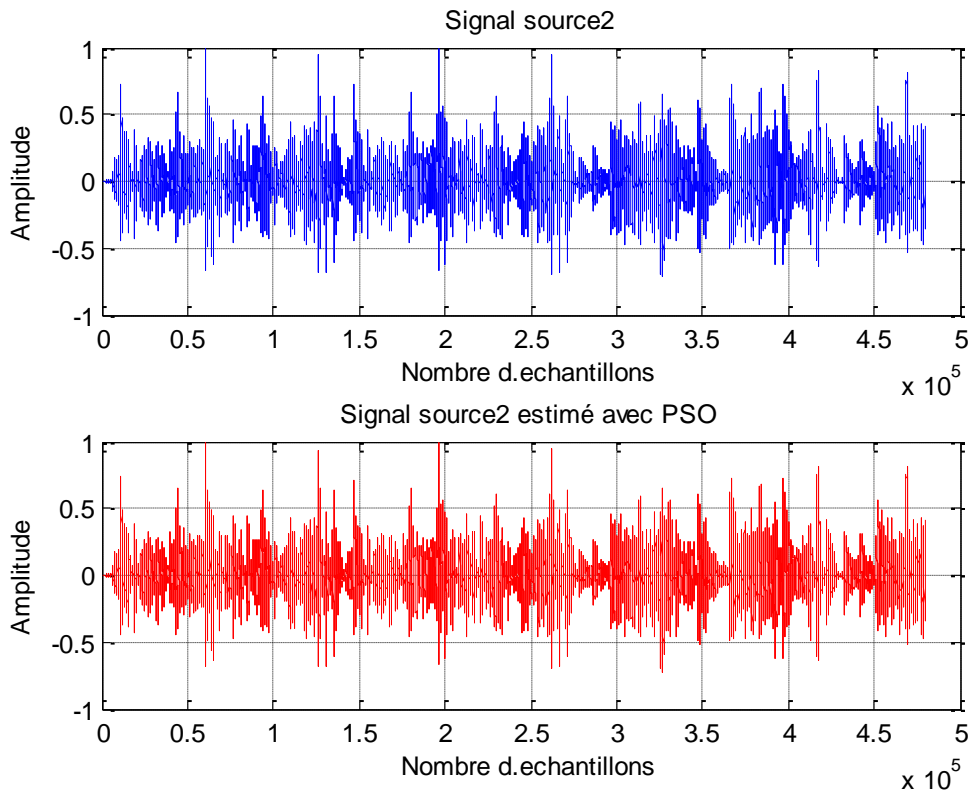


Figure IV.1. Signal original et signal estimé par PSO_(Music_Herbalizer.wav)

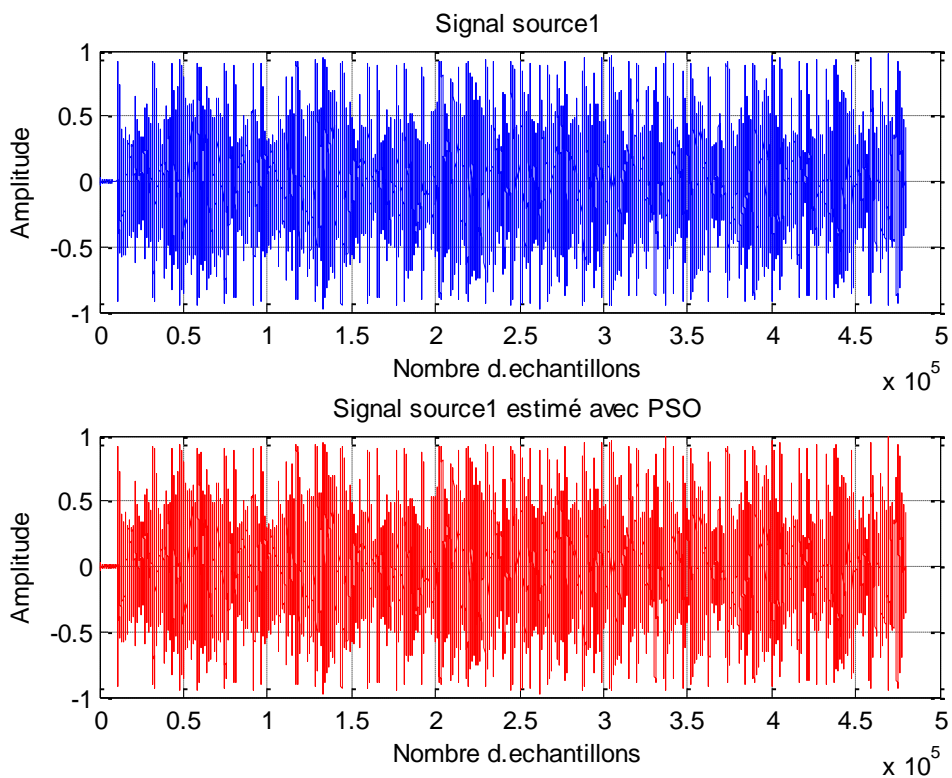


Figure IV.2. Signal original et signal estimé par PSO_(Music_Piano.wav)

Simulation2:Les signaux originaux (Paroles1.wav) et (Paroles2.wav) et leurs signaux estimés par PSO sont montrés dans les figure IV.3,figure IV.4.

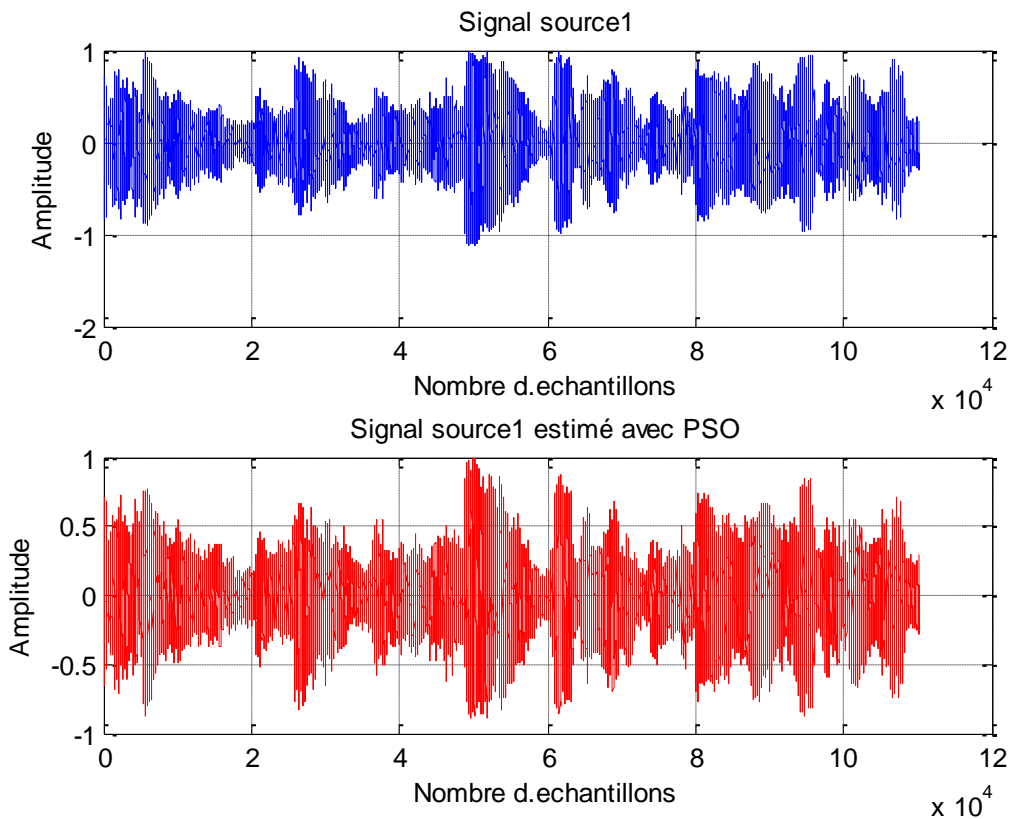


Figure IV.3. Signal original et signal estimé par PSO_(Paroles1.wav)

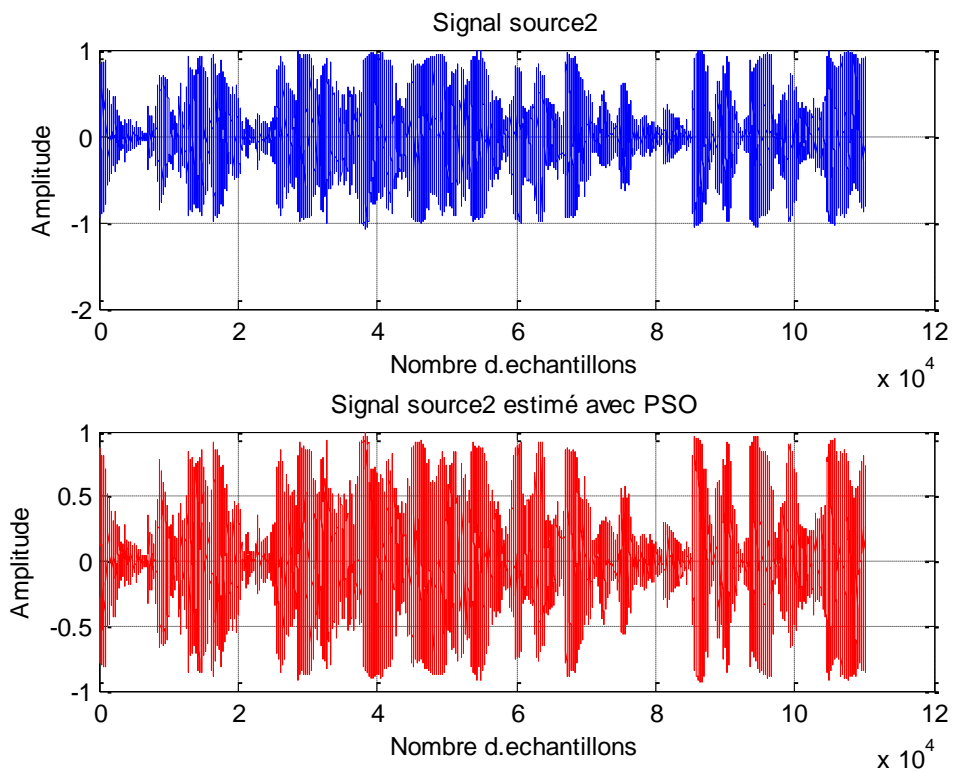


Figure IV.4. Signal original et signal estimé par PSO_(Paroles2.wav)

Simulation3: Les signaux originaux (Paroles2.wav) et (Music_Herbalizer.wav) et leurs signaux estimés par PSO sont montrés dans les figure IV.5, figure IV.6.

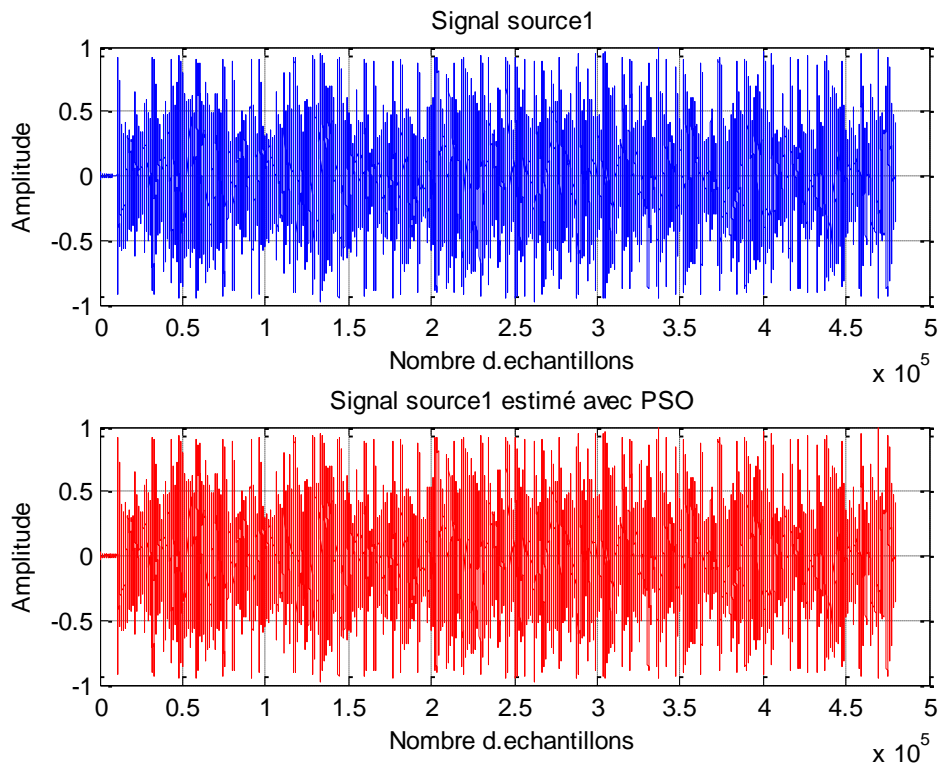


Figure IV.5. Signal original et signal estimé par PSO_ (Music_Herbalizer.wav)

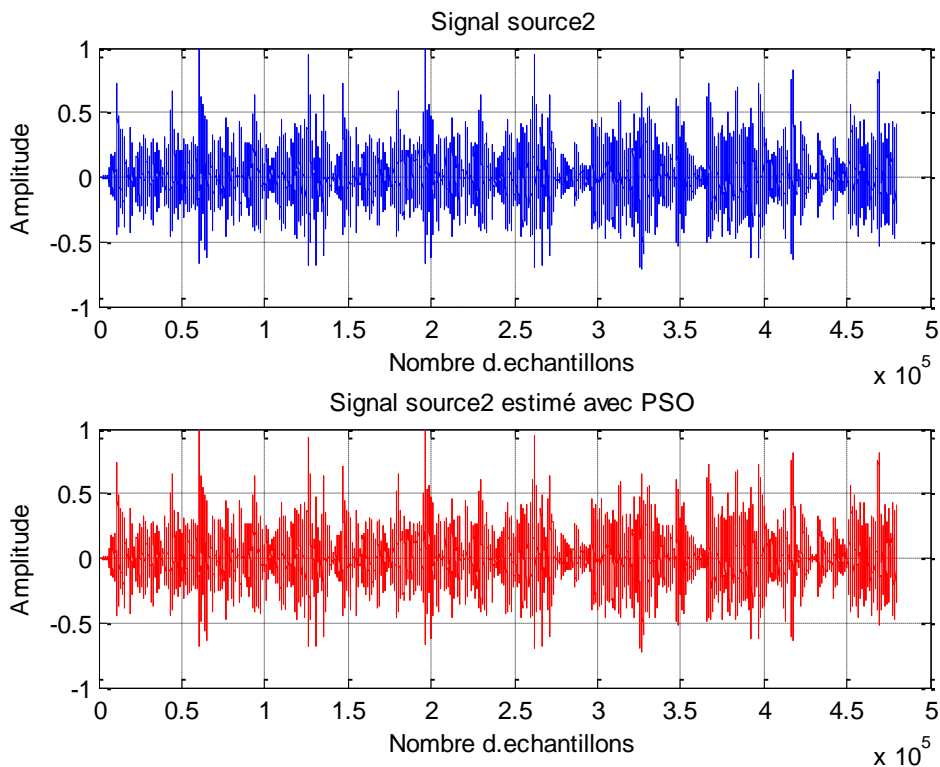


Figure IV.6. Signal original et signal estimé par PSO_ (Paroles2.wav)

IV.6. Interprétations

Qualitativement parlant, les figures obtenues suite aux simulations par l'algorithme d'optimisation PSO démontrent le degré de similarité remarquable entre les signaux estimés et leurs correspondants originaux, autant pour les signaux de parole que pour les tronçons musicaux.

Les résultats précédents sont confirmés par les indices de performance SDR, AVCC et MSE résumés dans le Tableau IV.1, ce qui démontre l'efficacité de cette méthode pour la séparation des sources audio.

Nous avons pensé à comparer nos résultats de simulation avec une des méthodes utilisées en séparation des sources audio. Nous avons choisi une méthode qui se base sur les statistiques d'ordre supérieur, connue pour son efficacité, la méthode de JADE (Joint Approximate Diagonalization of Eigenmatrices) [88].

Mais nous n'avons pas tous les fichiers audio tests pris dans [88] pour faire une comparaison objective ici.

A travers nos simulations et pour différents échantillons de fichiers audio, il paraît que la nature du signal source (parole ou music) à un effet important sur les indices de performances, ce qui le démontre la simulation 3_b.

Tableau IV.1 Nos résultats avec PSO _ Résultats avec JADE [88]

S_i	SDR (dB)	AVCC		MSE
		Y1	Y2	
JADE				
Mélange 1				
Music_Vivaldi_30s.wav	39.3474	-0.0108	0.9999	2.7508e-06
speech1.wav	43.2680	1.0000	0.0069	4.0259e-06
Mélange 2				
speech1.wav	27.9627	0.9992	-0.0420	1.3363e-04
speech2.wav	35.1048	0.0157	0.9998	1.9503e-05
Mélange 3				
Music_Vivaldi_30s.wav	43.4778	-0.0056	1.0000	2.2596e-07
Music_Mandolin_30s.wav	45.0526	1.0000	0.0068	6.9887e-07
PSO				
Simulation 1				
Music_Herbalizer.wav	79.0257	0.99986	0.075053	0.0003963
Music_Piano.wav	87.1278	-0.017689	0.99713	1.2902e-05
Simulation 2				
Paroles1.wav	66.6915	-0.99254	-0.0035614	0.0011173
Paroles2.wav	66.8241	0.12025	-1	0.0017148
Simulation 3_a				
Music_Piano.wav	89.3733	-0.010354	0.99943	1.0986e-05
Paroles2.wav	61.2686	0.99995	-0.033544	9.5851e-04
Simulation 3_b				
Music_Herbalizer.wav	67.4474	0.080954	-0.97948	0.0014835
Paroles2.wav	66.8337	-0.99619	-0.20771	0.0012439

IV.7. Conclusion

Dans ce chapitre, la méthode PSO est exploitée pour effectuer la séparation aveugle des sources audio considérées. Les sources distinctes ont été rigoureusement estimées à partir des mélanges observés. L'évaluation objective, effectuée par les métriques de performances, montre que la méthode PSO donne des résultats remarquables vis-à-vis des performances envisagées, des performances meilleures que la méthode de JADE.

CONCLUSION

GENERALE

CONCLUSION GENERALE

Ce travail fait appel à la technique de la séparation aveugle de sources qui constitue un grand champ d'intérêt de beaucoup de chercheurs à travers le monde et dans divers domaines. Nous nous sommes intéressés ici aux mélanges audio linéaires instantanés.

Dans ce mémoire, nous avons exposé les différentes approches proposées dans la littérature traitant ce problème de séparation de sources.

Nous avons par la suite introduit l'algorithme de la PSO dans le but d'améliorer les performances fournies.

Nous avons effectué des simulations sous environnement Matlab qui ont abouti à des résultats remarquables vis-à-vis des performances, aussi bien pour les signaux de parole que pour les tronçons musicaux.

Les performances sont évaluées par les indices SDR, AVCC et MSE qui ont montrés l'efficacité de la PSO dans la séparation aveugle de sources audio. La méthode PSO a abouti à de très bons résultats et les indices de performances ont été meilleurs que ceux obtenus par la méthode JADE.

Perspectives

Comme perspectives de ce modeste travail nous proposons de :

- Faire une comparaison objective entre la méthode PSO et la méthode JADE, voire même avec d'autres méthodes, en considérant les mêmes signaux tests.
- Appliquer la technique étudiée sur les modèles de mélanges non linéaires.
- Etendre la méthode aux signaux et images ECG.

- [1]J. Herault, C. Jutten, and B. Ans. Détection de grandeurs primitives dans un message composite pour une architecture de calcul neuromimétique en apprentissage non supervisé, In GRETSI, pages 1017-1022, Nice, France, Mai 1985.
- [2]J. L. Lacoume and P. Ruiz, Sources identification : a solution based on cumulants, in Proc. 4th Annual ASSP Workshop on Spectrum Estimation and Modeling, Minneapolis, MN, USA, août 1988.
- [3]J. F. Cardoso, A. Souloumiac, Jacobi angles for simultaneous diagonalization, SIAM Journal on Matrix Analysis and Applications (SIMAX), 161-4, 1996.
- [4]P. Comon, Independent component analysis, a new concept, Signal Processing, Elsevier ; 287 – 314, 1994.
- [5]M. Gaeta and J. L. Lacoume, Estimateurs du maximum de vraisemblance étendus à la séparation de sources non gaussiennes, Traitement du signal, vol. 7, no. 5, pp. 419-434, 1990.
- [6]B. Laheld and J. F. Cardoso, Séparation adaptative de sources en aveugle, implémentation complexe sans contraintes, in Actes 14ème colloque GRETSI sur le Traitement du Signal et des Images, Juan Les Pins, France, Septembre 1993, pp. 329-332.
- [7]N. Delfosse, P. Loubaton, "Adaptive blind separation of independent sources : A deflation approach, Signal Processing, Elsevier ; V. 45, 59-83, 1995.
- [8]A. Hyvarinen , E. Oja, A fast fixed-point algorithm for independent component analysis, Neural Comput 1997, 1483-92.
- [9]P. Comon, E. Moreau, Improved contrast dedicated to blind separation in communications, ICASSP 97, IEEE International Conference on Acoustics Speech and Signal Processing, Munich, 3453-3456, avril 1997.
- [10]E. Moreau, O. Macchi New self-adaptive algorithms for source separation based on contrast functions", HOS93, IEEE Signal Processing Workshop on Higher Order Statistics, Lake Tahoe, USA, 215-219, 1993.
- [11]N. Thirion, E. Moreau, New criteria for blind signal separation, IEEE Workshop on Statistical Signal and Array Processing, Pennsylvania, USA, 344-8, 2000.
- [12] L. Fety, Méthodes de traitement d'antenne adaptées aux radiocommunication, PhD, École Nationale Supérieure des Télécommunications, (ENST), 1988.
- [13]L. Tong, R. Liu, V. Soon, Y. Huang, Indeterminacy and identifiability of blind identification, IEEE Trans Circuits Systems, 499-509, 1991.
- [14]L. Tong, V. Soon, Y. Huang, R. Liu, Amuse : a new blind identification algorithm, New Orleans, LA., 1784-1786 1990.
- [15]A. Belouchrani, Séparation autodidacte de sources : algorithmes, performances et application à des signaux expérimentaux, PhD, École Nationale Supérieure des Télécommunications (ENST) Paris, 1995.
- [16]A. Belouchrani, K. Abed-Meraim, J. F. Cardoso, E. Moulines, A blind source separation technique using second-order statistics, IEEE Trans Signal Processing, 434-44, 1997.

- [17]A. Belouchrani, A. Cichocji, Robust whitening procedure in blind source separation context, *Electronics Letters, IEE Journal Electronic Engineering*, 2000.
- [18]A Ziehe, K. R. Müller, TDSEP : an efficient algorithm for blind separation using time structure, *International Conference on Artificial Neural Networks*, 675-680, 1998.
- [19]A. Belouchrani, M. G. Amin, Blind source separation based on time-frequency signal representation, *IEEE Trans Signal Processing*, 2888-2897, 1998.
- [20]C. Fevotte, *Approche temps – fréquence pour la séparation aveugle de sources non-stationnaires*, PhD, École centrale de Nantes et de l’université de Nantes, France, octobre 2003.
- [21]C. Fevotte, C. Doncarli, Two contributions to blind sources separation using time-frequency distributions, *IEEE Signal Processing Letter*, 386 – 389, 2004.
- [22]C. Fevotte, A. Holobar, C. Doncarli, D. Zazula , Single autotermes separation based on bilinear time-frequency representations, *EUSIPCO, Toulouse, France*, 565 – 568, 2002.
- [23]A. Prieto, B. Prieto, C. G. Puntonet, A. Ganas, P. Martin-Smith, Geometric separation of linear mixtures of sources : Application to speech signals, *ICA 2004, LNCS 3195*, pp. 798 – 806, 2004, Springer-Verlag Berlin Heidelberg.
- [24]C. G. Puntonet, A. Prieto, C. Jutten, M. Rodriguez-Alvarez, J. Ortega, Separation of sources : a geometry-based procedure for reconstruction of n-valued signals, *Signal Processing*, 267-2814 1995.
- [25]KI. Diamantaras Blind separation of multiple binary sources using a single linear mixture, *ICASSP 2000, IEEE International Conference on Acoustic Speech Signal Processing, Istanbul, Turkey*, 2889 – 2892, June, 2000.
- [26]D. T. Pham and J. F. Cardoso, Blind separation of instantaneous mixtures of non stationary sources, in *Proc. 2nd Symposium on Independent Component Analysis and Blind Source Separation ICA, Helsinki, Finland, juin 2000*, pp. 187-192.
- [27]D. T. Pham and J. F. Cardoso, Blind separation of instantaneous mixtures of non stationary sources, *IEEE Transactions on Signal Processing*, vol. 49, no. 9, 1837-1848, 2001.
- [28]K. Kokkinakis and A. K. Nandi, Multichannel blind deconvolution for source separation in convolutive mixtures of speech, *IEEE Trans. Audio, Speech, Lang. Proc.*, vol. 14, no. 1, pp. 200 – 212, janvier 2006.
- [29]I. Lee, T. Kim, and T. W. Lee, Complex fastica : A robust maximum likelihood approach of mica for convolutive bss, in *ICA’06*, pp. 625 – 632, 2006.
- [30]T. Kim, H. Attias, S. Y. Lee, and T. W. Lee, Blind source separation exploiting higher-order frequency dependencies, *IEEE Trans. Audio, Speech, Lang. Proc.*, vol. 15, no. 1, janvier 2007.
- [31]S. J.Qin, R. Dunia, Determining the number of principal components for best reconstruction, In *IFAC DYCOPS’98, Greece, juin1998*.
- [34]M. S. Pedersen, J. Larsen, U. Kjems, and L. C. Parra, “A survey of convolutive blind source separation methods”, *Springer Handbook of Speech*, Springer Press, septembre 2007.
- [32]E. Sourouchyari, Blind separation of source, Part III : Stability analysis, *Signal Processing*, Vol. 24, pp. 21-29, 1991.
- [33]J. C. Fort, Stabilité de l’algorithme de séparation de sources de JUTTENHERAULT, *Traitement du signal*, Vol.8, N°1, 35-42, 1991.

- [35]A. Weinstein, M. Feder, and A.-V. Oppenheim, Multi-Channel Signal Separation by Decorrelation, *IEEE Transactions on Speech and Audio Processing*, vol. 1, no. 4, pp. 405-413, octobre 1993.
- [36] L. Parra and C. Spence, Convolutional blind separation of non-stationary sources, *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 3, pp. 320-327, mai 2000.
- [37] N. Mitianoudis and M. E. Davies, Audio source separation of convolutional mixtures, *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 5, pp. 489-497, septembre 2003.
- [38]N. Murata, S. Ikeda, and A. Ziehe, An approach to blind source separation based on temporal structure of speech signals, *Riken Brain Science Institute, Technical Report 98-102*, avril 1998.
- [39]R. Mukai, S. Araki, and S. Makino, Separation and dereverberation performance of frequency domain blind source separation, *Proc. Independent Component Analysis (ICA'01)*, 230-235, décembre 2001.
- [40]S. C. Douglas, M. Gupta, H. Sawada, and S. Makino, Spatio–Temporal FastICA Algorithms for the Blind Separation of Convolutional Mixtures, *IEEE Trans. On Audio, Speech and Language Processing*, Vol. 15, no. 5, juillet 2007.
- [41]L. Nguyen-Thi, C. Jutten, and J. Caelen, Speech enhancement : Analysis and comparison of methods in various real situations, in *In J. Vandewalle, R. Boite, M. Moonen, and A. Oosterlinck, editors, Signal Processing VI, Theories and Applications*, Elsevier, Ed., 303-306, août 1992.
- [42]K. Torkkola, Blind separation of convolved sources based on information maximization, *Proc. IEEE Workshop on Neural Networks and Signal Processing*, Kyoto, Japan, 423-432, septembre 1996.
- [43]T. W. Lee, A. J. Bell, and R. Lambert, Blind separation of delayed and convolved sources, *Proc. Advances in Neural Information Processing Systems 9*, 758- 764, Cambridge, MA : MIT Press, 1997.
- [44]R. Mukai, S. Araki, H. Sawada, and S. Makino, Evaluation of separation and dereverberation performance in frequency domain blind source separation, *Journal of Acoustic Science and Technology*, vol. 25, no. 2, pp. 119-126, mars 2004.
- [45]P. Smaragdis, Information theoretic approaches to source separation, Master's thesis, MIT Media Lab, juin 1997.
- [46]P. Smaragdis, Blind separation of convolved mixtures in the frequency domain, *Neurocomputing*, vol. 22, pp. 21-34, novembre 1998.
- [47]H. Bousbia-Salah, A. Belouchrani, and K. Abed-Meraim, Jacobi-like algorithm for blind signal separation of convolutional mixtures, *IEE Electronics Letters*, vol. 37, no. 16, pp. 1049-1050, août 2001.
- [48]L. Lacoume, P. Ruiz , Separation of independent sources from correlated inputs, *IEEE Trans. on Signal Processing*, 3074-3078, décembre 1992.
- [49]W. Jun, Wide band blind identification and separation of sources, *Proceedings of Eusipco-96*, 2077-2080, Trieste, Italy, septembre, 1996.
- [50]V. Capdevielle, Ch. Servière, J. L. Lacoume, Blind separation of wideband sources in the frequency domain, *Proceedings of ICASSP-95*, vol. 2080-2083, Detroit (Michigan), USA, mai 1995.

- [51]V. Capdevielle, Ch. Servière, J. L. Lacoume, Blind separation of wideband sources : application to rotating machine signals, Proceedings of Eusipco-96, 2085–2088, Trieste, Italy, septembre, 10-13, 1996.
- [52]Ch. Servière, V. Capdevielle, Blind adaptive separation of wide-band sources, Proceedings of ICASSP-96, vol. 5, 2698–2701, Atlanta, Georgia, USA, mai, 7-10, 1996.
- [53]H. Attias and C. E. Schreiner, Blind source separation and deconvolution : The dynamic component algorithm, Neural computation, V. 10,1373–1424, 1998.
- [54] H. C. Wu and J. Principe, Simultaneous diagonalization in the frequency domain (sdif) for source separation, Proceedings of ICA 99, Aussois, France, Janvier, 11-15, 1999.
- [55]L. Parra and C. Spence, “Convulsive blind separation of non-stationary sources, IEEE Transaction on Speech and Audio Processing, 320-327, mai 2000.
- [56]A. Ghazdali, M. El Rhabi, H. Fenniri, A. Hakim, A. Keziou, Blind noisy mixture separation for independent/dependent sources through a regularized criterion on copulas , Signal Processing, 502-513, 2017.
- [57]M. El Rhabi, H. Fenniri, A. Keziou, E. Moreau, A robust algorithm for convolutive blind source separation in presence of noise, Signal Process. 93 (4) () 818-827, 2013.
- [58]A. Keziou, H. Fenniri, A. Ghazdali, E. Moreau, New blind source separation method of independent/dependent sources, Signal Process. 104 () 319-324, 2014.
- [59]Douglas, S.C. Blind Signal Separation and Blind Deconvolution, CRC Press, New York, 2002.
- [60]Vincent, E. Musical source separation using time-frequency source priors. IEEE Transactions on Speech and Audio Processing, 14(1), 91-98, 2005.
- [61]Zhang, X., and Hu, B. An application of blind signal separation in wireless digital communication. Journal of China Institute of Communications, 21(2),73-77, 2000.
- [62]Tukr, M., and Pentland, A. Face recognition using eigen faces. Proceedings of IEEE Conferences on Computer Vision and Pattern Recognition, Maui, HI, pp. 586-591, 1991.
- [63]Bartlett, M.S., Movellan, J.R., and Sejnowski, T.J.Face recognition by independent component analysis. IEEE Transactions on Neural Networks, 13(6),1450-1464, 2002.
- [64]Si, H.L., and Zhang, L.M. Intelligent Visual Image Processing, Shanghai Science and Technology Education Press, Shanghai , 2002.
- [65]Yu, X.C., Ren, J.M., and Zhang, N. An optimal ICA algorithm applied to fMRI data. Proceedings of 2007 International Conference on Machine Learning and Cybernetics, pp. 3565-3570, 2007.
- [66]Liu, X.W., Liu, H., and Zheng, T.Y. Separation of converted-wave from seismic multiples by independent component analysis. Journal of Seismology, 23(1), 11-19,2003.
- [67]Cichocki, A. and Amari, S.Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications, John Wiley & Sons, Ltd, Chichester, 2002.
- [68]Long, Z., Yao, L., Zhao, X. et al., Spatial independent component analysis of multitask-related activation in fMRI data. International Conference on Artificial Neural Networks, pp. 515-522, 2003.
- [69]Gao, Y., and Church, G, Improving molecular cancer class discovery through sparse non-negative matrix factorization. Bioinformatics, 21(21), 3970-3975,2005.
- [70]BERMOND, Olivier et CARDOSO, Jean-François, Méthodes de séparation de sources dans le cas sous-déterminé. In : 17° Colloque sur le traitement du signal et des

- images, FRA, 1999. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images, 1999.
- [71]ZHANG, L.-Q., CICHOCKI, Andrzej, et AMARI, Shun-ichi, Natural gradient algorithm for blind separation of overdetermined mixture with additive noise. *IEEE Signal Processing Letters*, vol. 6, no 11, p. 293-295, 1999.
- [72]J.F. Cardoso and A. Soudoumiac, Blind beamforming for non-Gaussian signals, *IEE Proceedings-F*, vol. 140, pp. 362-370, December 1993.
- [73]D. Yellin and E. Weinstein, Criteria for multichannel signal separation, *IEEE Trans. Signal Processing*, pp. 2158-2168, August 1994.
- [74]C. Simon, Séparation aveugle des sources en mélange convolutif, Ph. D.thesis, l'université de Marnela Vallee, Novembre 1999.
- [75]Roland Priemer, *Introductory Signal Processing*, World Scientific. p. 1. ISBN 978-9971509194. Archived from the original on 2013-06-02.
- [76]Pragnan Chakravorty, What Is a Signal? [Lecture Notes], *IEEE Signal Processing Magazine*, vol. 35, no. 5, pp. 175-177, Sept. 2018
- [77]Speech processing in embedded systems. Springer. p. 9. ISBN 978-0387755809. Archived from the original on 2013-06-02.
- [78]IEEE Transactions on Signal Processing. IEEE. Archived from the original on 2012-04-17.
- [79]L.Rabiner, M.J.Cheng, A.E.Rosenberg, and C.A.Mcgonagal, A comparative performance study of several pitch detections algorithms", *IEEE, Transactions, ASSP*, vol.24, pp.319-418, 1976.
- [80] D.Thierry, Introduction au traitement automatique de la parole, Première édition, faculté de polytechnique de Mons, 2000.
- [81] R.E Bekka, Analyse et synthèse du signal vocal appliqué à l'environnement des handicapés : application à la prévention des escarres, thèse de doctorat, université de Setif, 1994.
- [82] J.Wiley, *Speech coding Algorithms foundation and evolution of Standardized Coders*, Published simultaneously in Canada, 2003.
- [83]MAHDI SAMIR, *Optimisation Multiobjectif Par Un Nouveau Schéma De Coopération*, mémoire de Magister, Université Mentouri de Constantine, 2007.
- [84] tao.lri.fr/tiki-download_file.php?fileId=7.
- [85]Kennedy, J. and Eberhart, R. , "ParticleSwarm Optimization", In the *IEEE International Joint Conference on Neural Networks*, pages 1942-1948, Perth. 1995.
- [86]R. Eberhart, P. Simpson, and R. Dobbins. *Computational Intelligence PC Tools*. AP Professional, 1996.
- [87] Y. Yang, X. Wang, D. Zhang, Blind Source Separation Research Based on the Feature Distance Using Evolutionary Algorithms, *International Journal of Acoustics and Vibration*, Vol. 19, No 4. 1 DOI:0.20855/ijav.2014.19.4360, 2014.
- [88] W. Kadir, B. Khenache, Séparation Aveugle de Signaux Audio en Utilisant les Statistiques D'ordre Supérieur, mémoire de Master, Université Mohammed Boudiaf de M'sila, 2020.