REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE YAHIA FARES DE MEDEA

Faculté de **T**echnologie

Département du **G**énie **E**lectrique

Mémoire de fin d'études de Master en

Filière : Télécommunications.

Spécialité : Systèmes de télécommunications.

La séparation de sources audio aveugles en utilisant un banc de filtre Gammatone.

Présentés par :

HAMADI Yousra

ERRAHMANI Fedoua

Proposés et Dirigés par :

Mr. ESSAID Bilel

Promotion 2022-2023

Remerciement

En premier lieu, *nous tenons à remercier* ***LE DIEU*** *le* tout puissant qui nous a donné, durant toutes ces années, la santé, le courage et la foi pour avoir finaliser ce modeste travail.

Nos vifs remerciements et notre profonde gratitude s'adressent à notre encadreur monsieur "Mr. Essaid Billel" *qui nous a proposé le sujet et encourager à la réalisation de ce travail et pour sa patience, et surtout pour sa confiance, ses remarques et ses conseils*

Veuillez trouver ici l'expression de notre respectueuse considération.

Nos sincères remerciements vont également aux membres de jury *pour avoir accepter d'évaluer et examiner ce travail, notamment,* "Mr. LAACHENANI" qui nous fait l'honneur de présider le jury et "Mr. BOUCHERITE " en tant que membre examinateur de jury.

Nous nous devons à remercier particulièrement nos parents notre succès demeure de loin.

Nous remercions tous ceux qui de près ou de loin nous ont soutenus et aidés pour achever ce travail.

Fedoua et Yousra

Dédicace

*A la mémoire de mes grands-mères, mamie « **khadidja** » qui nous a quitté très tôt je dédie ce travail a son angélique âme et mani « **Fatima Zohra** » que dieu les accueille dans son propre paradis.*

Je dédie ce modeste travail accompagné d'un profond d'amour :

*A celle qui m'a arrosé de tendresse et d'espoirs, Quoi que je fasse ou que je dise, je ne saurai point te remercier comme il se doit. Ton affection me couvre, ta bienveillance me guide et ta présence à mes côtés a toujours été ma source de force pour affronter les différents obstacles, ma source de vie ma maman « **Fatiha** ».*

*A mon très cher père « **Benaissa**» Tu as toujours été pour moi un exemple du père respectueux, honnête, de la personne méticuleuse, je tiens à honorer l'homme que tu es. Grâce à toi papa j'ai appris le sens de la responsabilité. Je voudrais te remercier pour ton amour, ta générosité, ta compréhension. Ton soutien fut une lumière dans tout mon parcours.*

*A mon âme sœur « **Nour El Houda** » et mon bien-aimé mon frère « **Alaa** » et mes petites nièces « **Maria**» et « **Meriem»** et mon neveu mon petit prince « **Abderzzak**».*

*A mon cher oncle « **Chaker** » Tu as toujours offert soutien et réconfort, j'exprime envers toi une profonde admiration, reconnaissance et attachement inconditionnels.*

*A mes chères tantes en particulier « **Ikram** » et « **Souad** ».*

*Mes chères cousines « **Zola Ilhem Nadjet Mezoura Manel** ».*

*A toute ma famille « **Hamadi** » et « **Ameur** » sans exception.*

*A mes copines « **Nesrine Khouloud Chaima Malak Soumia Madina Meriem Manal** ».*

*Sans oublier ma très chère partenaire « **Fedoua** » pour son soutien moral, sa patience et sa compréhension tout au long de ce projet.*

Et à toutes mes collègues de promotion 2023.

Yousra

Dédicace

Je dédie ce travail :

*A mes chers parents ma mère « **Amel** » et mon père « **Mohamed** » pour leur amour, leur soutien et leur encouragement.*

*A mes frères « **Sifeddine** » et « **Abdelaziz** » et ma sœur « **Romaissa** » et mon petit neveu « **Dhiaa** » que je admire, sans oublier ma grand-mère « **Hafsa** » que j'aime et à la mémoire de ma grand-mère défunte « **aicha**. »*

« Que Dieu protège ma grande famille »

*Et à mes meilleurs amis « **CHahinez** » et « **Batoul** »*

*A ma très chère binôme « **Yousra** »*

A tous mes amis de promo Télécom 2023 pour leur sincère amitié.

A toute personne qui de près ou de loin a participé à ma carrière.

Je Vous Remercie.

Fedoua

ملخص

يعد فصل المصدر الأعمى موضوعًا ذا أهمية كبيرة في مجال معالجة الإشارات. يتضمن فصل المصدر الأعمى استعادة إشارات المصدر غير المعروفة فقط من المخاليط المرصودة. في الواقع، تعني كلمة أعمى أننا لا نحتاج إلى معلومات عن إشارات المصدر الأصلية باستثناء استقلاليتها الثابتة.

الهدف من هذه المذكرة هو اكتشاف طريقة جديدة لفصل مصدر الصوت باستخدام خوارزمية MGD IS-NMF2D. يعتمد تطوير خوارزمية الفصل هذه على تمثيل التردد الزمني باستخدام مرشحات غاما. على وجه التحديد، نظهر أن الإشارات الصوتية الأحادية المختلطة أكثر قابلية للفصل في مجال غير المنتظم.

سنقدم أيضًا بعض نتائج المحاكاة التي تسمح بتقييم أداء هذه الطريقة.

الكلمات المفتاحية: فصل المصدر الأعمى، بنك مرشحات غاماتون، خوارزمية MGD IS-NMF2D ، إشارات صوتية.

Résumé

La Séparation Aveugle de Sources (SAS) est un sujet de grand intérêt dans le domaine du traitement du signal. La séparation aveugle de source (SAS) consiste à récupérer des signaux de source inconnus uniquement à partir des mélanges observés. En effet le mot aveugle signifie on n'a pas besoin d'information sur les signaux source original sauf leurs indépendances statiques.

L'objectif de ce mémoire est de découvrir une nouvelle méthode de séparation de sources audio en utilisant l'algorithme MGD IS-NMF2D. Le développement de cet algorithme de séparation est basé sur une représentation Temps-Fréquence (TF) en utilisant le banc de filtres Gammatone. Particulièrement, nous montrons que les signaux audio mono-mixtes sont plus séparables dans le domaine TF non uniforme.

Nous présenterons également, quelques résultats de simulation qui permettant d'évaluer les performances de cette méthode.

Mots clés : La séparation Aveugle de Source, le banc de filtres Gammatone, l'algorithme MGD IS-NMF2D, les signaux audio.

Abstract

Blind Source Separation (BSS) is a subject of great interest in the field of signal processing. Blind source separation (BSS) involves recovering unknown source signals only from the observed mixtures. Indeed, the word blind means we do not need information on the original source signals except their static independences.

The objective of this thesis is to discover a new method of audio source separation using the MGD IS-NMF2D algorithm. The development of this separation algorithm is based on a Time-Frequency (TF) representation using the Gammatone filter bank. Specifically, we show that mono-mixed audio signals are more separable in the non-uniform TF domain.

We will also present some simulation results that allow to evaluate the performance of this method.

Keywords : Blind Source Separation, Gammatone filter bank, MGD IS-NMF2D algorithm, audio signals.

Sommaire

Liste des figures I

Liste des tableaux III

Liste des Acronymes et des sigles IV

Introduction générale 1

Chapitre I : L'état de l'art de la séparation de sources aveugles.

I.1. Introduction 4

I.2. Description générale de la séparation aveugle de sources (SAS) 4

I.3. Historique de la Séparation Aveugle de Sources 5

I.4. Les types de mélanges 7

I.4.1. Les mélanges linéaires 7

I.4.1.1. Le mélange convolutif 8

I.4.1.2. Le mélange instantané 10

I.4.2. Les mélanges non-linéaires 11

I.5. Applications de la séparation aveugle de sources (SAS) 11

I.5.1. Séparation de signal parole 11

I.5.2. Communication et traitement du signal 12

I.5.3. Traitement et reconnaissance d'image 12

I.5.4. Traitement de l'information géologique spatiale 12

I.5.5. Traitement du signal biomédical 13

I.6. Algorithmes de séparation aveugle de sources 13

I.6.1. L'algorithme SOBI 13

I.6.2. L'algorithme COM2 14

I.6.3. L'algorithme COM1 15

I.6.4. L'algorithme JADE 15

I.7. Conclusion 16

Chapitre II : Le banc de filtre Gammatone.

II.1. Introduction 18

II.2. Description du système auditif humain 18

II .2.1. L'Oreille Externe 19
Le pavillon 19
Le conduit auditif externe 20
II .2.2. L'Oreille Moyenne 20
II .2.3. L'Oreille Interne 21
Le vestibule 21
La cochlée 22
II.3. Filtres auditifs 22
II.4. Le masquage auditif 26
II.4.1. Les types de masquage 27
II.4.1.1. Le masquage simultané (ou fréquentiel) 27
II.4.1.2. Le masquage non-simultanée (ou temporel) 29
II.5. Le banc de filtres Gammatone 30
II.5.1. Introduction sur les filtres Gammatone 30
II.5.2. Quatre filtres de la famille des Gammatone 32
II.5.2.1. Le filtre Gammatone (GTF) 32
II.5.2.2. Le filtre Gamma Chirp (GCF) 32
II.5.2.3. Le filtre Gammatone à tous les pôles (APGF) 32
II.5.2.4. Le filtre gammatone à un zéro (OZGF) 33
II.5.3. Le filtre Gammatone dans le domaine temporel 33
II.5.4. Le filtre Gammatone dans le domaine fréquentiel 33
II .5.5. Largeur de bande rectangulaire équivalente du filtre Gammatone 34
II .5.6. Compensation de phase 36
II.6. Conclusion 36
Chapitre III : la simulation et les résultats de la SAS.
III.1. L'introduction 38
III.2. C'est quoi le logiciel MATLAB ? 38
III.2.1. Le son sous MATLAB 39
III.2.2. Lecture d'un fichier son 39

III.2.3. L'écriture d'un fichier son 39
III.3. Les données de traitement 39
III.4. Banc de filtre Gammatone 40
III.5. Choix de l'échelle fréquentielle 41
III.6. Réalisation d'un filtre Gammatone 41
Résultats 42
La réponse impulsionnelle d'un filtre Gammatone 42
Réponse en fréquence d'un filtre Gammatone 44
III.7. Transformée de Fourier d'un filtre Gammatone 45
Résultats 46
III.8. Le cochléagramme 48
Résultats 49
L'énergie d'un filtre Gammatone 49
Le cochleagramme d'un filtre Gammatonne 51
III.9. Algorithme de séparation de sources 52
III.10. Estimation des sources 53
Résultats de simulation 54
III.11. Conclusion 55
Conclusion générale 56
Bibliographie 57

Liste des figures :

Chapitre I : L'état de l'art de la séparation de sources aveugles.

Figure I.1. Principe de la séparation aveugle de sources 5
Figure 1.2. Modèle du mélange linéaire.....7
Figure I.3. Modèle des mélanges convolutifs. 9
Figure I.4. Mélange linéaire instantané.....11

Chapitre II : Le banc de filtre Gammatone.

Figure II.1. Le système auditif 19
Figure II.2. Anatomie du pavillon 20
Figure II.3. Anatomie de conduit auditif externe 20
Figure II.4. Anatomie de l'oreille moyenne 21
Figure II.*5.* Section axiale de la cochlée 22
Figure II.6. Modèle de seuil de détection pour un neurone du nerf auditif. Le seuil à la fréquence cible est de 20 dB SPL. D'après Hartmann 23
Figure II.7. Pattern d'excitation de deux sons purs de niveau modéré. F1 = 500Hz (50 dB SPL) et F2 = 1200Hz (20 dB SPL) 25
Figure II.8. Pattern d'excitation de deux sons purs de fort niveau. F1 = 500Hz (90 dB SPL) et F2 = 1200Hz (60 dB SPL) 26
Figure II.9. Représentation de l'effet de masque 27
Figure II.10. Courbes d'effet de masque de bruits à bande étroite. Le signal masquant est un bruit de bande centré à 1 kHz de largeur 160 Hz présenté aux niveaux LG = 100, 80,60, 40 et 20 dB 28
Figure II.11. Courbes de sélectivité de fréquence Psycho acoustique (psychophysical tuning curves) 29
Figure II.12. Masquage non-simultané 30
Figure II.13. Réponse en fréquence des fonction G(z) (en bleu) et K(z) (en rouge) pour $\tilde{a}$=0.9693 + 0.2104i, correspondant à une fréquence centrale de 1,5kHz et une largeur de bande de 50Hz 31

Figure III.1. Le tracé du mélange (piano et trompette) à séparés 41
Figure III.2. Réponse impulsionnelle de la bande passante N°3. 43
Figure III.3. Réponse impulsionnelle de la bande passante N°50 43
Figure III.4. Réponse impulsionnelle de la bande passante N°100 44
Figure III.5. Réponse en fréquence d'un banc de filtre Gammatone 12 bandes..... 45
Figure III.6. Le Signal de filtre Gammatone filtré de la bande passante N°3 46
Figure III.7. Le Signal de filtre Gammatone filtré de la bande passante N°50 47
Figure III.8. Le Signal de filtre Gammatone filtré de la bande passante N°100 47
Figure III.9. La sortie de l'énergie d'un filtre Gammatone de la bande passante N°3 49
Figure III.10. La sortie de l'énergie d'un filtre Gammatone de la bande passante N°50 50
Figure III.11. La sortie de l'énergie d'un filtre Gammatone de la bande passante N°100 50
Figure III.12. Le cochléagramme de piano mélangé au trompette 51
Figure III.13. Le cochléagramme de source estimé 1 51
Figure III.14. Le cochléagramme de source estimé 2 52
Figure III.15. Les deux sources estimées (séparées). 54

Liste des tableaux

Tableau II .1. Les données pour le filtre Gammatone. 35

La liste des acronymes et des sigles

APGF	The all-pole gammatone filter.
BSS	Blind Separation Sources.
CDMA	Code division multiple access.
COM2	Contrast Maximization2.
ECG	les signaux d'électrocardiogramme.
ERB	Equivalent Rectangular Bandwidth.
FFT	FastTransformede Fourier
GCF	The gammachirp filter .
GTF	The gammatone filter .
ICA	Independent component analysis.
IIR	Infinité Impulse Réponse.
JADE	Joint Approximate Diagonalization of Eigen-matrices.
OZGF	The one-zero gammatone filter.
PTC	Psychophysical Tuning Curves.
SAS	Séparation aveugle de sources.
SOBI	Second Order Blind Identification.
TF	Temps-Fréquence

Les sigles

$\boldsymbol{\Sigma}$: Sommation

$\propto$: proportionnalité

$|\bullet|$ *:* module ou Valeur absolue.

$\boldsymbol{A}$: matrice de mélange instantané composée de coefficients scalaires réels $\boldsymbol{aij}$.

$\boldsymbol{Cumr(.)}$: cumulant d'ordre r.

$\boldsymbol{H}$ *:* entropie.

$\boldsymbol{\lambda}$ *:* valeurs propres d’une matrice.

$\boldsymbol{T}$: nombre d'échantillons de chaque observation.

$\boldsymbol{t}$: variable du temps.

Le traitement de la parole fait désormais partie intégrante des sciences de l'ingénieur. Il se trouve entre le traitement numérique du signal et le traitement du langage (c'est-à-dire du traitement des données symboliques), cette discipline scientifique a connu une expansion fulgurante depuis les années **60**, liée au développement des moyens et des techniques de télécommunications. L'importance particulière du ce premier dans ce cadre plus général s'explique par la position privilégiée de la parole comme vecteur d'information dans notre société humaine. Il fait partie de la discipline du traitement de signal.

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.

Le cerveau de l'être humain est l'unité centrale de traitement de l'information et à partir de ces neurones, plusieurs signaux proviennent respectivement que ce soit par voie visuelle ou auditive. La modélisation de la transmission des signaux d'information de mouvement à partir des articulations, par exemple au niveau d'un bras ou d'un pied, comme la vitesse ou la position, a donné lieu à un problème connu sous le nom de problème de séparation aveugle des sources.

La séparation aveugle des sources a reçu beaucoup d'attention au cours des dernières décennies. Dans de nombreux cas, il est souhaitable de récupérer toutes les sources à partir des mélanges enregistrés, ou du moins de séparer une source spécifique.

La séparation aveugle des sources (SAS) 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.

Ce manuscrit est organisé en trois chapitres comme suit :

- Le premier chapitre, nous présenterons l'état de l'art de la séparation de sources en expliquant son principe, et on donne une description générale de la SAS, les types de mélanges et leurs modèles mathématiques, ainsi que les applications de la SAS.
- Le deuxième chapitre, nous avons défini le système auditif humain, ensuite on introduit la notion de filtres auditifs qui indique que l'oreille traite les sons d'une manière comparable à un banc de filtres. Et à la fin du chapitre nous avons détaillé le filtre

Gammatone le principe de fonctionnement, la réponse impulsionnelle et la réponse fréquentielle.

- Le troisième chapitre, on montre l'algorithme de séparation de sources avec les résultats de simulations obtenus.

Une conclusion générale donnera une synthèse en ce qui concerne les différentes démarches suivies dans ce travail et expose également les perspectives envisagées pour les travaux futurs.

Chapitre 1 :

L’état de l’art de la séparation des sources

I.1. Introduction

La séparation aveugle des sources occupe un volume très important dans le traitement du signal, elle utilise de nombreuses parties de la théorie de signal. C'est un problème qui consiste à trouver des signaux statistiquement indépendants de leur mélange (observations), Les approches de solution de ce problème se décomposent essentiellement en deux grandes parties, statistiques et neuronales. Ces méthodes sont proposées pour trouver les signaux sources reçus par des capteurs qui reçoivent un mélange de signaux sources. On dit que la séparation est aveugle en raison d'absence d'information propre sur les signaux sources et la nature de matrice de mélange des signaux reçus.

I.2. Description générale de la séparation aveugle de sources (SAS)

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.

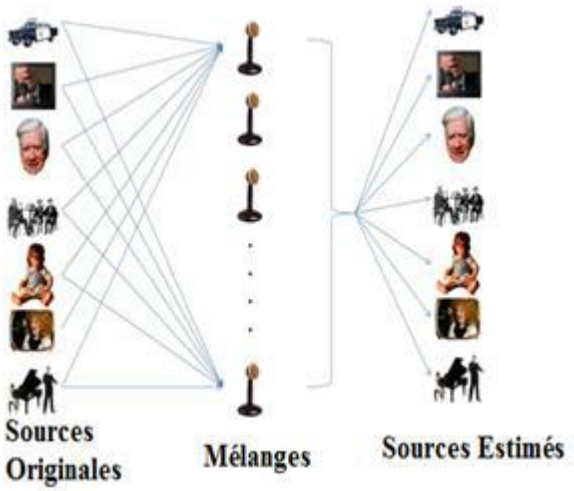


Figure I.1. *Principe de la séparation aveugle de sources[1]*

I.3. Historique de la Séparation Aveugle de Sources

La séparation aveugle de sources a été formalisée pour la première fois par des chercheurs français en traitement de signal au milieu des années 80. En 1985 par C.Jutten et J. Hérault. Ces derniers ont développé, à partir d'un problème de décodage du mouvement en neurobiologie, un algorithme de séparation à base d'un réseau neuronale. Cet algorithme cherche à résoudre le problème du mélange instantané sans bruit avec une matrice de mélange carrée et des signaux à valeurs réelles.

Depuis, une multitude d'algorithmes a été proposée. Ils peuvent être classés selon différents critères : mélange instantané ou mélange convolutif, linéaire ou non linéaire, statistiques d'ordre deux ou d'ordre supérieur.

En 1988 J. L. Lacoume et P. Ruiz ont proposé un algorithme minimisant une fonction objective basée sur les cumulant croisés des sorties d'un filtre inverse, qui sont les outils naturels de mesure d'indépendance.

En 1989, J. F. Cardoso et Solomiac proposent de considérer les propriétés algébriques des cumulant d'ordre quatre en utilisant l'information d'ordre deux pour rendre unitaire la matrice de mélange puis de l'identifier par des techniques de diagonalisation. En revanche, ils montrent que les cumulant d'ordre quatre sont suffisants pour permettre, à eux seuls, d'estimer la matrice de mélange.

L'approche développée par P. Comon, en 1990, élargit la théorie de l'analyse en composantes principales (en anglais : PCA, Principale Components Analysais) à l'analyse en

composantes indépendantes et introduit la notion de fonction de contraste. L'auteur propose comme contraste à maximiser la somme des cumulant d'ordre quatre. Comon montre par ailleurs, qu'il suffit de reconstruire des sources indépendantes deux à deux pour estimer, à une matrice triviale près, des sources mutuellement indépendantes. Des travaux de P. Comon, M. Gaeta et J. L. Lacoume proposent de résoudre le problème de séparation par une technique du maximum de vraisemblance approchée.

B. Laheld et J. F. Cardoso ont développé en 1993 une famille d'algorithmes adaptatifs basés sur la notion de gradient matriciel relatif toujours en 1993, survient la méthode JADE (Joint Approximante Diagonalisation of Eigen-matrices), au travers de la quelle Cardoso et Solomiac présentent une solution algébrique à la maximisation de leur contraste basé sur les cumulant d'ordre 4. Ils étendent par la même occasion l'algorithme de Jacobi dans le but de diagonaliser conjointement un ensemble de matrices.

En 1995, Delafosse et Loubaton introduisent le concept de déflation en proposant une manière originale d'optimiser le contraste uni-source d'ordre 4 sur lequel s'appuie leur méthode. Deux années plus tard, Hyvärinen et Oja présentent l'algorithme Fast-ICA. La version initiale de cet algorithme exploite pleinement le principe de déflation et la mesure de non-gaussianité, plus connue sous le nom denéguentropie, approximée par la valeur absolue du kurtosis de la source estimée.

Un certain nombre d'auteurs parmi eux K. Kokkinakis en 2006, 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 (2006) propose une version d'algorithme multidimensionnel pour l'estimation de la densité (2007) et en exploitant les dépendances en fréquence d'ordre supérieur [1].

I.4. Les types de mélanges

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

1.4.1. Les mélanges 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.

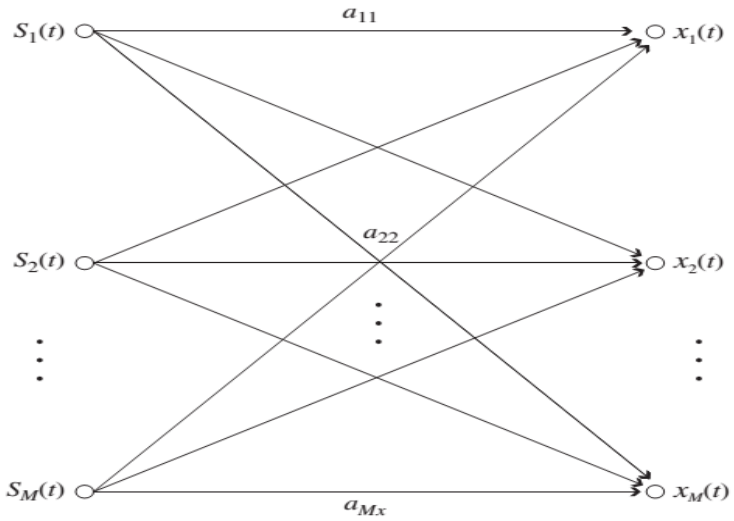


Figure I.2. *Modèle du mélange linéaire [1]*

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

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

Avec

$\boldsymbol{a_{ji}}(\boldsymbol{i} \in \{\boldsymbol{1}, \boldsymbol{2}, \dots, \boldsymbol{N}\}, \boldsymbol{j} \in \{\boldsymbol{1}, \boldsymbol{2}, \dots, \boldsymbol{M}\})$ sont les paramètres du mélange,

$\boldsymbol{s_i(t)}(\boldsymbol{i} \in \{\boldsymbol{1}, \boldsymbol{2}, \dots, \boldsymbol{N}\})$ sont les signaux de source

$\boldsymbol{x_i(t)(i \in \{1, 2, \dots, M\})}$ sont les signaux observés.

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

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

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

$$\boldsymbol{x(t) = A.s(t)} \qquad \textbf{(I.2)}$$

Où

$\boldsymbol{A \in R^{MxN}}$ est la matrice de mélange et

$\boldsymbol{s \in R^{MxN}}$ est le vecteur source dont les éléments sont les coefficients de la matrice $\boldsymbol{A}$.

Comme le canal de transmission et le réseau de capteurs peuvent inclure du bruit additif qui est à prendre en compte, l'équation (I.2) devient :

$$\boldsymbol{x(t) = A.s(t) + n(t)} \qquad \textbf{(I.3)}$$

Où

$\boldsymbol{n(t) = [n_1(t), \dots \dots \dots, n_M(t)]^T}$ 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. [2]

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. [3]

I.4.1.1. Le mélange Convolutif

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 [4] . Le modèle de mélange convolutif est plus compliqué que le modèle de mélange instantané linéaire. Cependant, il a été prouvé dans [5] 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 [6] , 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 $\boldsymbol{s_i(t), i = 1, 2, \dots N}$ reçues par $\boldsymbol{M}$ capteurs après le processus de mélange par convolution (figure I.3).[1]

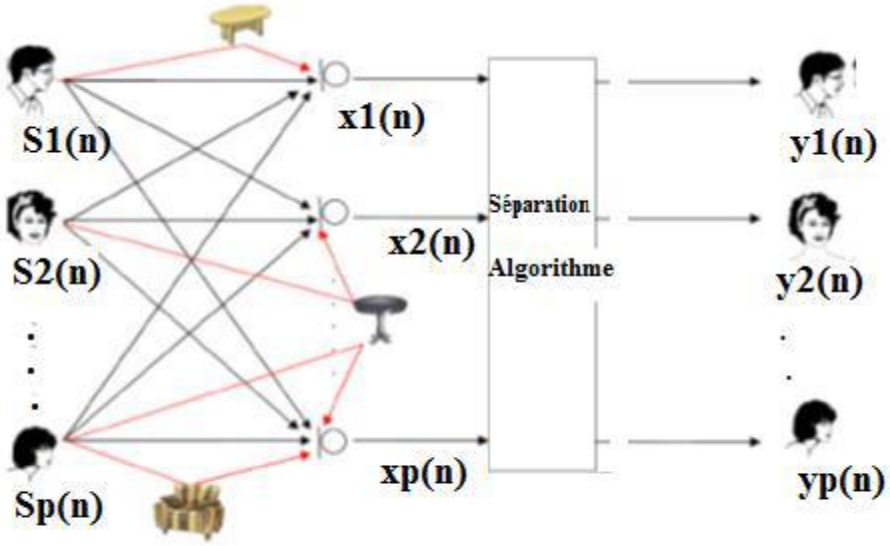


Figure I.3. *Modèle des mélanges convolutifs.[1]*

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

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

Où * désigne l'opération de convolution et $\boldsymbol{a_{ji}(\tau)}$ 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).

$$\boldsymbol{x = As} \qquad (\boldsymbol{I.5})$$

Où A est une matrice FIR de la forme :

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a_{11}^T} & \cdots & \boldsymbol{a_{1n}^T} \\ \vdots & \ddots & \vdots \\ \boldsymbol{a_{1n}^T} & \cdots & \boldsymbol{a_{nn}^T} \end{bmatrix} \qquad (\boldsymbol{I.6})$$

Où $\boldsymbol{a_{ji}}$ est une colonne de dimension L qui désigne l'ordre du filtre FIR.

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

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

où

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

pour L=1, le mélange convolutif devient un mélange instantané.

I.4.1.2. Le mélange instantané

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 $sj(t)$, les observations $xi(t)$ sont de la forme :

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

Où aij 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.

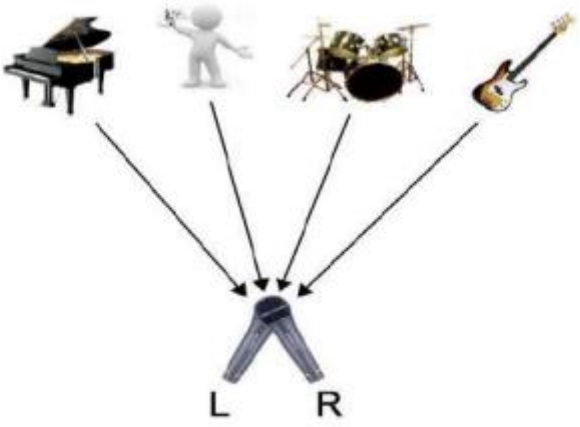


Figure I. 4. *Mélange linéaire instantané.* [1]

I.4.2. Les mélanges non-linéaire :

Dans certains cas, les signaux sources peuvent être mélangés de manière non linéaire, ce qui signifie que la combinaison des sources ne peut pas être représentée par une simple opération linéaire.

I.5. Applications de la séparation aveugle de sources (SAS)

La séparation à la source est un domaine de recherche très actif en raison de son intérêt théorique et de ses diverses applications pratiques. Dans cette liste non-exhaustive d'applications, nous nous référons à :

I.5.1. Séparation de signal parole :

La situation en temps réel de la position du microphone par rapport à la source sonore pouvant être très arbitraire, le processus de mélange est inconnu et doit donc être estimé à l'aveugle. Dans ce cas, l'algorithme SAS est important.

De plus, les algorithmes SAS convolutifs sont souvent préférentiellement appliqués dans le domaine musical. Douglas a utilisé la technique SAS convolutive pour séparer deux voix [7], tandis que Vincent a appliqué la technique ICA dans [8] pour mettre les voix de différents instruments dans une partition.

I.5.2. Communication et traitement du signal :

Le SAS a été largement utilisé dans les communications numériques sans fil [9]. Divers chercheurs ont proposé d'utiliser un système de réception de réseau d'antennes et une technologie de filtrage spatial combinés à l'algorithme SAS pour réaliser une séparation de signal multi-utilisateur, de sorte que plusieurs signaux d'utilisateur peuvent être transmis sur le même canal en même temps, améliorant ainsi considérablement les performances. Capacité du canal. Dans un système à accès multiple par répartition en code (CDMA), plusieurs utilisateurs partagent la même bande passante de canal et transmettent des signaux simultanément. Les utilisateurs peuvent distinguer leurs propres signaux grâce à des codes d'étalement orthogonaux plus longs, par conséquent l'utilisation de l'ICA peut aider à améliorer le SNR (rapport signal sur bruit) à l'extrémité de réception.

En 2006, l'institut de recherche japonais RIKEN et la société Chaos ware ont développé une nouvelle technologie capable de séparer avec précision les ondes électromagnétiques mixtes du téléphone, améliorant ainsi considérablement la vitesse de transmission et la clarté de la communication.

I.5.3. Traitement et reconnaissance d'image :

À l'heure actuelle, la technologie SAS est largement utilisée dans le domaine du traitement d'image, et la technologie de séparation aveugle est utilisée dans l'extraction de caractéristiques d'image, la reconnaissance faciale, la détection d'objets en mouvement, le filigrane d'écran, l'image numérique, le débruitage d'image, la séparation d'image et la restauration d'image [10] [11].

Par exemple, dans les applications de restauration et de reconstruction d'images, les méthodes BSS disponibles peuvent aider à restaurer le contenu original de l'image à partir d'images contaminées et à supprimer les effets négatifs tels que le bougé de l'appareil photo.

I.5.4. Traitement de l'information géologique spatiale :

Dans le traitement de l'information géospatiale, la technologie SAS a été appliquée à de nombreux aspects du traitement des données de télédétection, y compris la fusion d'images de télédétection, la décomposition mixte de pixels, la classification et la reconnaissance d'images, etc., la télédétection [12] [13].

En prévision minérale, la méthode SAS a été appliquée avec succès à l'analyse de données géophysiques et géochimiques. La technologie SAS peut également être utilisée dans

les levés sismiques. L'utilisation de techniques SAS dans le dynamitage explosif des sols peuvent aider à distinguer différentes caractéristiques lithologiques à partir des signaux de réflexion ou de diffraction reçus [14].

I.5.5. Traitement du signal biomédical :

Il s'agit d'un domaine d'application très encourageant pour la technologie SAS, non seulement parce qu'il s'agit d'un domaine en croissance rapide et important, mais aussi parce que certains types de données d'imagerie cérébrale semblent être bien décrits par les modèles BSS. En particulier, les signaux d'électrocardiogramme (ECG), d'EEG, d'électromyographie (EMG), de MEG et d'IRMf, qui sont des enregistrements des champs électriques et magnétiques des signaux électriques neuronaux dans le cerveau [15].

Par exemple, dans le traitement des signaux ECG fœtaux, les signaux ECG mesurés par les femmes enceintes incluent en fait les signaux des femmes enceintes et des bébés à naître, alors que les paramètres du support de transmission sont inconnus, ce qui est une application typique. En utilisant la méthode SAS, nous pouvons distinguer et séparer les signaux ECG des femmes enceintes et des bébés à naître, fournissant ainsi des données précises pour un diagnostic ultérieur. De plus, l'ICA a été appliqué avec succès au traitement du signal IRMf du cerveau pour obtenir des informations sur les caractéristiques cachées, fournissant des données utiles pour l'étude scientifique du comportement humain [11].

I.6. Algorithmes de séparation aveugle de sources :

I.6.1. L'algorithme SOBI :

L'algorithme SOBI utilise plusieurs matrices de covariance associées à différents délais non nuls après les observations de blanchiment. De plus, les sources doivent être temporellement corrélées, avoir des spectres distincts et le bruit doit être temporellement blanc. La matrice de covariance des observations blanchies$\{x[m]\}$, $m \in \mathbb{N}$ est alors donnée pour un retard $\tau_j \in \mathbb{N}$ par :

$$R_z[\tau_j] = E[z[m]z[m-\tau_j]^T] - E[z[m]]E\left[[m-\tau_j]^T\right] = \hat{H}R_s[\tau_j]\hat{H}^T \qquad (I.9)$$

Où

$\hat{H}$ *:* représente la matrice de mélange blanchie orthogonale

$R_s[\tau_j]$: désigne la matrice de covariance des sources associée au retard τj, qui a la propriété d'être diagonale.

I.6.2. L'algorithme COM2 :

La première classe de méthodes est basée sur la technique de l'Analyse en Composantes Indépendantes, utilisant des outils d'ordre supérieur pour construire des fonctions de contraste, les quelles exploitent les propriétés des cumulant des sources supposées indépendantes. C'est P.Comon qui a étendu la notion de contraste utilisée auparavant dans les systèmes de déconvolution aveugle au cas des vecteurs aléatoires.

Il a proposé dans une méthode baptisée COM2 (COntrast Maximization2) consistant à maximiser une fonction de contraste notée ψ_R^2 (y) (Il s'agit de l'une des premières fonctions de contraste), et définie comme la somme des modules au carré des cumulant d'ordre $R \geq 3$ des sources estimées $y(t)$:

$$\psi_R^2(y) = \sum_{i=1}^{n} |C_R\{y_i\}|^2 \qquad (I.10)$$

Où n est le nombre de sources et $CR\{.\}$ représente le cumulant d'ordre R d'un signal stationnaire. La maximisation de la fonction contraste ψ_R^2 (y) a conduit à l'algorithme ICA.
De même que SOBI, COM2 nécessite un blanchiment des observations. Puis il faut maximiser le contraste. Une optimisation de ce contraste repose sur le théorème suivant :
Soit ***s*** un vecteur aléatoire de composantes indépendantes, parmi lesquelles au plus une est gaussienne, et dont les densités de probabilité ne sont pas réduites à une masse ponctuelle. Soit ***T*** une matrice orthogonale de taille ***(P× P)*** etc. le vecteur défini par ***z=T.s.***

Alors, les trois propriétés suivantes sont équivalentes :
- les composantes de **z** sont indépendantes deux à deux.
- les composantes de **z** sont mutuellement indépendantes.
- **T** est une matrice triviale.

Ce théorème nous dit qu'au vu des indéterminations intrinsèques au problème de la Séparation Aveugle de Sources, il suffit juste de chercher à reconstruire des sources indépendantes deux à deux.

Une autre façon d'introduire les cumulants marginaux en tant que critère d'optimisation consiste à remarquer que les variables aléatoires sont indépendantes si tous Leurs cumulants

croisés sont nuls. On peut donc être tenté de minimiser l'ensemble des cumulants croisés à un ordre donné.

I.6.3. L'algorithme COM1 :

Après un blanchiment des observations le contraste suivant est optimisé :

$$\psi_R^2(y) = \lambda \sum_{i=1}^{n} |C_R\{y\}| \ , o\grave{u}\ \lambda = \pm 1 \qquad (I.11)$$

La solution itérative proposée pour **COM2** peut être utilisée afin d'optimiser le contraste précédent. Cependant, une solution algébrique a été proposée. Notons toutefois que **COM1** nécessite que les sources aient des kurtosis de même signe.

L'idée de l'algorithme **COM1(α1)** ou **COM2(α2)** dans $\boldsymbol{\psi_R^2(y) = \sum_{i=1}^{n} |C_R\{y_i\}|^{\alpha}}$est de remarquer que ce maximum peut être calculé entièrement analytiquement en dimension2, dans un certain nombre de cas. L'intérêt de l'algorithme **COM1** réside dans une propriété établissant que la valeur absolue peut être supprimée lorsque les kurtosis des sources sont tous de même signe. C'est-ce qui sera toujours implicitement supposé dans toutes les mises en œuvre de **COM1**.

I.6.4. L'algorithme JADE :

Il est basé sur l'utilisation de cumulants d'ordre quatre impliquant à la fois des auto-cumulants et des cumulants croisés :

$$\boldsymbol{J(y) = \sum_{i1,i2,i3=1}^{n} |cum\{y_{i1} y_{i2} y_{i3} y_{i4}\}|^2} \qquad (\boldsymbol{I.12})$$

Où l'opérateur *cum{.}* désigne un cumulant d'ordre quatre. La fonction de contraste précédente est appelée « **contraste JADE** ». Maximiser cette fonction revient à diagonaliser conjointement un ensemble de *n2* matrices.

L'algorithme **JADE** nécessite aussi un blanchiment des observations, puis la maximisation du contraste.

Une solution algébrique consiste à diagonaliser conjointement l'ensemble des matrices Propres *Mp* construites à partir des vecteurs propres associés aux *P* plus grandes valeurs propres de la matrice de quadri-covariance [1].

I.7. 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'utilisation des bancs de filtres Gammatone ainsi que l'algorithme MGD IS-NMF2D dans le but de faire la séparation aveugle de sources audio.

Chapitre 2 :

Le banc de filtre Gammatone

II.1. Introduction

Dans ce chapitre nous allons présenter le principe de base des filtres Gammatone. Un filtre Gammatone est un filtre numérique conçu pour modéliser la réponse du système auditif humain aux sons. Il tire son nom de la fonction de réponse en fréquence d'un filtre Gammatone qui ressemble à une onde Gammatone.

Le filtre Gammatone est largement utilisé dans le traitement du signal audio, en particulier dans les domaines de la reconnaissance vocale, de la synthèse sonore et de l'analyse de la parole.

Il peut être utilisé pour décomposer un signal audio en plusieurs bandes de fréquences, chacune correspondant à une bande passante spécifique du filtre.

II.2. Description du système auditif humain

Le système auditif humain possède des capacités remarquables. Sa sensibilité s'étend sur une plage de fréquences audibles allant de **(20Hz à 20 kHz)** et couvre un ensemble de pressions acoustiques de **(20 µPa à 20 Pa)** ce qui représente une incroyable dynamique **1** de **120 dB**.

En outre, les capacités de discrimination aussi bien en fréquence qu'en intensité sont excellentes. Cette sensibilité est étonnante au vu du faible nombre de cellules sensorielles dont dispose la cochlée (**3500** cellules ciliées internes) en comparaison avec les millions de cellules photosensibles de l'œil.

Le système auditif humain est un système complexe qui permet à l'oreille de détecter, de traiter et de comprendre les sons. [17]

Il se compose de trois parties principales : l'oreille externe, l'oreille moyenne et l'oreille interne.

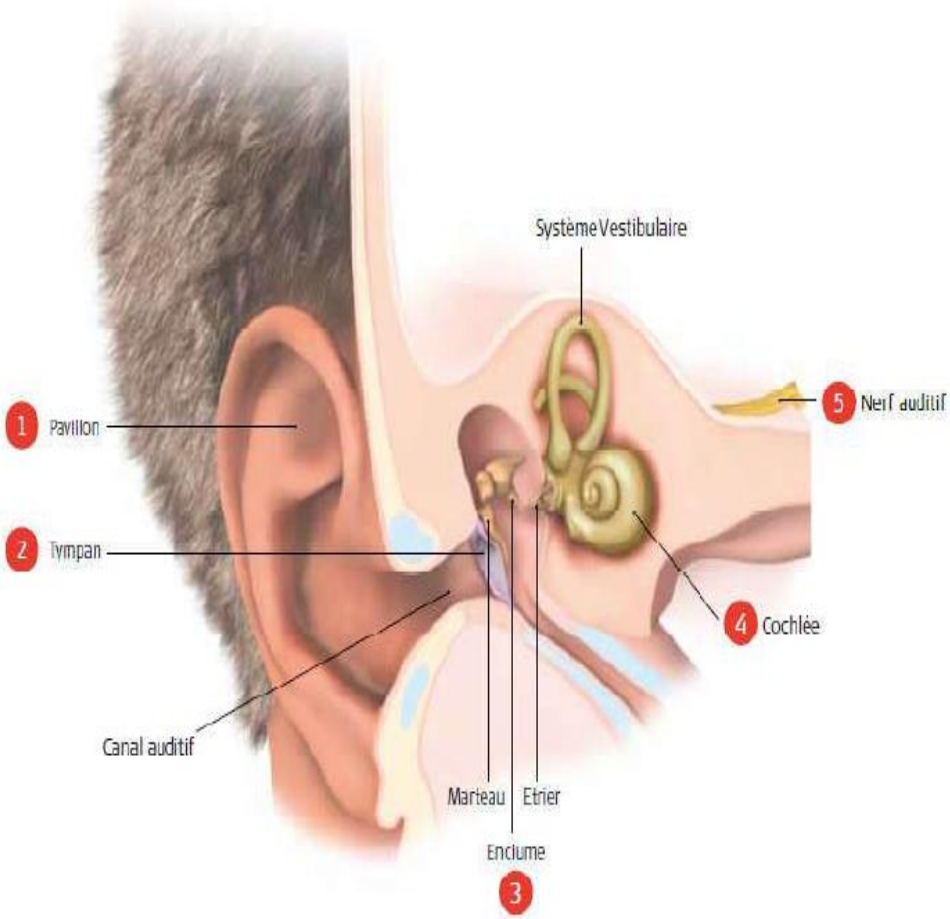


Figure II.1 : *Vue d'ensemble des 3 parties de l'oreille. [17]*

II .2.1. L'Oreille Externe

L'oreille externe est la partie visible de l'oreille et se compose du pavillon et du conduit auditif externe sachant que le pavillon recueille les ondes sonores et les canalises vers le conduit auditif externe, qui mène au tympan.

Et son rôle est d'amplifier et de transmettre les sons à l'oreille moyenne.

Le pavillon

Le pavillon est principalement composé de cartilage, qui peut filtrer sélectivement les sons en fonction de leur azimut et de leur fréquence. Au centre de ce pavillon on trouve la conque qui mène au conduit auditif externe. [18]

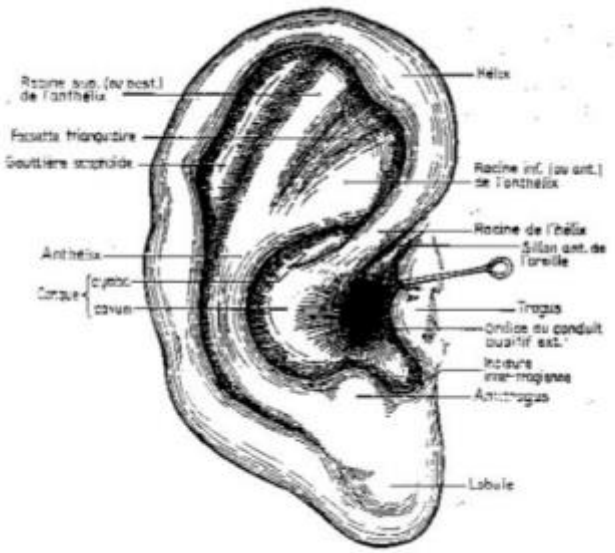


Figure II.2 : *Anatomie du pavillon. [18]*

Le conduit auditif externe

Le conduit auditif externe va de la conque à la membrane tympanique. Son rôle est de guider les ondes acoustiques jusqu'au tympan et de la protéger. Il est composé d'un tiers externe fibrocartilagineux et de deux tiers internes osseux. Il est en forme de S et sa longueur est d'environ 25 à 30 mm chez l'adulte. [19]

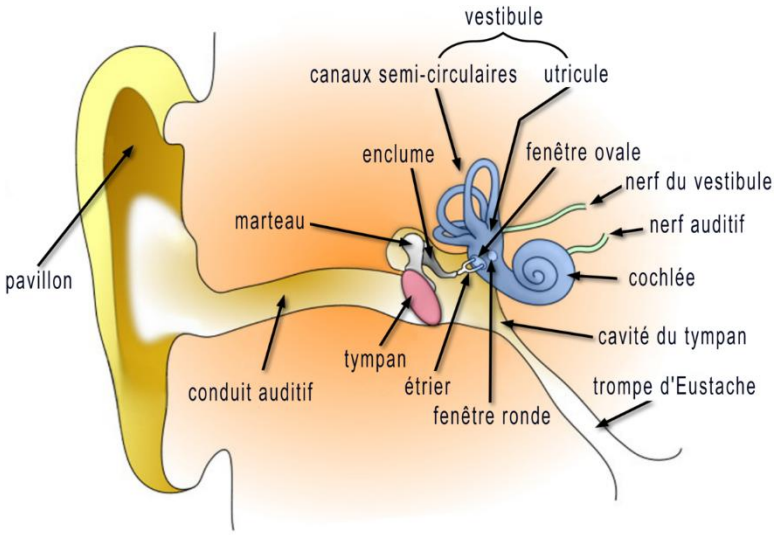


Figure II.3 : *Anatomie de conduit auditif externe. [19]*

II .2.2. L'Oreille Moyenne

L'oreille moyenne est une cavité remplie d'air située derrière le tympan. Elle est composée de trois petits os appelés osselets de l'oreille :

- Le marteau, dont le manche est inséré dans la membrane tympanique.
- L'enclume.
- L'étrier, inséré dans la fenêtre ovale.

Les osselets de l'oreille transmettent les vibrations sonores du tympan à l'oreille interne. [20]

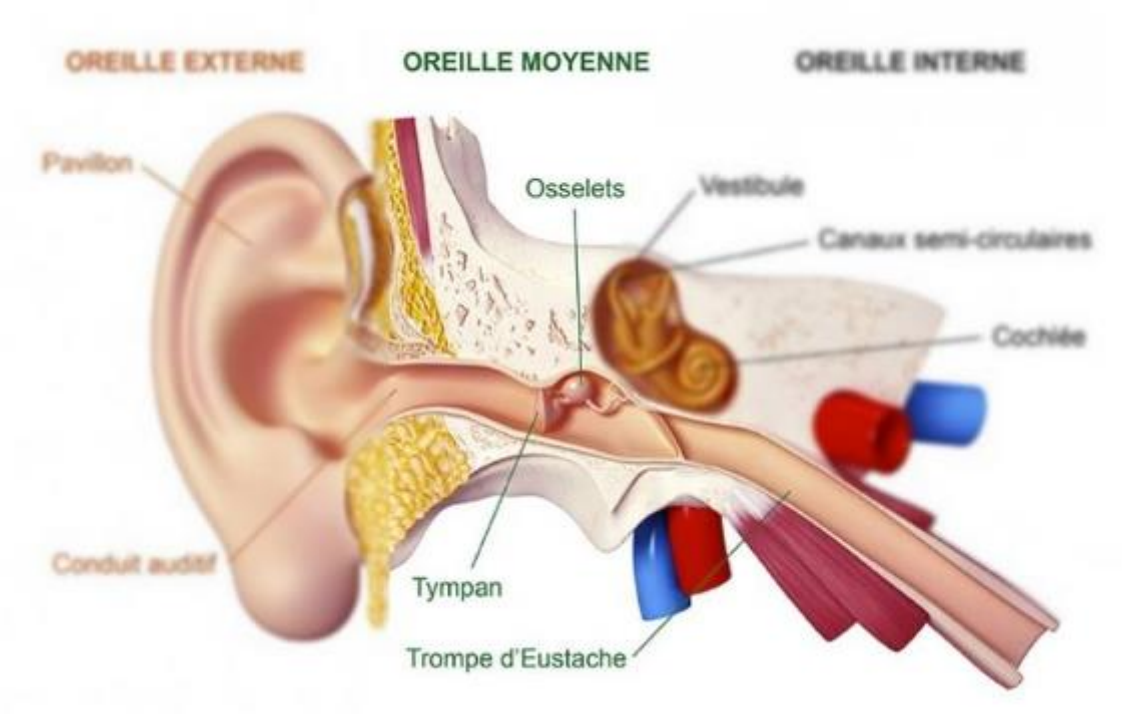


Figure II.4 : *Anatomie de l'oreille moyenne [20]*

II .2.3. L'Oreille Interne

L'oreille interne est la partie de l'oreille responsable de la transmission des signaux sonores au cerveau elle est composée de deux parties principales : la cochlée et le vestibule. [20]

Le vestibule

Le vestibule est responsable de la détection de l'équilibre et de la position du corps.

Il est situé dans le labyrinthe osseux à l'intérieur duquel, nous trouvons le labyrinthe membraneux. Le labyrinthe membraneux est essentiellement composé de la cochlée, de trois canaux semi-circulaires ainsi que l'utricule et le saccule qui servent à l'équilibre du sujet, conjointement avec la vue et la proprioception. [21]

La cochlée

La cochlée est un organe en forme de coquille qui se trouve dans l'oreille interne de l'oreille humaine. Elle est responsable de la conversion des vibrations sonores en signaux électriques qui sont transmis au cerveau pour être interprétés comme des sons.

La cochlée est remplie d'un liquide appelé liquide endolymphatique. Elle est également composée de cellules ciliées, qui sont des cellules sensorielles qui détectent les vibrations sonores. Les vibrations sonores sont transmises par l'oreille moyenne à travers la fenêtre ovale de l'oreille interne et sont ensuite transmises à la cochlée.

À l'intérieur de la cochlée, les vibrations sonores sont transmises aux cellules ciliées, qui sont organisées en rangées et sont sensibles à des fréquences sonores spécifiques. Les cellules ciliées convertissent les vibrations sonores en signaux électriques qui sont transmis au cerveau via le nerf auditif.

La cochlée est donc un organe vital pour l'audition humaine, car elle permet la conversion des vibrations sonores en signaux électriques qui sont interprétés par le cerveau comme des sons.[22]

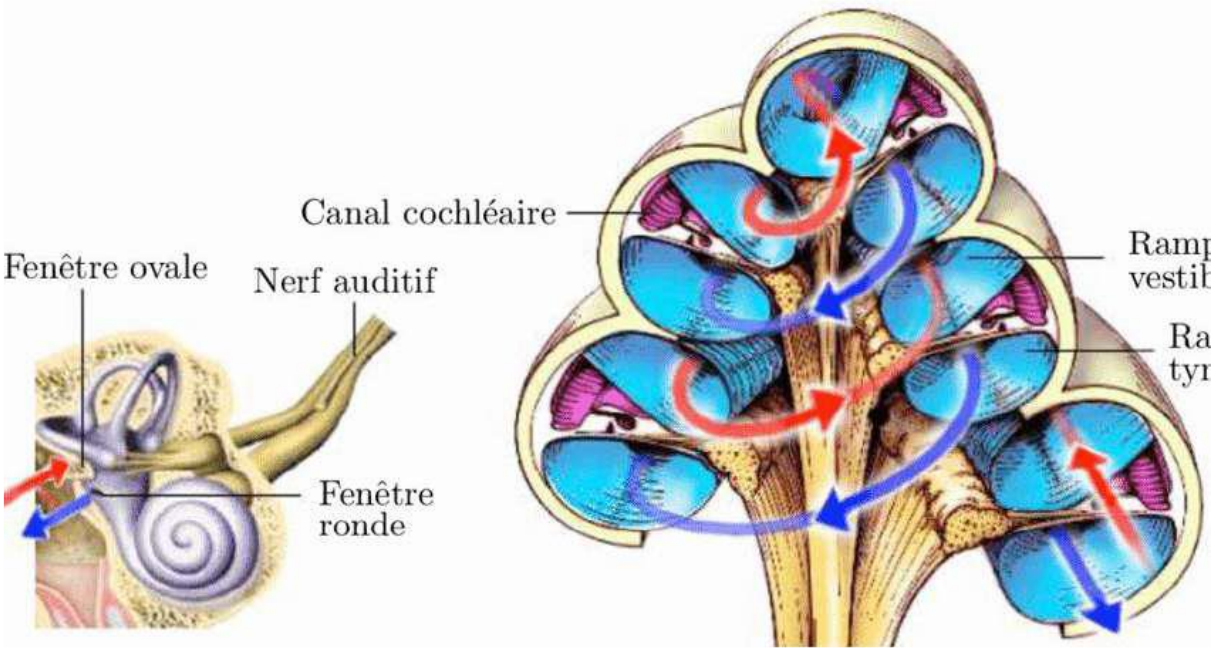


Figure II.5 : *Section axiale de la cochlée [22]*

II.3. Filtres auditifs

La perception du son commence par la conversion d'un champ de pression acoustique en impulsions neuronales grâce aux cellules ciliées disposées le long de la membrane basilaire.

Chacune de ces cellules correspond à une fréquence, et les caractéristiques du système sont très sélectives. A l'aide d'une fréquence pure, une expérience de neurophysiologie chez l'animal (où l'on enregistre l'activité d'une fibre du nerf auditif) permet de construire le seuil de détection pour un seul neurone : c'est la plus faible amplitude nécessaire pour exciter le nerf auditif en fonction de la fréquence. Une approximation est montrée figure (**II.6**). On remarque un passe-bande pointu centré sur la fréquence d'excitation. L'idée de la modélisation est donc de considérer la membrane basilaire et les cellules ciliées comme un banc de filtres passe-bande, appelés filtres auditifs. Dans le cadre de filtres linéaires, il est possible de retrouver le transfert en gain de ces filtres en inversant la courbe de seuil de détection.

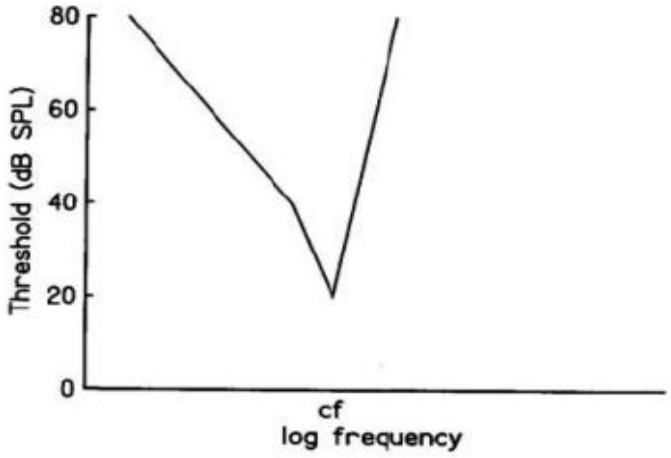


Figure II.6 : *Modèle de seuil de détection pour un neurone du nerf auditif. Le seuil à la fréquence cible est de 20 dB SPL. D'après Hartmann. [23]*

Une telle approche est cohérente avec les phénomènes de masquage simultané. En effet, pour une fréquence d'excitation donnée, les fréquences proches sont sollicitées. Il est donc prévisible que le seuil d'audition sera ainsi augmenté pour un autre son pur aux abords de cette première fréquence, si celui-ci n'a pas assez d'amplitude, il est alors masqué.

Historiquement, les filtres auditifs ont été caractérisés en premier lieu par leur largeur de bande, sous l'hypothèse qu'ils pouvaient être modélisés par des filtres passe-bande idéaux, c'est-à-dire rectangulaires. Etant donné la complexité des phénomènes mis en jeu au niveau de la cochlée, cette hypothèse est évidemment très simplificatrice. En effet, des mesures plus précises des filtres auditifs ont montré leur caractère non linéaire et en particulier leur

dépendance vis-à-vis du niveau des stimuli. Les filtres auditifs ont en réalité une bande passante plus étroite pour les niveaux modérés que pour les niveaux forts [24].

Cependant, dans ce cadre très complexe, une modélisation par des filtres passe-bande idéaux constitue une première approche pratique et permet de quantifier approximativement la résolution fréquentielle du système auditif. Moore et Glasberg [25]. Ont proposé une mesure de la largeur de ces filtres idéaux appelée largeur de bande rectangulaire équivalente ou Equivalent Rectangular Bandwidth (ERB). Pour chaque filtre auditif, la valeur d'ERB est définie comme la largeur d'un filtre passe bande idéale de même fréquence centrale qui, alimenté par un bruit blanc, produirait la même énergie en sortie. La mesure des ERB illustre la résolution fréquentielle du système auditif.

La formule suivante donne la relation moyenne pour plusieurs normo-entendants entre la fréquence centrale f et l'ERB en Hz :

$$\boldsymbol{ERB(\ f\ ) = l + \frac{f}{q} avec, \qquad l = 24,7\ et\ q = 9,265} \qquad (\boldsymbol{II.1})$$

On peut définir une échelle des ERB. Chaque numéro d'ERB correspond alors à une distance fréquentielle d'une bande rectangulaire équivalente.

$$\boldsymbol{N°ERB = q \log_{10}\ (1 + \frac{f}{l.q})} \qquad (\boldsymbol{II.2})$$

Les filtres auditifs sont définis pour toute fréquence audible et la largeur des filtres augmente avec la fréquence, de telle sorte que cette largeur vaut approximativement 11% à 17% de la valeur de la fréquence centrale. Ces caractéristiques influent sur la résolution fréquentielle du système auditif, en effet, la discrimination fréquentielle de l'oreille est meilleure pour les basses fréquences que pour les hautes.

Schématiquement, on peut résumer ainsi les propriétés générales des filtres auditifs :

- Le sommet de chaque filtre est étroit et à peu près symétrique en fréquence.

- Les bords sont larges et asymétriques, en particulier leur pente du côté des basses fréquences est moins forte.

- Plus le niveau du stimulus est élevé, plus les filtres sont larges.

La modélisation par les filtres auditifs permet également une représentation schématique de l'excitation de la membrane basilaire en réponse à un stimulus. Cette représentation est

appelée pattern d'excitation, anglicisme tiré de l'expression excitation pattern. Moore et Glassberg ont proposé une méthode simple pour estimer le pattern d'excitation à partir de la sortie des filtres auditifs. Celui-ci est obtenu en relevant le niveau de sortie de chaque filtre en fonction de sa fréquence centrale.

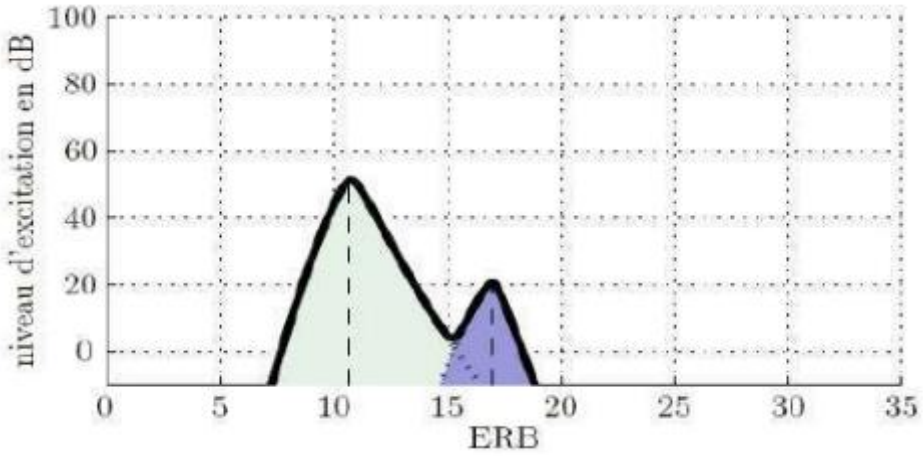


***Figure II.7** : Pattern d'excitation de deux sons purs de niveau modéré. F1 = 500Hz (50 dB SPL) et F2 = 1200Hz (20 dB SPL).[23]*

La figure II.7 montre un exemple de différents patterns d'excitation :

- celui d'un son pur à 500Hz (10,7 en échelle ERB) et d'intensité 50 dB SPL indiqué par la courbe englobant la zone claire centrée à 10,7 ERB,

- celui d'un son pur à 1200Hz (17 en échelle ERB) et d'intensité 20 dB SPL indiqué par la courbe englobant la zone d'aire foncée centrée à 17 ERB,

- enfin le pattern d'excitation complexe composé des deux fréquences précédentes, représenté par la courbe supérieure en trait gras.

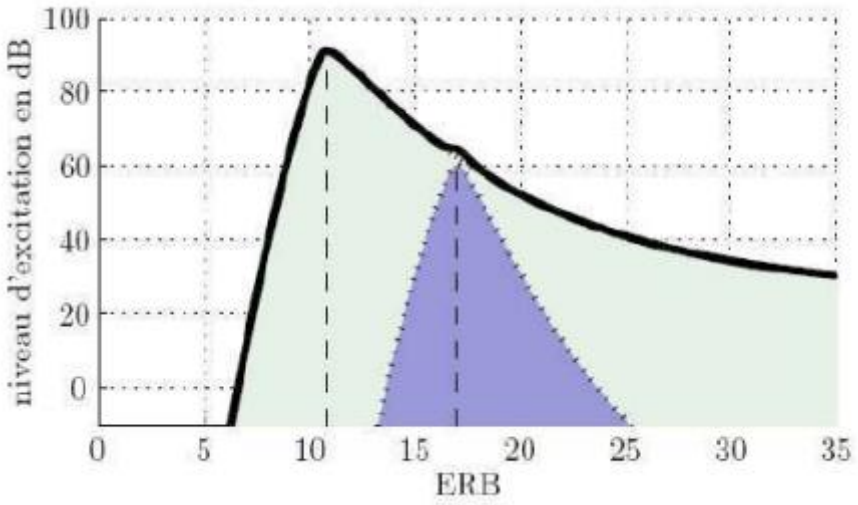


Figure II.8 : *Pattern d'excitation de deux sons purs de fort niveau. F1 = 500Hz (90 dB SPL) et F2 = 1200Hz (60 dB SPL).[23]*

Le pattern d'excitation est un outil intéressant car il permet d'une part de retrouver la tonotopie de la membrane basilaire, et d'autre part d'illustrer avec clarté le phénomène de masquage auditif. En effet, le stimulus de la figure II.7 présente deux pics fréquentiels distincts correspondant aux deux fréquences d'excitation. Si on considère maintenant le même exemple mais pour des sons purs rehaussés de 40dB chacun, on obtient le pattern d'excitation de la figure II.8. A ce niveau de stimulation, les filtres auditifs sont plus larges. On remarque alors que le pattern d'excitation du premier son pur indique une excitation beaucoup plus étendue sur la membrane basilaire que dans le premier exemple et surtout qu'en tout point de la membrane basilaire, ce premier apporte plus d'énergie que le second. Le second pattern d'excitation est ainsi complètement recouvert par le premier.

Le pattern d'excitation résultant 10 est pratiquement confondu avec le premier et n'indique plus qu'un seul pic fréquentiel : on voit donc ici clairement apparaître le phénomène de masquage. Notons que l'effet d'élargissement des filtres avec le niveau d'excitation ne sera pas explicitement simulé dans le banc de filtre mais plutôt dans un étage de compression non linéaire qui sera placé après, et qui produit des effets qualitativement similaires. [26]

II.4. Le masquage auditif

Le masquage auditif est un phénomène perceptif qui fait qu'un son n'est pas perçu lorsqu'il est en présence d'un autre son plus fort que lui, surtout si ces deux sons sont de fréquences proches.

La courbe de masquage démontrant ce phénomène présente une pente raide du côté des fréquences graves et une pente plus faible du côté des fréquences aigües, ainsi le masquage est plus important dans cette zone, c'est-à-dire que les graves masquent plus facilement les aigus. [27]

De manière générale, l'effet de masque se caractérise de deux manières :

- Un son fort masque plus facilement un son faible.

- Un son fort masque ceux dont la fréquence est supérieure.

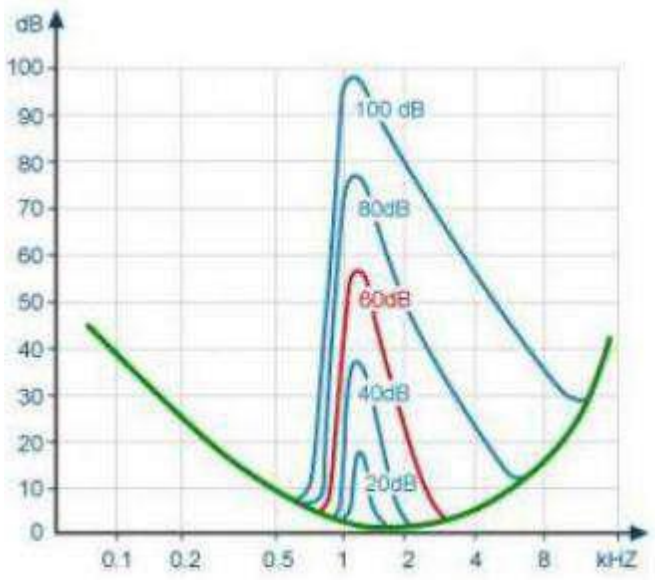


Figure II.9 : *Représentation de l'effet de masque[27]*

(Seuil auditif sans masquage en vert puis avec un masquage par un bruit (de 1 100 à 1 300 Hz) à divers niveaux d'intensité).

II.4.1 Les types de masquage

Il existe deux types de masquage :

II.4.1.1 Masquage simultané (ou fréquentiel)

Lorsque deux sons purs sont présentés à l'oreille, le ton le plus faible ne sera audible que si le rapport des puissances est suffisamment élevé dans un des filtres auditifs. Ce phénomène est connu sous le nom de phénomène de masquage et révèle le rôle important des filtres auditifs pour la compréhe-nsion de la sélectivité fréquentielle du système auditif.

Schématiquement, supposons que la distance fréquentielle entre deux tons soit inférieure à une largeur de bande de filtre auditif et considérons un filtre auditif de fréquence

centrale proche de celle du signal le plus faible (appelé signal masqué contrairement au premier appelé signal masquant). Les deux signaux se situent tous deux dans la bande passante du filtre qui a donc une faible influence sur le rapport des puissances des signaux. Le seuil d'audibilité du son masqué est donc relativement élevé.

Inversement, supposons que la distance fréquentielle entre les deux tons soit supérieure à une largeur de bande de filtre auditif. Le masqueur se situe alors dans la bande atténuée du filtre et le rapport des puissances est donc accru. Par conséquent le seuil d'audibilité du signal masqué est relativement faible.

Ainsi, l'effet d'un signal masquant est prépondérant dans la bande critique auquel il appartient. Cependant s'il est de niveau acoustique suffisant, il peut également modifier les seuils d'audition dans les bandes adjacentes (spread of masking).

Ce mécanisme, décrit très simplement ici, illustre le pouvoir de séparabilité ou de résolution fréquentielle du système auditif.

La figure II.10 donne un exemple de courbes d'effet de masque indiquant les seuils d'audibilité d'un son pur en fonction de sa fréquence pour un bruit de bande masquant présenté à différents niveaux d'intensité.

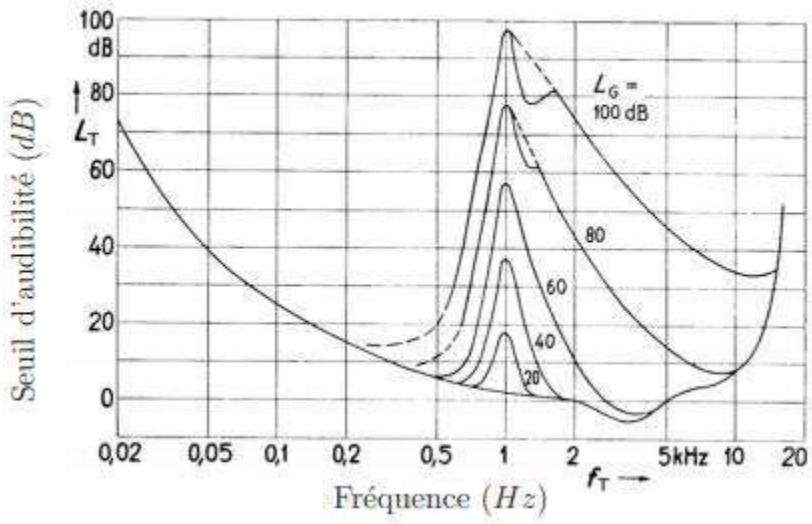


Figure II.10 : *Courbes d'effet de masque de bruits à bande étroite. Le signal masquant est un bruit de bande centré à 1 kHz de largeur 160 Hz présenté aux niveaux LG = 100, 80,60, 40 et 20 dB[27]*

Dans l'expérience illustrée par la figure précédente le masqueur est fixe, tandis que le signal masqué varie en fréquence. Une alternative consiste à garder le signal constant tandis qu'on fait varier le masqueur. Les courbes ainsi obtenues sont appelées courbes de sélectivité

de fréquence Psycho acoustique (psychophysical tuning curves, PTC). Plusieurs exemples de PTC sont représentés sur la figure suivante. Pour chaque courbe, le signal de référence est présenté à 10 dB SL (indiqué par un point) et le niveau maximal du signal masquant tel que le signal de référence soit juste audible est tracé en fonction de sa fréquence.

Enfin, le signal étant fixe, on peut supposer approximativement qu'un seul et même filtre auditif est sollicité. Chaque PTC représente donc approximativement la forme inverse du filtre auditif dont la fréquence centrale est celle du signal. Si on fait l'hypothèse que les filtres auditifs sont linéaires, on peut alors obtenir leurs formes en inversant l'allure des PTC obtenues.

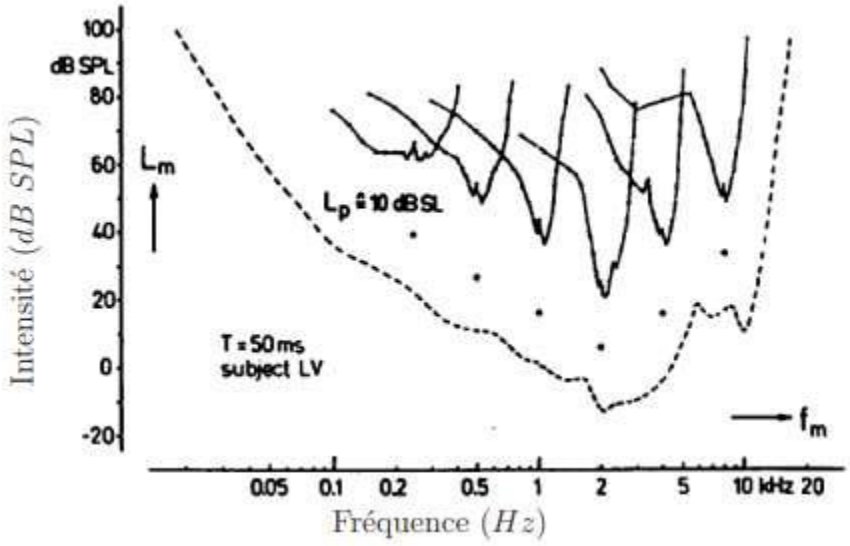


Figure II.11 : *Courbes de sélectivité de fréquence Psycho acoustique (psychophysical tuning curves)[27]*

Remarque

Le phénomène de masquage a été récemment utilisé avec succès dans les applications de codage audio. L'originalité de ce type de codage est de faire en sorte que le bruit de quantification soit masqué par le signal utile.

II.4.1.2. Masquage non-simultané (ou temporel)

Le phénomène de masquage est également observable pour des sons non-simultanés. On parle alors de masquage non-simultané ou masquage temporel. Dans ce type de masquage, un son fort présenté avant ou après un stimulus peut masquer ce dernier. La figure II.12 illustre ce phénomène. [32]

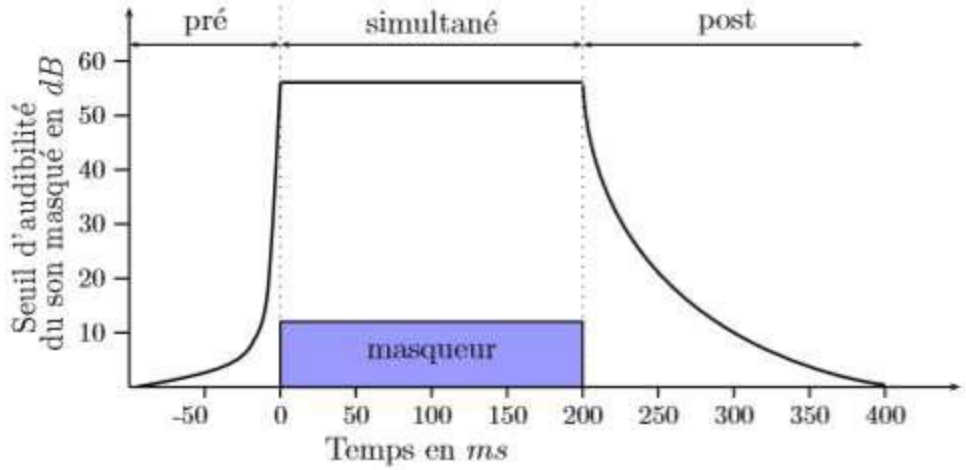


Figure II.12: *Masquage non-simultané[32]*

II.5. Le banc de filtres Gammatone

II.5.1. Introduction sur les filtres Gammatone

Les filtres Gammatone sont souvent utilisés pour la modélisation du traitement du son dans la cochlée [28]. Plusieurs implémentations de ces filtres sont connues, toutes proviennent de diverses approximations du filtre Gammatone original.

Pour notre modèle auditif, il a été retenu un banc de filtres Gammatone tout pôle d'ordre 4, à sortie complexe, proposée par Hofmann [29]. Afin de pouvoir générer rapidement le filtre, ce dernier a donc été simplifié par omission du numérateur, en effet, il est dérivé de la version échantillonnée de la réponse impulsionnelle Gammatone analogique :

$$\mathbf{g_\gamma(n) = n^{\gamma-1} \cdot \tilde{a}^n, \quad n \geq 0 \quad avec \quad \tilde{a} = \lambda \cdot exp(i\beta)} \qquad \mathbf{(II.3)}$$

Où λ traduit l'atténuation, β la fréquence d'oscillation, γ l'ordre du filtre de n le temps échantillonné, $\tilde{\alpha}$ est un coefficient complexe. Pour γ=4, on obtient la transformée en z suivante :

$$\boldsymbol{G(z) = \frac{\tilde{a}z^{-1} + 4(\tilde{a}z^{-1})^2 + (\tilde{a}z^{-1})^3}{(1 - \tilde{a}z^{-1})^4}} \qquad \mathbf{(II.4)}$$

On simplifie alors cette fonction de transfert pour ne garder finalement que le dénominateur :

$$\boldsymbol{k(z) = \frac{1}{(1 - \tilde{a}z^{-1})^4}} \qquad \mathbf{(II.5)}$$

Les pôles d'une fonction de transfert ayant généralement une influence bien plus importante sur la réponse en fréquence que les zéros pour les filtres à réponse impulsionnelle infinie (Infinité Impulse Réponse, IIR), tout laisse à penser que la réponse en fréquence ne sera que légèrement modifiée.

On peut effectivement le vérifier sur la figure (II.13) : ne pas prendre en compte le numérateur de la fonction de transfert ne modifie que très peu la réponse en fréquence, et la plupart de ces différences sont situées dans la bande coupée. Par contre, comme les effets sur la réponse temporelle peuvent être plus prononcés, nous allons calculer l'erreur temporelle commise par cette approximation.

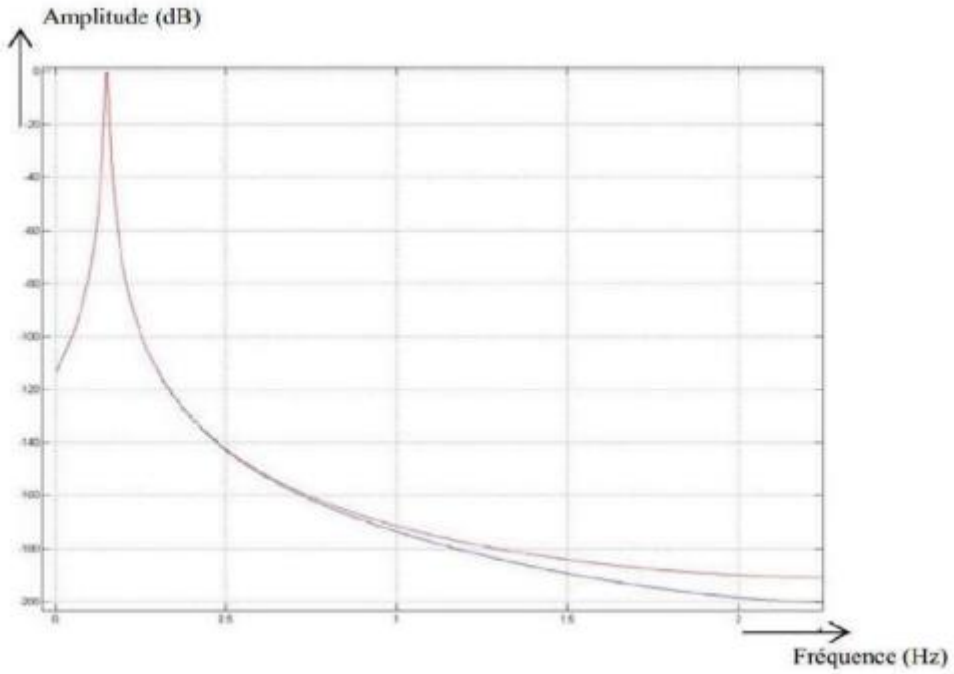


Figure II.13 : *Réponse en fréquence des fonction G(z) (en bleu) et K(z) (en rouge) pour* $\tilde{a}$*=0.9693 + 0.2104i, correspondant à une fréquence centrale de 1,5kHz et une largeur de bande de 50Hz.[26]*

L'expression temporelle k(n) de la fonction de transfert K(z) est la suivante :

$$\boldsymbol{k(n) = \frac{\tilde{a}}{6}(n^3 + 6n^2 + 11n + 6)} \qquad (\boldsymbol{II.6})$$

En négligeant le facteur 1/6, l'erreur absolue e(n) et relative r(n) de k(n) par rapport à g(n) est :

$$\boldsymbol{e(n) = 6 \cdot k(n) - g(n) = g(n)\left(\frac{6}{n} + \frac{11}{n^2} + \frac{6}{n^3}\right) \qquad n > 0}$$

$$r(n) = \frac{e(n)}{g(n)} = \frac{6}{n} + \frac{11}{n^2} + \frac{6}{n^3} \quad (\mathbf{II.7})$$

L'erreur relative étant en 1/n, la forme et la structure temporelle fine du filtre Gammatone, qui ont une importance majeure pour la modélisation, notamment la caractéristique de phase de la cochlée, sont préservées par cette approximation tout-pôle. [26]

II.5.2. Quatre filtres de la famille des Gammatone

II.5.2.1. Le filtre Gammatone (GTF)

Ce filtre a été extrêmement populaire, principalement en raison de sa description simple dans le domaine temporel comme une enveloppe de distribution gamma multipliée par une tonalité. Il a été mis en œuvre (rendu exécutable) par un certain nombre de méthodes et d'approximations, mais généralement pas dans le domaine temporel. Approximations, mais généralement pas de la manière la plus directe en se basant sur sa décomposition pôle-zéro dans le domaine de Laplace, car cette décomposition n'a pas été largement comprise jusqu'au au moins milieu des années 1990. Le GTF a une réponse en fréquence intrinsèquement très symétrique, ce qui n'est pas une bonne aux données auditives, mais il a une meilleure forme de crête/jupe que les filtres roex.

II.5.2.2. Le filtre GammaChirp (GCF)

Est une généralisation du filtre GTF qui permet une asymétrie réaliste et contrôlable dans le domaine de la fréquence, et une asymétrie correspondante dans le domaine de l'énergie.

Fréquence et une asymétrie réaliste correspondante dans le domaine temporel, ou un "glissement" dans la fréquence instantanée de sa réponse impulsionnelle. D'impulsion. Mais il ne dispose pas d'une décomposition pôle-zéro, et nécessite d'autres approximations dans sa mise en œuvre.

II.5.2.3. Le filtre Gammatone à tous les pôles (APGF)

Est une autre approche de fournir une asymétrie réaliste, tout en simplifiant le domaine de Laplace.

Tout en simplifiant la description du domaine de Laplace et la mise en œuvre du GTF. L'APGF a été utilisé comme une approximation pour la mise en œuvre du GTF, mais il présente des avantages propres.

II.5.2.4. Le filtre gammatone à un zéro (OZGF)

Est une légère généralisation de l'APGF qui, en ajoutant un seul zéro réel dans le domaine de Laplace, permet d'obtenir un bon contrôle. Dans le domaine de Laplace, il permet un bon contrôle de la forme de la queue basse fréquence. [30]

II.5.3. Le filtre Gammatone dans le domaine temporel

Le filtre Gammatone est défini dans le domaine temporel (la réponse impulsionnelle) par :

$$\boldsymbol{G(t)t\alpha t^{n-1}\exp(-2\pi bt)\cos(2\pi f_0 t+\phi)} \qquad \boldsymbol{(t>0)} \qquad \textbf{(II.8)}$$

Il s'agit donc d'un filtre causal avec une réponse temporelle infinie. La forme de cette réponse impulsionnelle est d'une porteuse (Tone) de fréquence $fo[Hz]$modulée en amplitude, avec une enveloppe proportionnelle à $\boldsymbol{t^{(n-1)}exp\,(-2bt)}$, qui est la distribution Gamma. Ces deux fonctions donnent naissance au nom Gammatone (de Boer et de Jongh, 1978)

Les paramètres du filtre Gammatone sont :

- l'ordre du filtre $\boldsymbol{n}$,qui (pour une valeur de **b** fixe) contrôle la forme relative de l'enveloppe qui devient oblique lorsque $\boldsymbol{n}$ augmente;
- $\boldsymbol{b}$ (en $\boldsymbol{Hz}$) qui (pour **n** fixe) contrôle la durée de la réponse impulsionnelle, l'augmentation de b implique la diminution de la durée;
- $\boldsymbol{f_0}(en\ \boldsymbol{Hz})$ qui détermine la fréquence de la porteuse;
- Ø (en radians), la phase du porteuse, Qui détermine la position relative de la structure fine du porteuse sur l'enveloppe.[26]

II.5.4. Le filtre Gammatone dans le domaine fréquentiel

Si GT(f) représente le filtre Gammatone dans le domaine fréquentiel (fonction de réponse en fréquence), alors :

$$\boldsymbol{GT(f)\alpha\left[1+\frac{j(f-f_0)}{b}\right]^{-n}+\left[1+\frac{j(f+f_0)}{b}\right]^{-n},(-\infty<f<\infty)} \qquad \textbf{(II.9)}$$

Cela peut être obtenu soit directement à partir de la définition dans le domaine temporel décrit ci-dessus, par l'application de la transformée de Fourier, soit en utilisant la propriété que le produit dans le domaine temporel d'une fonction Gamma et une fonction cosinus correspondra à la convolution dans le domaine fréquentiel de la transformée de Fourier $\boldsymbol{(1+jf/b)^{-n}}$ de la fonction Gamma avec une distribution à deux points à $\boldsymbol{\pm f_0}$. Ici pour

simplifier nous avons mis $\emptyset = \mathbf{0}$, car il n'a pas d'effet important sur les caractéristiques du filtre dans le domaine fréquentiel.

Le rôle des paramètres dans le domaine fréquentiel est clair dans la formule ci-dessus. f_0 Est la fréquence centrale du filtre ; Pour un ordre du filtre **n** fixe, b agit comme un paramètre d'échelle tel que le la bande passante du filtre augmente avec **b** ; Le paramètre **n** contrôle la forme globale du filtre. Pour une valeur **b** fixe, la largeur de bande diminue lorsque **n** augmente. GT(f) est approximativement symétrique par rapport à f_0 sur une échelle de fréquence linéaire.

Comme indiqué par de Boer et Krudenier (**1988**), le deuxième terme de l'équation (**II.9**) peut être ignoré lorsque f_0/b est suffisamment grand, ce qui est toujours le cas lors de la modélisation du filtre auditif humain.

- donnant une fonction de réponse de fréquence approximative

$$GT(f) \approx \left[1 + \frac{j(f - f_0)}{b}\right]^{-n}, 0 < f < \infty \qquad (II.10)$$

Pour les filtres Gammatone nécessaires à la modélisation du système auditif humain ($4 < f0/b < 8$), Le deuxième terme de l'équation **(II .9)** est typiquement inférieur de **25 dB** au niveau du premier terme, même à deux Bande passante en dessous de la fréquence centrale.

En raison de la **n-ième** forme de puissance de l'équation (II .10), le filtre Gammatone d'ordre **n** peut être approximé par une cascade de **n** filtres Gammatone identiques d'ordre **1** avec la réponse fréquentielle :

$$[1 + j(f - f_0)/b]^{-n} \qquad (II.11)$$

Chacun ayant la forme d'un filtre passe-bas à décalage de fréquence - un filtre de premier ordre qui a une application numérique récursive bien connue.

II .5.5. Largeur de bande rectangulaire équivalente du filtre Gammatone :

Lorsque fo / b est grand, la largeur de bande rectangulaire équivalente du filtre Gammatone est proportionnelle à **b**, et la constante de proportionnalité an ne dépend que de **n**.

Nous pouvons calculer la bande passante rectangulaire équivalente du filtre Gammatone en appliquant le théorème de Parseval.

Si l'on suppose que $|GT(f)|^2$ est symétrique dans f, et que sa valeur maximale se produit lorsque $f = fo$ alors :

$$\int_{-\infty}^{+\infty} gt\,(t)^2 dt = \int_{-\infty}^{+\infty} |GT(f)|^2\, df = 2|GT(fo)|^2 ERB[GT] \qquad (II.12)$$

Où **ERB [GT]** est la bande passante rectangulaire équivalente du filtre Gammatone.

L'application de l'équation **(II .12)** au filtre Gammatone conduit à une intégrale traitable, ce qui donne la bonne approximation pour le constant a_n :

$$a_n = \frac{\Pi(2n-2)!2^{-(2n-2)}}{(n-1)!} \qquad (II.13)$$

La bande passante à ***3dB*** d'un filtre Gammatone d'ordre n est $c_n b$ ***[Hz]*** où

$$c_n = 2\sqrt{2^{1/n}-1} \qquad (II.14)$$

Le tableau **(II .1)** donne les valeurs de a_n et c_n pour le filtre Gammatone décrit ci-dessus.

n	a_n	a_n^{-1}	c_n	c_n^{-1}
1	3,141	0,318	2,000	0,500
2	1,570	0,637	1,288	0,777
3	1,178	0,849	1,020	0,980
4	0,982	1,019	0,870	1,149
5	0,889	1,164	0,772	1,296
6	0,773	1,293	0,700	1,429
7	0,709	1,411	0,646	1,550
8	0,658	1,520	0,602	1,662
9	0,617	1,621	0,566	1,767

Tableau II .1 : *Les données pour le filtre Gammatone.*

Moore et Glasberg (**1983**) donnent la formule suivante pour la largeur de bande rectangulaire équivalente (en **Hz**) d'un filtre auditif humain d'une fréquence centrale f_0 **[Hz]**:

$$ERB[F0] = 6,23 \times 10^{-6} f_0^2 + 93,39 \times 10^{-3} f_0 + 28,52 \quad (100 < f_0 < 10000)$$

(II.15)

Ainsi, pour faire correspondre un filtre Gammatone d'ordre **4** et d'une fréquence centrale **1000 Hz**, la valeur de **b** est $\boldsymbol{a_n^{-1} ERB\ [1000] = 1,019 \times 128,14 = 130,57}$. La largeur de bande à **3dB** de ce filtre est $\boldsymbol{c_n\ b = 0,870 \times 130,57 = 113,59\ HZ}$.

II .5.6. Compensation de phase

Lors de la représentation graphique de la sortie d'un banc de filtres, il est souvent utile d'aligner les pics des enveloppes des fonctions de la réponse impulsionnelle et d'aligner le pic de la structure fine avec le pic de chaque enveloppe. Tout d'abord, les sommets des enveloppes peuvent être alignés en introduisant un facteur $\boldsymbol{t_c = (n - 1) / 2\pi b}$ à la sortie du filtre.

Deuxièmement, un pic dans la structure fine peut être déplacé vers le pic de l'enveloppe en introduisant une correction de phase $\boldsymbol{\emptyset_c = -2\pi f_0 t_c}$.Cela conduit au filtre (non causal)

$$\boldsymbol{\tilde{g}t(t) \propto (t + t_c)^{n-1} \exp[-2\pi b(t + t_c)] \cos(2\pi f_0 t) \qquad t \geq -t_c} \qquad \textbf{(II.16)}$$

Qui aligne le pic de la réponse impulsionnelle de tous les filtres à **t =0** [31].

II.6. Conclusion :

Dans ce chapitre, nous avons décrit le mécanisme de l'audition, ainsi que la modélisation de la cochlée pathologique par un banc de filtres plus précisément le filtre Gammatone utilisé dans notre projet,

Chapitre 3 :

La simulation et les résultats de la SAS

III.1. L'introduction

Dans ce chapitre, nous nous présentons les résultats de simulation d'un algorithme de séparation de sources.

Nous avons utilisé le logiciel MATLAB pour obtenir un bon résultat. On commence par la première partie qui expliquer le principe du banc de filtres Gammatone en générant sa réponse impulsionnelle et en calculant son cochléagramme dans l'algorithme MGD IS-NMF2D afin d'estimer les sources à partir d'un mélange.

III.2. C'est quoi le logiciel MATLAB ?

C'est le diminutif de Matrix Laboratory, Matlab est un environnement de programmation orienté calcul numérique. Articulé autour du langage de script du même nom, il est doté d'un éditeur permettant d'exécuter des séquences de commandes encapsulées dans des fonctions. C'est une plateforme de calcul numérique et de programmation utilisée par des millions d'ingénieurs et de scientifiques pour analyser des données, développer des algorithmes et créer des modèles.

Les utilisateurs de MATLAB (environ 4 millions en 2019) sont de milieux très différents tels que l'ingénierie, les sciences et l'économie, dans un contexte aussi bien industriel que pour la recherche. Matlab peut s'utiliser seul ou bien avec des toolboxes (« boîte à outils »).[33]

MATLAB est constitué d'un noyau relativement réduit, capable d'interpréter puis d'évaluer les expressions numériques matricielles qui lui sont adressées :

- Soit directement au clavier depuis une fenêtre de commande ;
- Soit sous forme de séquences d'expressions ou scripts enregistrées dans des fichiers-texte appelés m-files et exécutées depuis la fenêtre de commande ;
- Soit plus rarement sous forme de fichiers binaires appelés mex-files ou fichiers .mex générés à partir d'un compilateur C ou fortran.

Ce noyau est complété par une bibliothèque de fonctions prédéfinies, très souvent sous forme de fichiers m-files, et regroupes en paquetages ou toolboxes. A côté des toolboxes requises local et Matlab, il est possible d'ajouter des toolboxes spécifiques à tel ou tel problème mathématique, Optimization Toolbox, Signal Processing Toolbox par exemple ou encore des toolboxes crées par l'utilisateur lui-même.

Un système de chemin d'accès ou path permet de préciser la liste des répertoires dans lesquels MATLAB trouvera les différents fichiers m-files.

III.2.1. Le son sous MATLAB

Windows propose un standard des fichiers son : Le format '.wav' ('.au' est le Standard chez unix). Ce format est rudimentaire et non compressé. Comme tout fichier formaté, un fichier '.wav' comporte un entête : Champs de 44 octets, suivie des données (le son proprement dit).

III.2.2. Lecture d'un fichier son

Pour lire un fichier '.wav' dans Matlab, nous devons exécuter la commande suivante :

[x,fs]=audioread('pianotrump.wav');

Où *x* est un vecteur colonne ; et *fs* est la fréquence d'échantillonnage en Hertz.

III.2.3. Ecriture d'un fichier son

La commande suivante :

audiowrite(x,fs,'son.wav'); Permet de crée un fichier son a un nom « *son.wav* » dont les échantillons (compris entre -1 et 1) se trouvent dans le vecteur *x* avec une fréquence d'échantillonnage *fs*.

III.3. Les données de traitement

Dans notre projet, un fichier audio contient un mélange de deux sons instruments (piano et trompette), lire comme suit :

[x,fs]=wavread('pianotrump.wav') ;

Avec :

x : représente le signal d'entrée qui contient le mélange de deux signaux

fs : est la fréquence d'échantillonnage en Hertz.[34]

III.4. Banc de filtre Gammatone

Le banc de filtres Gammatone est un modèle de filtrage cochléaire qui décompose un signal d'entrée dans le domaine temps-fréquence (TF) à l'aide d'un ensemble de filtres Gammatone. La réponse impulsionnelle d'un filtre Gammatone centré en fréquence est mathématiquement exprimée par :

$$g(f,t) = \begin{cases} t^{h_1} e^{-2\pi v t} \cos(2\pi f t), & t \geq 0 \\ 0, & else \end{cases} \qquad (III.1)$$

Où :

h : désigne l'ordre du filtre,

v : représente la bande passante rectangulaire qui augmente à mesure que la fréquence centrale ***f*** augmente.

En considérant un canal de filtre particulier ***c*** avec ***f*** comme fréquence centrale, la réponse de sortie du filtre $x(c,t)$ peut être exprimée comme suit :

$$x(c,t) = x(t) * g(f_c,t) \qquad (III.2)$$

Où $"*"$ désigne la convolution dans le domaine temporel. La réponse est décalée vers l'arrière de $(h-1)/(2\pi v)$ pour compenser le retard du filtre. La sortie de chaque canal de filtrage est divisée en trames temporelles avec un chevauchement de 50% entre les trames consécutives.

Il y a certaines différences cruciales existent dans la représentation TF de la façon dont le son est analysé par l'oreille. En particulier, les sous-bandes de fréquences de l'oreille s'élargissent pour des fréquences plus élevées, Puisque les signaux vocaux sont caractérisés comme hautement non stationnaires et non périodiques alors que la musique change continuellement ; par conséquent, l'application de la transformée de Fourier produira des erreurs lorsque des phénomènes transitoires compliqués tels que le mélange de deux instruments sont contenus dans le signal analysé. [35]

III.5. Choix de l'échelle fréquentielle

Les bandes passantes du banc de filtres, supposés être idéaux, doivent être disjointes mais leur union doit couvrir la bande utile du signal. La largeur de ces bandes peut être uniforme, ou non uniforme, elle suit dans ce cas, les caractéristiques des bandes critiques, la formule qui permet de convertir une fréquence *f* en Hertz de sa valeur *ERB* est :

$$\boldsymbol{cf = (10.\hat{}(erb/21.4) - 1)/4.37e - 3} \qquad (\boldsymbol{III.3})$$

Le système auditif humain ne perçoit pas les signaux selon une échelle uniforme. Notamment, il effectue une analyse plus précise dans les fréquences graves que dans les bandes hautes du spectre. Ainsi, réalisant l'analyse spectrale selon une échelle en fréquence conforme à notre perception auditive, l'échelle *ERB* est la largeur des bandes critiques et elle modélise la résolution fréquentielle des filtres auditifs. La relation moyenne pour plusieurs sujets "normaux" (à propos de l'audition) entre la fréquence centrale *f* en Hz et l'*ERB* en Hz est la suivante [34]:

$$\boldsymbol{b = 1.019 * 24.7 * \left(4.37 * \frac{cf}{1000} + 1\right)} \qquad (\boldsymbol{III.4})$$

III.6. Réalisation d'un filtre gammatone

Pour réaliser notre banc de filtres Gammatone, nous prenons :

- Nombre de canaux=**128** implique **N=2^7** qui représente la taille de signal.
- Une bande fréquentielle de travail entre **50** et **8000 Hz** (le spectre audible d'une oreille saine est compris entre **20** et **20000Hz**)
- $\boldsymbol{f_s = 16kHz}$

La figure suivante montre le tracé échantillonné du signal de mélange choisi.

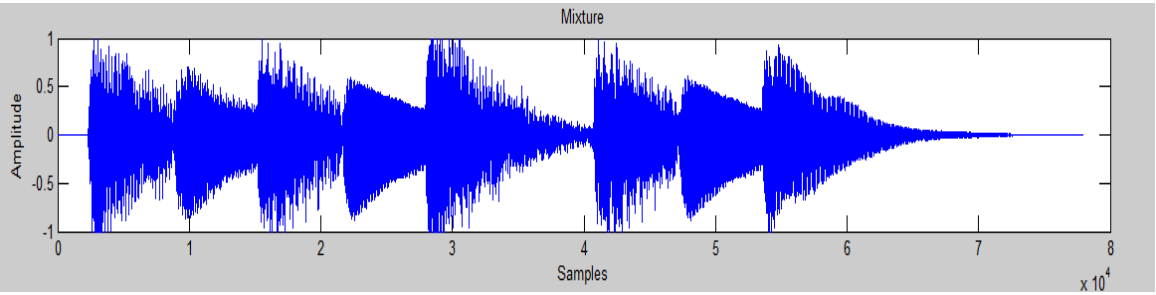


Figure III. 1. *Le tracé du mélange (piano et trompette) à séparés.*

Résultats

La réponse impulsionnelle d'un filtre Gammatone

On va générer un banc de filtre Gammatone avec trois filtres de différentes fréquence (fréquence centrale) pour dessiner la réponse impulsionnelle.

```
%créer un banc de filtre gammatonne
% écrit par Yang Shao, et adapted par Xiaojia Zhao en Oct'11
sampFreq =8000; % fréquence d'echantillonnage ,
numChannel = 128; % nombre de canaux,
stepBound=600;
filterOrder=4; % commande de filtre
phase(1:numChannel)=zeros(numChannel,1); % phase original
erb_b=hz2erb([50,sampFreq/2]); % limite suppérieure et inférieure de l'ERB
erb=[erb_b(1):diff(erb_b)/(numChannel-1):erb_b(2)]; % segment ERB
cf=erb2hz(erb); % fréquence centrale
b=1.019*24.7*(4.37*cf/1000+1); % Taux de décomposition
gt=zeros(numChannel,stepBound);
tmp_t=[1:stepBound]/sampFreq;
for i=1:numChannel
gain=10^((loudness(cf(i))-60)/20)/3*(2*pi*b(i)/sampFreq).^4;
gt(i,:)=gain*sampFreq^3*tmp_t.^(filterOrder-1).*exp(-
2*pi*b(i)*tmp_t).*cos(2*pi*cf(i)*tmp_t+phase(i));
end
figure
plot(tmp_t,gt(100,:))
xlabel('Temps(s)')
ylabel('Amplitude')
```

La fonction "*loudness*" dans le code précédant permet d'ajuster le signal à l'effet de l'oreille moyenne selon le standard *ISOBS3383*.

Les figures suivantes démontrent la réponse impulsionnelle du canal 3, 50 et 100.

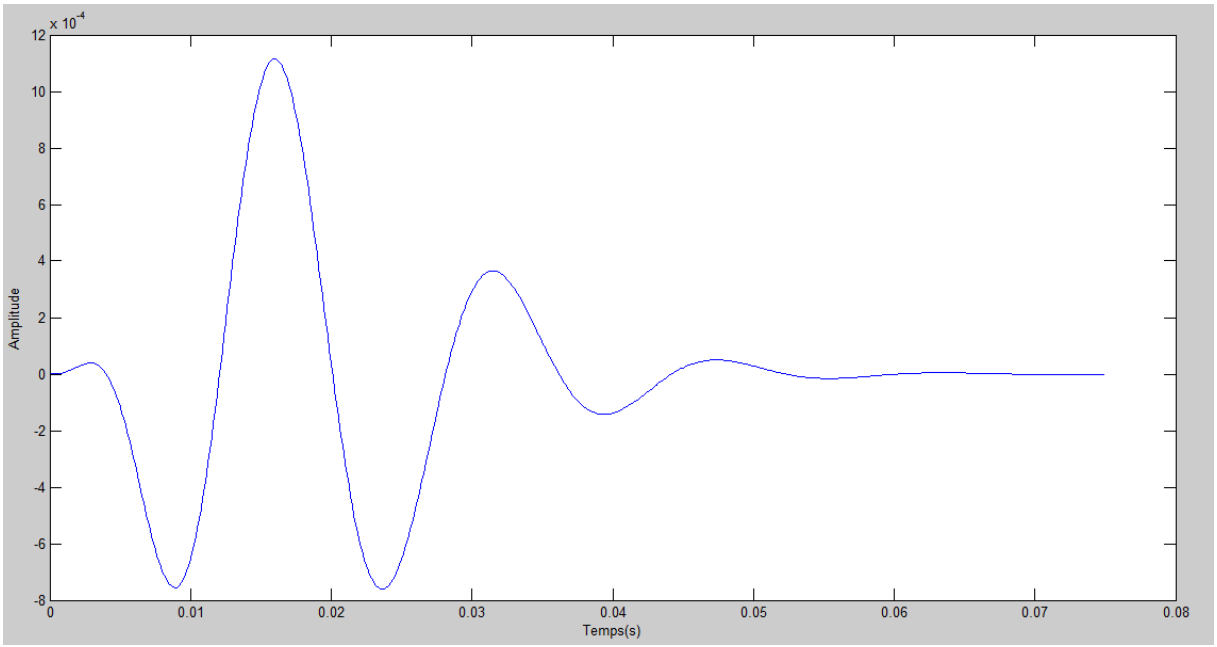


Figure III. 2. *Réponse impulsionnelle de la bande passante N°3.*

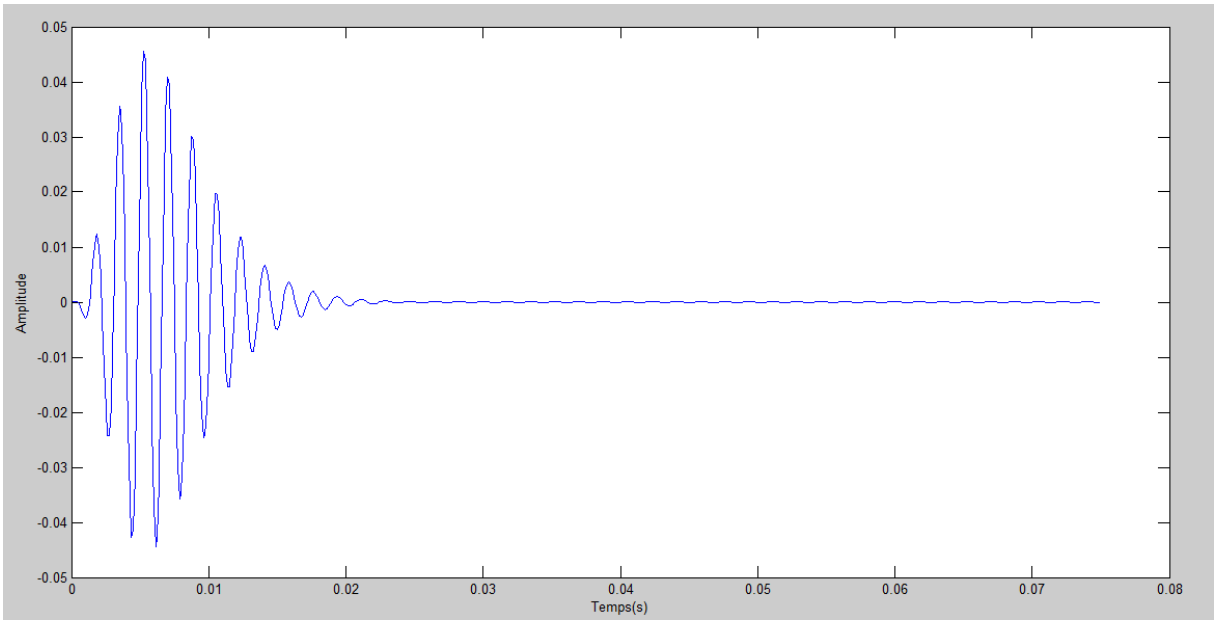


Figure III. 3. *Réponse impulsionnelle de la bande passante N°50.*

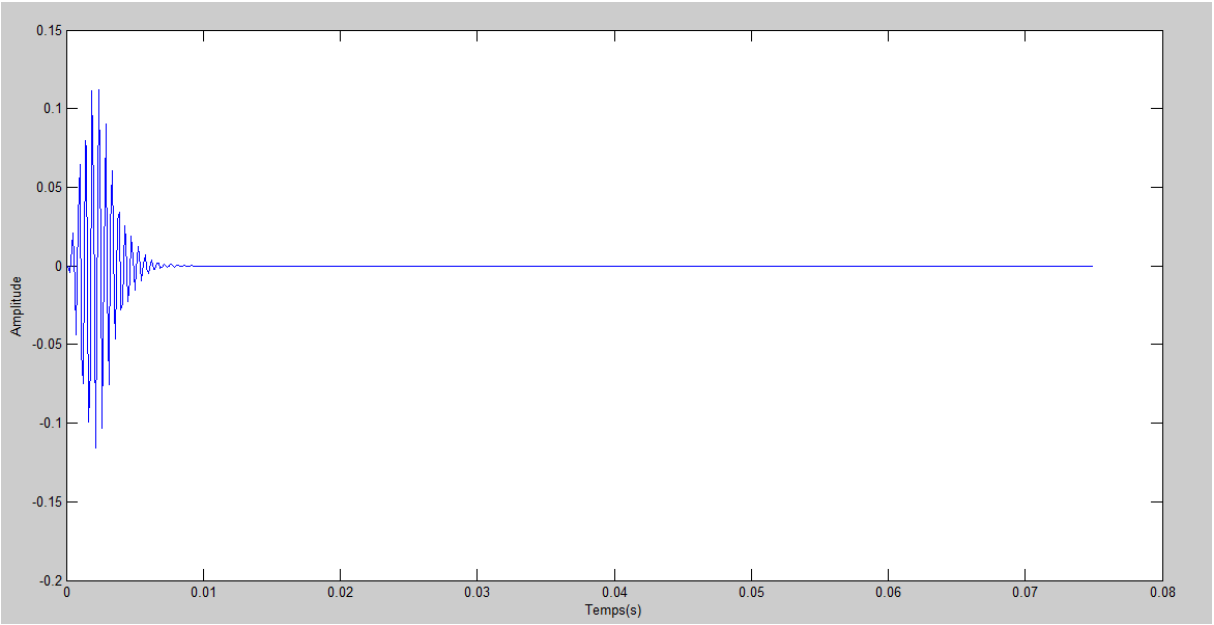


Figure III. 4. *Réponse impulsionnelle de la bande passante N°100.*

Réponse en fréquence d'un filtre Gammatone

La réponse en fréquence est généralement caractérisée par la transformé de fourrier de la Réponse impulsionnelle.

Alors on va réaliser la réponse d'un banc de filtre de 12 bandes :

```
nfft=1024;
fs=16000; % fréquence d'échantillonnage
nb=12; % nombre de bande
minfreq=50;
maxfreq=8000;
long_max=512;
gtm = fft2gammatonemx(nfft,fe,nb,1,minfreq,maxfreq,long_max);
f = [0:511]/512*8000;
figure
plot(f,20*log10(gtm)','')
xlabel('Fréquence(Hz)')
ylabel('Amplitude (db)')
axis([0 8000 -60 0])
grid
```

Avec : la fonction fft2gammatone est une fonction qui génère une matrice des poids de filtre Gammatone à partir des coefficients du FFT.

La figure suivante montre la réponse en fréquence d'un banc générale de filtres à gammatone :

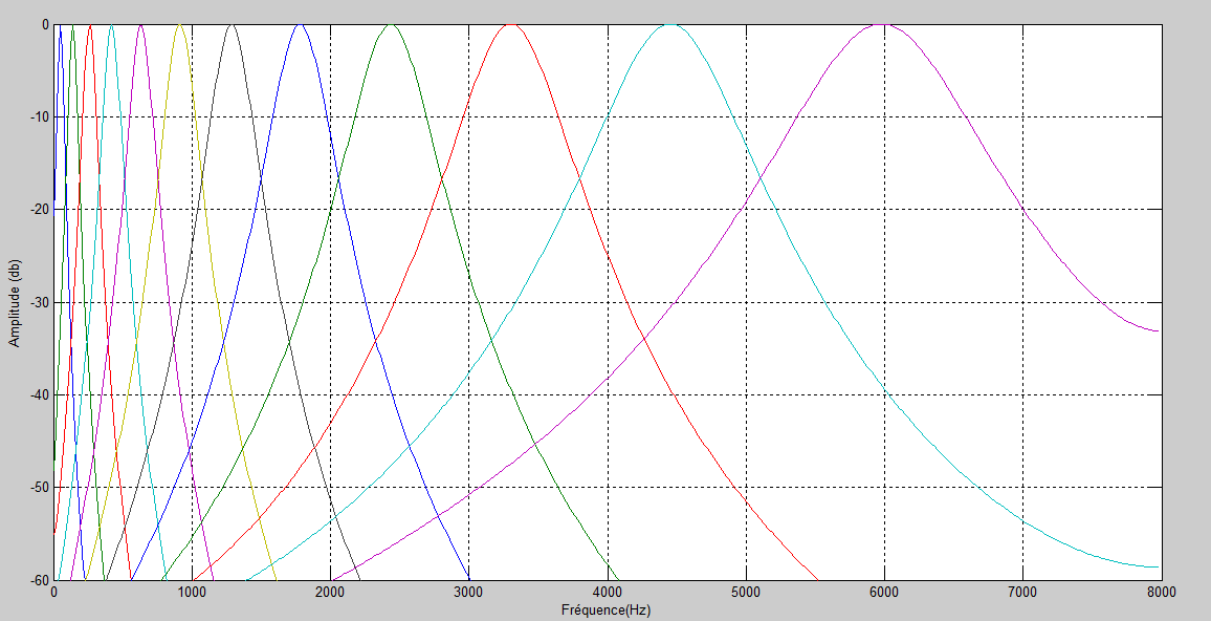


Figure III.5 : *Réponse en fréquence d'un banc de filtre Gammatone 12 bandes*

On voit que les fréquences plus élevées correspondent aux sous-bandes de fréquences plus larges qui ressemblent étroitement à la perception humaine des fréquences.

III.7. Transformée de Fourier d'un filtre Gammatone

La transformée de Fourier d'une fonction g(t) est définie comme suit :

$$\boldsymbol{G(f) = F\,[g(t)](f) = \int_{-\infty}^{+\infty} g(t)e^{-2\pi ift}\,(t)dt} \qquad (\boldsymbol{III.5})$$

En utilisant la propriété stipulant que la transformée de Fourier d'un produit est le produit de convolution des transformées de Fourier, on remarque pour la fonction Gammatone que :

$$\boldsymbol{G(f) = F\,[cos\,(2\pi\, fc\, t)]\, * F\,[t^{n-1}e^{-2\pi bt}\,]} \qquad (\boldsymbol{III.6})$$

Sachant que $F\,[cos\,(2\pi fCt)]\, =\, 1\,/2\,(\delta\,(f\, -\, fc)\, +\, \delta\,(f\, +\, fc))$ [36]

Alors après la génération des réponses impulsionnelles des filtres Gammatone, nous devons filtrer le signal du mélange par le banc de filtres en utilisant la commande *fftfilt* du Matlab.

```
clear
clc
fRange = [50, 8000];
fs=16000;
numChan=128;
%% lecture d'un signal audio
x=audioread('pianotrump.wav');
%% filtrage du signal audio en utilisant un banc de 128 filtres Gammatones
r = gammatoneFast(x, numChan, fRange, fs)
%% visualisation des tracés temporelles
t=(length(x)-1)/fs;
figure
plot(r(3,:))
xlabel('Echantillon')
ylabel('Amplitude')
```

Résultats :

On prend ici 3 filtres Gammatone parmis 128 pour trouver la visualisation des tracés temporelles, donc les figures suivantes montrent le signal de sortie du la 3ème ,50ème et 100ème filtres dans le banc réalisé, après filtrage :

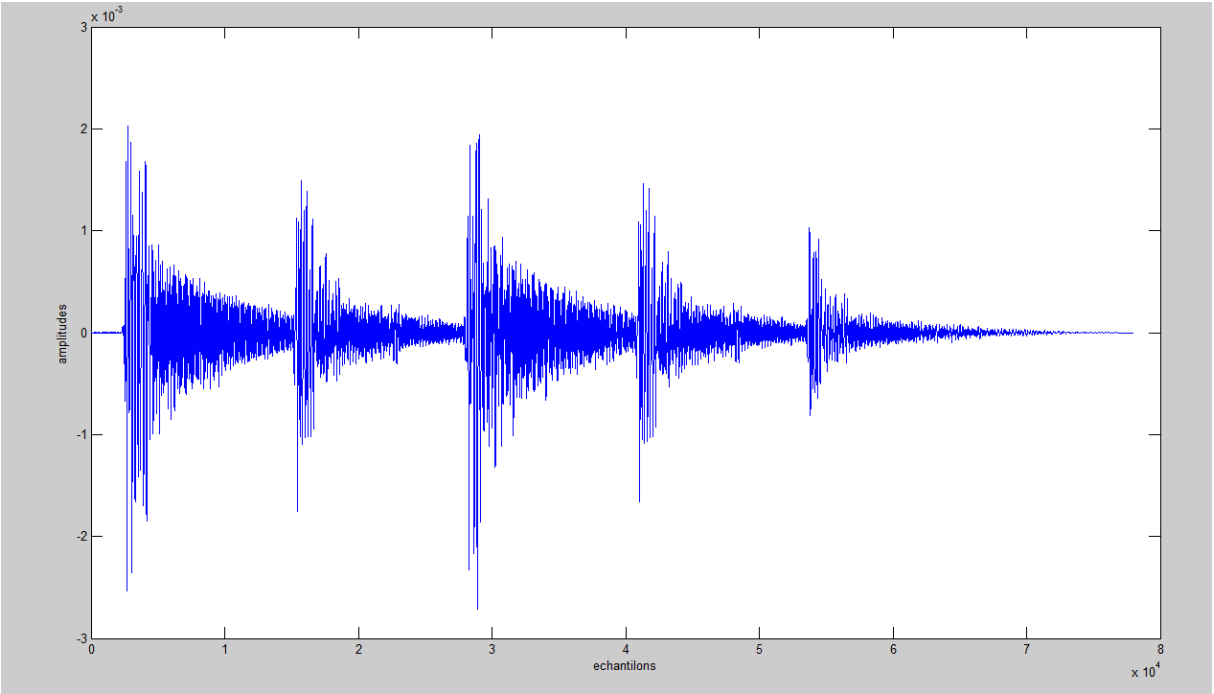


Figure III.6. *Le Signal de filtre Gammatone filtré de la bande passante N°3.*

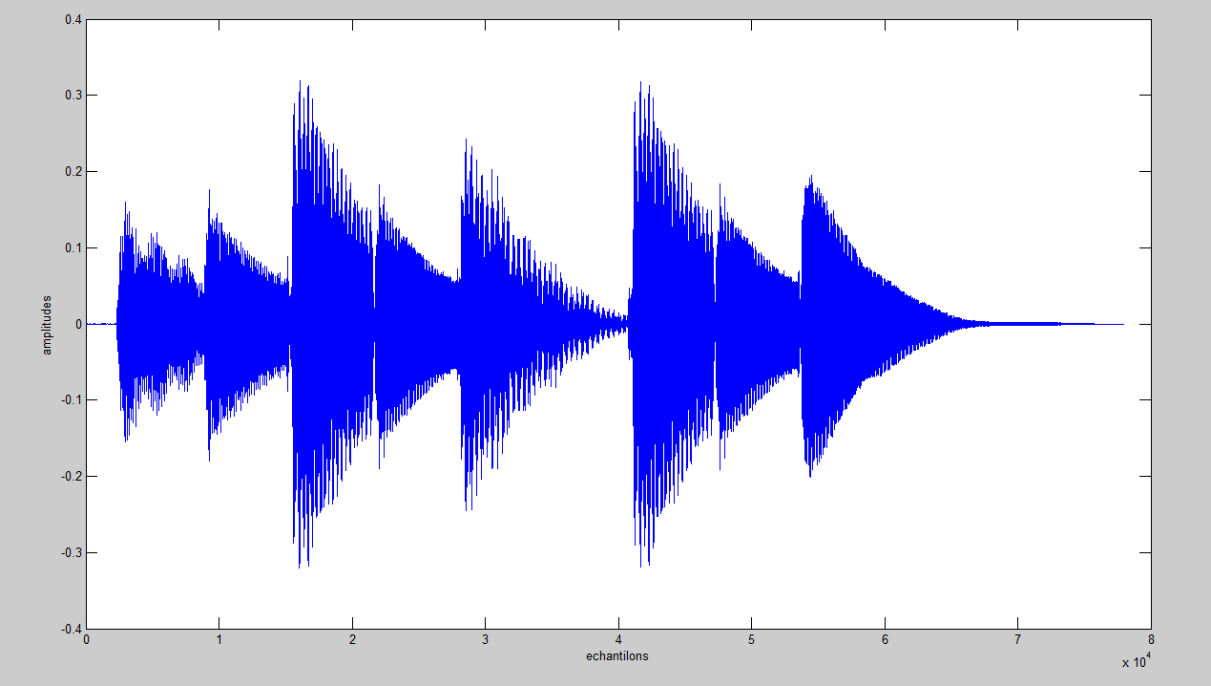


Figure III. 7. *Le Signal de filtre Gammatone filtré de la bande passante N°50.*

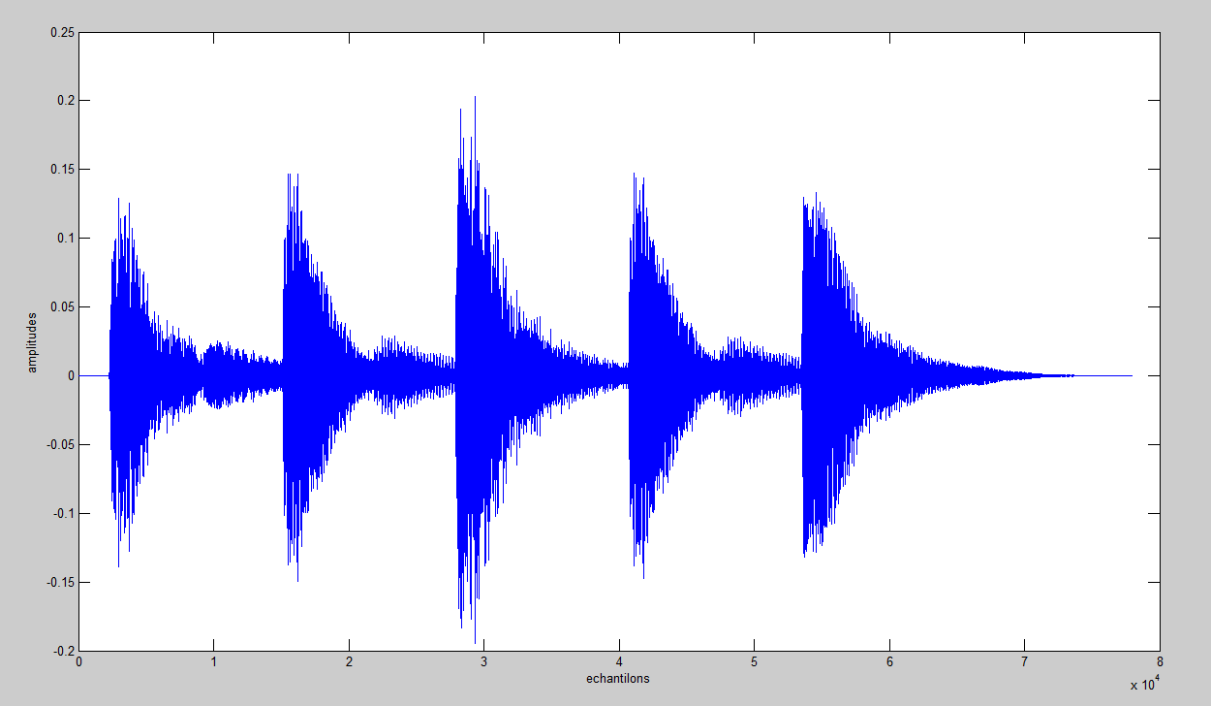


Figure III. 8. *Le Signal de filtre Gammatone filtré de la bande passante N°100.*

III.8. Le cochléagramme

Le cochléagramme représente le comportement de la membrane basilaire dans la cochlée. Cette dernière est la partie interne de l'oreille.

Le cochléagramme est une représentation temps-fréquence du son d'entrée

L'analyse de cochléagramme est un outil utile pour la perception auditive. Selon certains chercheurs, le cerveau humain peut effectuer une opération d'autocorrélation sur les signaux neuronaux et d'analyser les informations d'histogramme pour obtenir les périodes de hauteur fondamentales du son d'entrée.

Le code Matlab permet de :

- Générer un cochléagramme à partir des réponses d'un banc de filtres Gammatone.
- Il donne l'énergie logarithmique des unités TF [35]

Il est décrit dans le code suivant :

```
function a = cochleagram(r, winLength)
winLength = 320; % Longueur de fenêtre par défaut dans les %points
d'échantillonnage qui est de 20 ms pour une fréquence
%d'échantillonnage de 16 KHz
[N_can,sigLong] = size(gf); % Nombre de canaux et longueur du signal
winShift = winLength/2;% décalage dans les trames (par défaut est %demi trame)
increment = winLength/winShift; % increment dans les trames
M = floor(sigLong/winShift); % Nombre de périodes
% Calcule de l'énergie pour chaque trame dans chaque canal
a = zeros(N_can,M);
for m = 1:M
    for i = 1:N_can
        if m < increment
            a(i,m) = gf(i,1:m*winShift)*gf(i,1:m*winShift)';%longueurs de
trame plus courtes pour les trames de début
        else
            startpoint = (m-increment)*winShift;

a(i,m)=gf(i,startpoint+1:startpoint+winLength)*gf(i,startpoint+1:startpoint
+winLength)';
        end
    end
end
```

Résultats

L'énergie d'un filtre Gammatone

Les figures suivantes montrent la variation de l'énergie du signal à la sortie de la 3ème ,50ème et 100ème bandes.

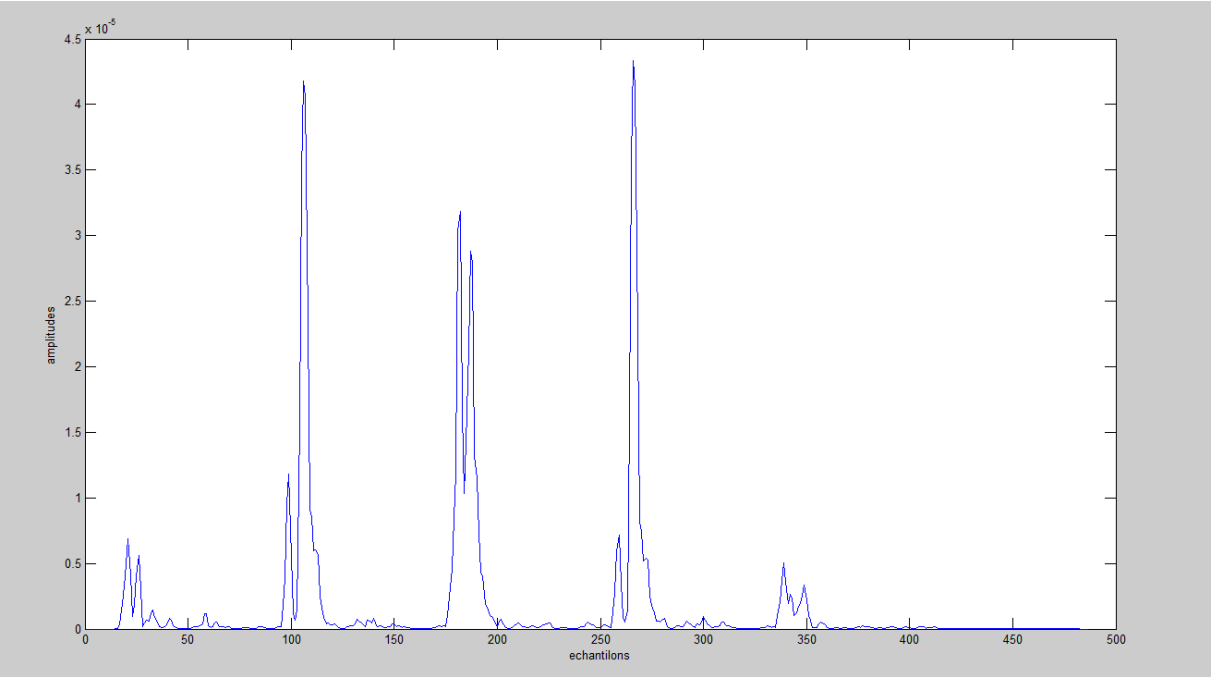


Figure III.9. La sortie de l'énergie d'un filtre Gammatone de la bande passante N°3.

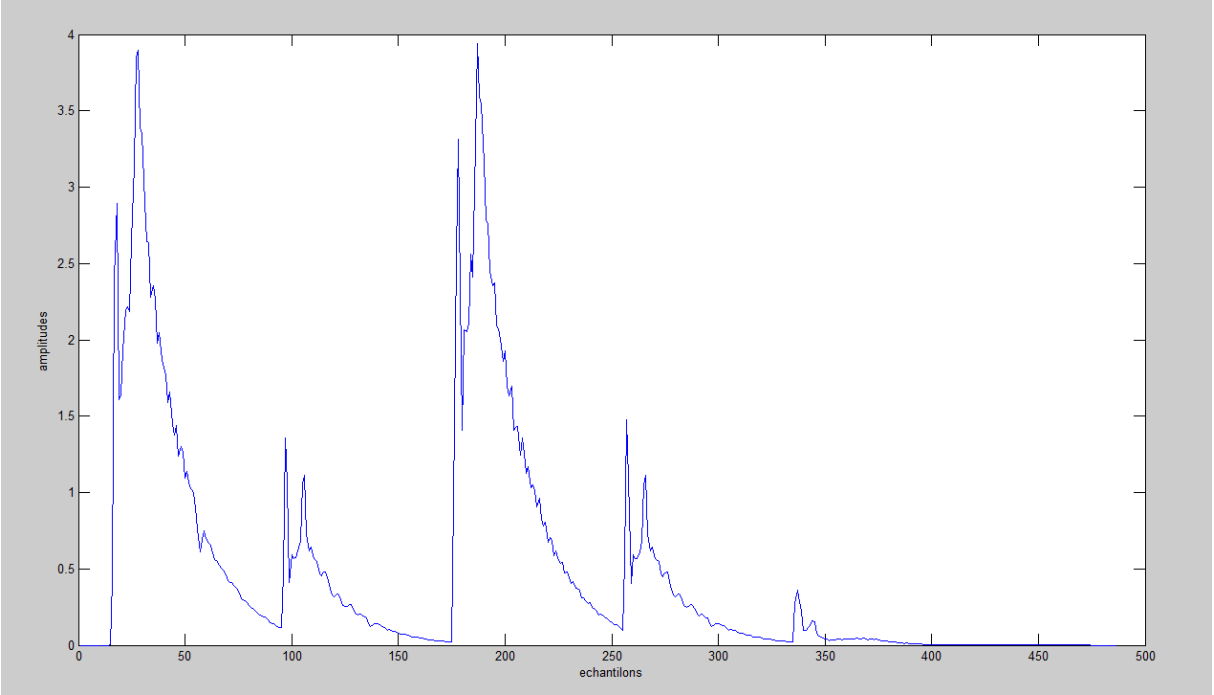


Figure III.10. *La sortie de l'énergie d'un filtre Gammatone de la bande passante N°50*

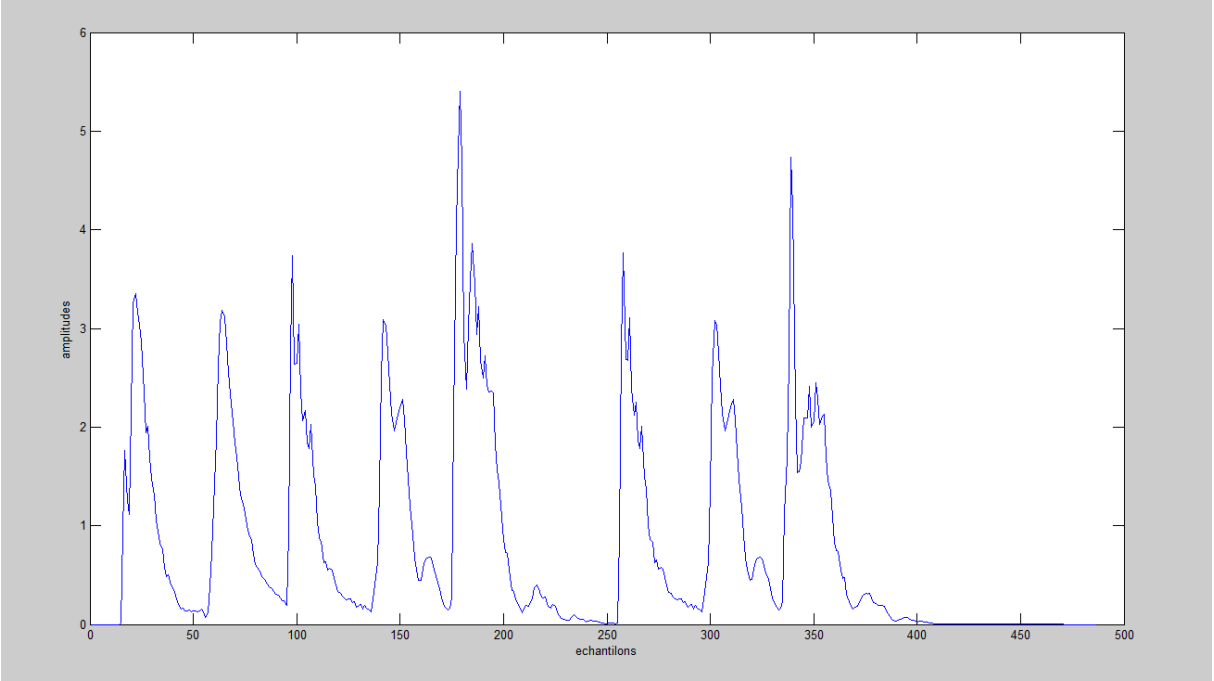


Figure III. 11. *La sortie de l'énergie d'un filtre Gammatone de la bande passante N°100.*

Le cochleagramme d'un filtre Gammatonne

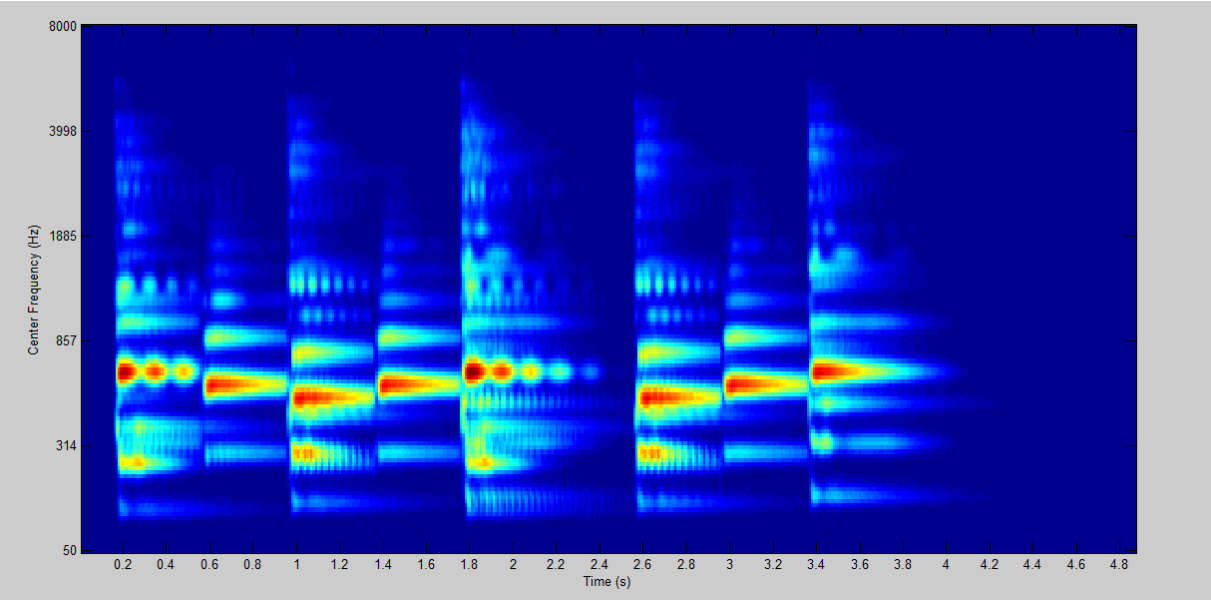


Figure III.12. *Le cochléagramme de piano mélangé au trompette.*

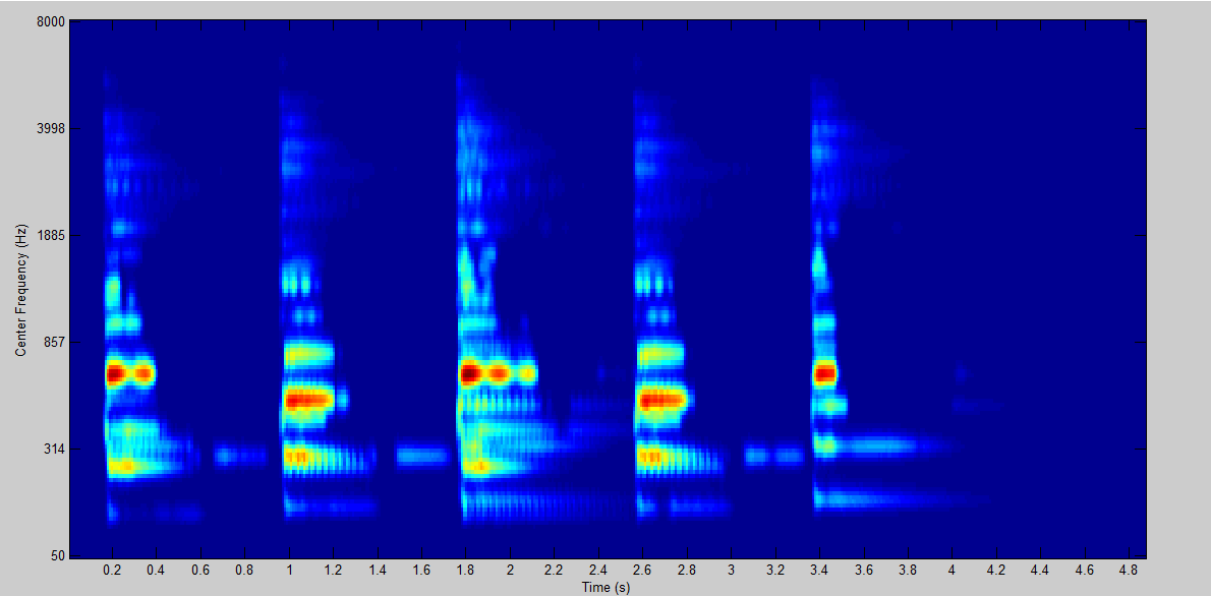


Figure III.13. *Le cochléagramme de source estimé 1.*

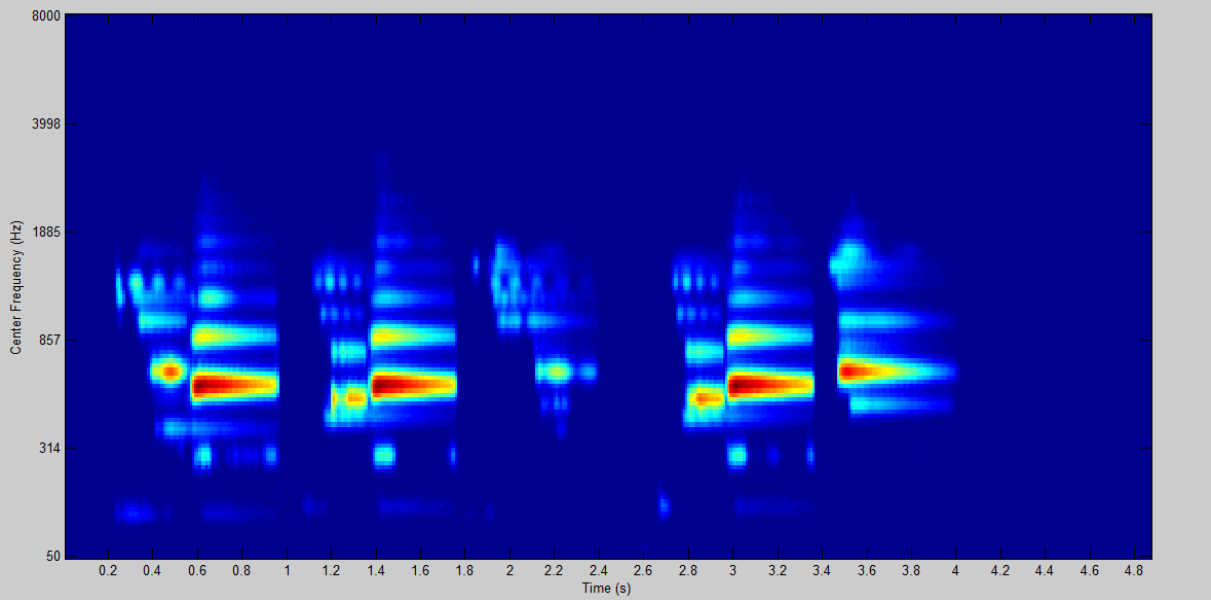


Figure III.14. *Le cochléagramme de source estimé 2.*

Par conséquent, le cochléagramme est développé comme un outil alternatif d'analyse TF pour la séparation des sources afin de surmonter les limitations associées à l'approche par transformée de Fourier.

III.9. Algorithme de séparation de sources

L'algorithme MGD IS-NMF2D est basé sur la descente de gradient multiplicatif.

L'algorithme est décrit comme suit :

- Entrée:$|Y|^2$, matrices non négatives aléatoires D^{τ} et H^{φ},φ,τ.
- Sortie: D^{τ} et H^{φ}
- Procédure:

Calculer la valeur initiale de la fonction de coût 'cost' en utilisant la fonction suivante :

$$\log_p(Y|D,H) \doteq d_{IS}\left(|Y|^2_{f.ts}\left|\sum_{i=1}^{I}\sum_{\tau=0}^{\tau max}\sum_{\varphi=0}^{\varphi max} D^{\tau}_{f-\varphi,i}H^{\varphi}_{i,ts-\tau}\right|\right) \quad \text{(III. 7)}$$

- Itérer pour n=1 : Nombre maximum d'itérations

- Calculer $Z = \sum_{\tau}\sum_{\varphi} D^{\tau}_{f-\varphi}H^{\varphi}_{ts-\tau}$

Mise à jour $D^{\tau}_{f',i'}$ en utilisant (III.7) pour tous les τ, φ.

$$D^{\tau}_{f',i'} \leftarrow D^{\tau}_{f',i'}\frac{\sum_{\varphi,ts}\left(Z_{f'+\varphi,ts}\right)^{-2}|Y|^2_{f'+\varphi,ts}H^{\varphi}_{i',ts-\tau'}}{\sum_{\varphi,ts}\left(Z_{f'+\varphi,ts}\right)^{-1}H^{\varphi}_{i',ts-\tau}} \quad \text{(III. 8)}$$

- Normaliser $D^{\tau}_{f'-i'}$.

- Calculer $Z = \sum_{\tau} \sum_{\varphi} D_{f-\varphi}^{\tau} H_{ts-\tau}^{\varphi}$

- Mise à jour$H_{i'-ts'}^{\varphi'}$ en utilisant (III.8) pour tous les τ, φ.

$$H_{i',ts'}^{\varphi'} \leftarrow H_{i',ts'}^{\varphi'} \frac{\sum_{\varphi,ts} (Z_{f,ts+\tau})^{-2} |Y|_{f,ts'+\tau}^{2} D_{f-\varphi',i'}^{\tau}}{\sum_{\tau,f} (Z_{f,ts'+\tau})^{-1} D_{f-\varphi',i'}^{\tau}} \quad \text{(III. 9)}$$

- Normaliser $H_{i',ts'}^{\varphi'}$

- Calculer la valeur '*cost*' en utilisant la fonction (III.7).

- Fin : Critère d'arrêt $\frac{cost(n-1)-cost(n)}{cost(n)} < \varepsilon$

III.10. Estimation des sources

Les deux matrices dont nous cherchons à séparer à partir de $|Y_{F,ts}|^2$ Sont $|\hat{X}1(f,ts)|^{.2}$ *et* $|\hat{X}2(f,ts)|^{.2}$.

Ces matrices sont estimées en utilisant les deux formules suivantes :

$$|\hat{X}1(f,ts)|^{.2} = \sum_{\tau=0}^{\tau max} \sum_{\varphi=0}^{\varphi max} D_{f-\varphi,1}^{\tau} H_{1,ts-\tau}^{\varphi} \quad \text{(III. 10)}$$

$$|\hat{X}2(f,ts)|^{.2} = \sum_{\tau=0}^{\tau max} \sum_{\varphi=0}^{\varphi max} D_{f-\varphi,2}^{\tau} H_{2,ts-\tau}^{\varphi} \quad \text{(III. 11)}$$

Qui sont ensuite utilisés pour générer le masque binaire :

$$\begin{cases} mask_i(f,ts) = 1 \; si \; |\hat{X}i(f,ts)|^{.2} > |\hat{X}j(f,ts)|^{.2} \\ mask_i(f,ts) = 0 \quad si \; non \end{cases} \quad \text{(III. 12)}$$

Enfin, les sources estimées du domaine temporel sont obtenues comme

$$\hat{X}i = Resynthesize(mask \bullet Y) \text{ pour i = 1,2 où} \hat{x}i = [\hat{x}i(1), \dots, \hat{x}i(T)]^T \quad \text{(III. 13)}$$

Désigne la ième source estimée.

Les sources estimées du domaine temporel sont resynthétisées en pondérant le mélange de cochléagramme par le masque et la correction des changements de phase introduits pendant le filtrage Gammatone. [35]

Le code MATLAB permettant de :

- Synthétiser un signal de forme d'onde à partir d'un signal de mélange et d'un masque binaire.
- Il s'agit d'un filtre inverse du banc de filtres Gammatone.

Les deux sources séparément sont donnés par :

```
for c = 1:N_can
 temp1(c,:) = fliplr(temp1(c,:))/midEarCoeff(c); % Temps de %sortie du filtre
inversé et normaliser les coefficients de l'oreille %moyenne
```

```
temp2 = fftfilt(gt(c,:),temp1(c,:));
temp1(c,:) = fliplr(temp2(1:sigLong))/midEarCoeff(c); % Le %temps revient à
nouveau et normalisation des coefficients de %l'oreille moyenne
weight = zeros(1, sigLong); % Calculer la pondération
for m = 1:numFrame-increment/2+1 % La valeur du masque peut % être binaire
ou rationnelle
startpoint = (m-1)*winShift;
if m <= increment/2
weight(1:startpoint+winLength/2) = weight(1:startpoint+winLength/2)
+ mask(c,m)*coswin(winLength/2-startpoint+1:end);
else
weight(startpoint-winLength/2+1:startpoint+winLength/2) = weight(startpoint-
winLength/2+1:startpoint+winLength/2) +mask(c,m)*coswin;
     end
end
```

Résultats de simulation

Le banc de filtre Gammatone, associé à l'algorithme de synthèse permet d'effectuer des traitements audio. Le son est analysé et décomposé par bande de fréquence, modifié, puis resynthétisé.

La figure suivante nous donne les deux sources séparées du mélange :

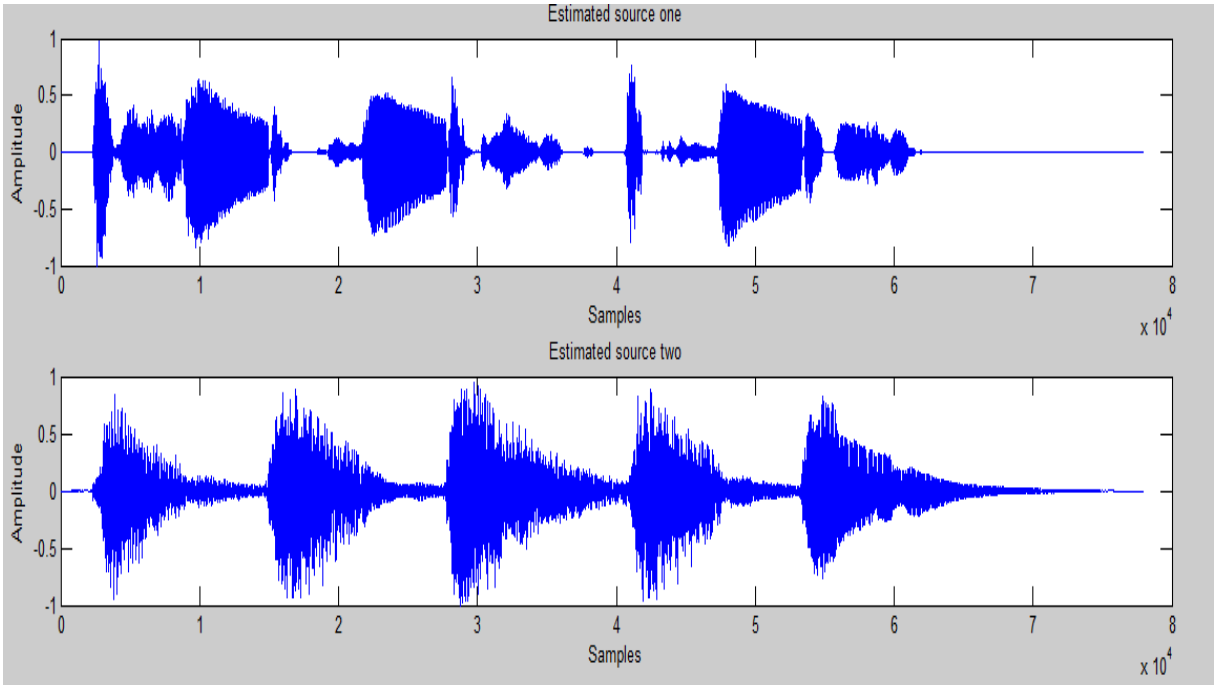


Figure III.15. *Les deux sources estimées (séparées).*

Et finalement pour lire les deux sources mélangées et séparées on utilise :

```
sound(r(1,:),fs);
sound(r(2,:),fs);
```

III.11. Conclusion

Dans ce chapitre, une nouvelle méthode pour résoudre le problème de la séparation des sources audio a été étudiée c'est l'algorithme MGD IS-NMF2D. De plus le système de séparation utilisant le banc de filtres Gammatone avec cette algorithme a montré un rendement considérable d'après les résultats obtenus.

Dans ce mémoire, nous avons abordé la problématique de la séparation aveugle de sources (SAS).

Après une étude rigoureuse du problème de la séparation aveugle de sources, sa définition, son concept et ses diverses applications, plusieurs techniques (qui sont immenses et remarquables) ont été proposées pour apporter une solution à ce problème. Dans notre travail, on a étudié et simulé une nouvelle méthode pour séparer deux sources audio provenant d'un seul canal. En outre, l'algorithme de factorisation bidimensionnelle de matrices non négatives optimisée a été présenté : MGD IS-NMF2D.

Au cours de l'étude, nous avons vu aussi que le filtre Gammatone est très important chez les personnes conçus comme un ajustement simple aux observations expérimentales de la cochlée et la stimulation cochléaires qui sont basées sur le banc de filtre utilisé.

La méthode qu'on a proposée bénéficie d'au moins trois Avantages importants :

- ✓ Tout d'abord, il évite de fortes contraintes de séparation de sources sans connaissances d'information apriori.
- ✓ Deuxièmement, le cochléagramme rendu par le banc de filtres Gammatone a une résolution TF non uniformes, et qui permet de séparer le signal mixte et d'améliorer ainsi l'efficacité de la séparation des sources.
- ✓ Finalement, La méthode détient une propriété souhaitable d'invariance à l'échelle qui permet aux composants à faible énergie dans le cochléagramme d'avoir la même importance relative que les énergies élevées.

Les résultats de notre étude nous permettent de conclure que la simulation dans MATLAB a été réussite ce qui nous laisse de visualiser les signaux sources, les signaux mélanges ainsi que les signaux estimés.

Finalement, nous pouvons dire que notre but de cette réalisation a été achevé en succès et nous proposons d'utiliser le principe de cet algorithme dans la séparation de plus de deux sources, vue les résultats que nous avons obtenus et qui montrent l'efficacité de cette méthode.

Bibliographie

[1] Yasmina BENABDERRAHMANE, `` SEPARATION AVEUGLE DE SIGNAUX DE PAROLE UTILISANT LES STATISTIQUES D`ORDRE SUPRIEUR ET LA DECONPOSISION EN SOUS-ESPACES``, Benabderrahmane_2011.

[2] 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.

[3] 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.

[4] J.F. Cardoso and A. Souloumiac, Blind beamforming for non-Gaussian signals, IEE Proceedings-F, vol. 140, pp. 362–370, December 1993.

[5] D. Yellin and E. Weinstein, Criteria for multichannel signal separation, IEEE Trans. Signal Processing, pp. 2158–2168, August 1994.

[6] C. Simon, Séparation aveugle des sources en mélange convolutif, Ph. D.thesis, l'université de Marnela Vallee,Novembre 1999.

[7] Douglas, S.C. Blind Signal Separation and Blind Deconvolution, CRC Press, New York, 2002.

[8] Vincent, E. Musical source separation using time-frequency source priors. IEEE Transactions on Speech and Audio Processing, 14(1), 91–98, 2005.

[9] 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.

[10] 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.

[11] 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.

[12] Si, H.L., and Zhang, L.M. Intelligent Visual Image Processing, Shanghai Science and Technology Education Press, Shanghai, 2002.

[13] 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.

[14] 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.
[15] Cichocki, A. and Amari, S.Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications, John Wiley & Sons, Ltd, Chichester, 2002.
[16] 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.

[17] Fillon, T., Traitement numérique du signal acoustique pour une aide aux malentendants, 2004, Télécom ParisTech.
[18] Toussaint, P., Etude des filtres auditifs des normo-entendants, des malentendants avec une surdité légère ou surdité moyenne, 2015, Université de Lorraine.
[19] Bastian, L., Etude et comparaison du positionnement microphonique d'un appareil de type contour à écouteur déporté et d'un appareil à microphone externe, 2014, Université de Lorraine.
[20] Site internet : https://www.discret-son.fr/page-type/
[21] C. PARIETTI-WINKLER, Anatomie macroscopique du labyrinthe postérieur, DE d'audioprothésiste 1ère année, Faculté de pharmacie de Nancy, 2013
[22] Site internet : https://www.laboratoires-unisson.com/faq/audition/fonctionnement-systeme-auditif
[23] W. M. Hartmann : Signals, sound and sensation, AIP Press.
[24] B. R. Glasberg & B. C. J. Moore : Frequency selectivity as a function of level and frequency measured with uniformly exciting notched noise, *J. Acoust. Soc. Am.*, 108: p 2318–2328, 2000.
[25] B. R. Glasberg & B. C. J. Moore : Derivation of auditory filter shapes from notchednoise data. *Hearing Research*, 47:p 103–138, August 1990.
[26] Dan Gnansia ;``Modèle auditif en temps réel ``, Mémoire de stage, Rapport de stage du Master ATIAM, Université Pierre et Marie Curie, juillet 2005.
[27] Site internet : http://amajeka.free.fr/analyses/masquage.html
[28] R. D. Patterson, J. Nimmo-Smith, J. Holdsworth, P. Rice: An efficient auditory filterbank based on the Gammatone function, *Paper presented at a meeting of the IOC Speech Group on Auditory Modeling at RSRE*, December 1987.
[29] V. Hohmann : Frequency analysis and synthesis using a Gammatone filterbank, *Acta*

Acustica united with Acustica Vol 88 (2002) 433 – 442.

[30] Richard F. Lyon, Andreas G. Katsiamis, Emmanuel M. Drakakis : History and Future of Auditory Filter Models Department of Bioengineering The Sir Leon Bagrit Centre Imperial College London South Kensington Campus London SW7 2AZ, UK

[31] John Holdsworth ;``Implementing a GammaTone Filter Bank*``, SVOSAnnex C 1988.

[32] COLOMBINI, D., devant l'INSTITUT NATIONAL des SCIENCES APPLIQUEES de LYON, 1999, Université de Savoie.

[33] Site Internet https://blogingenieurs.com/cest-quoi-le-logiciel-matlab/formation/matlab

[34] Faiza Benmoulud, Meriem Euldji " La Séparation Aveugle des Signaux non stationnaire" mémoire de master , Université Yahia Fares Médéa, 2017.

[35] Bin Gao, , W. L. Woo, , S. S. Dlay, "Unsupervised Single-Channel Separation of Nonstationary Signals Using Gammatone Filterbank and Itakura–Saito Nonnegative Matrix Two-Dimensional Factorizations" ,MARCH 2013

[36] Joubaud, T., Prédiction objective de l'effet des systèmes tactiques de communication et protection sur les performances de localisation sonore, 2017, Conservatoire national des arts et metiers-CNAM.