



## **Implémentation d'un Algorithme de Localisation Binaurale sur Audioprothèses : Contraintes et Perspectives**

G. Courtois<sup>a</sup>, P. Marmaroli<sup>a</sup>, H. Lissek<sup>a</sup>, Y. Oesch<sup>b</sup> et W. Balande<sup>b</sup>

<sup>a</sup>Ecole Polytechnique de Lausanne, EPFL STI IEL LEMA ELB 033 - Station 11, CH-1015 Lausanne, Suisse

<sup>b</sup>Phonak Communications AG, Laenggasse 17, 3280 Murten, Suisse  
gilles.courtois@epfl.ch

La capacité à localiser les sources de sons peut être significativement amoindrie pour une personne malentendante. Au quotidien, cela se traduit par une perte de confort acoustique (difficulté à identifier la personne qui parle dans un groupe) voire une augmentation du risque d'accident (incapacité à estimer la direction d'approche d'un danger). A l'heure actuelle, les audioprothèses permettent de compenser partiellement la dégradation de la réponse fréquentielle du système auditif de l'utilisateur. En revanche, elles n'autorisent généralement pas le rétablissement d'indices de localisation tels que la différence de temps et/ou de niveau interaural, ainsi que d'autres indices spectraux. La littérature montre même que les traitements numériques implémentés (filtres, réduction du bruit...) affectent d'avantage ces indices de localisation.

Cet article traite des possibilités d'implémentation d'un algorithme de localisation binaurale (ALB) sur audioprothèses. Celui-ci permettrait de restituer plus précisément les informations spatiales naturellement présentes dans l'environnement sonore de l'utilisateur. Après une revue de l'état de l'art des ALB proposés dans la littérature, nous explicitons plus en détails les contraintes matérielles et logicielles imposées par les prothèses auditives, puis nous proposons une solution applicable pour certains cas particuliers d'appareillage. Enfin, nous engageons une discussion sur la base de nos premiers résultats de mesures en conditions réelles.

## 1 Introduction

Les personnes malentendantes souffrent d'une capacité à localiser les sons pouvant être significativement amoindrie. Cela a été mis en évidence par plusieurs tests de localisation, comme ceux reportés dans [1]. Différents facteurs permettent d'expliquer cette dégradation, tels que le degré de la perte auditive, la symétrie du déficit ou encore le type de perte auditive (conductive ou neurosensorielle). Les performances réduites des personnes malentendantes ne sont pas uniquement dues à l'élévation du seuil d'audition, mais également à d'autres distorsions observées dans le système auditif [2, 3].

Les trois principaux indices permettant au système auditif d'estimer la direction d'un son sont la différence interaurale de temps (ITD en anglais, valable en basses fréquences), la différence interaurale d'intensité (ILD en anglais, valide en moyennes fréquences) et les indices spectraux monauraux (essentiellement utilisés dans les hautes fréquences). Van den Bogaert *et al.* ont montré que les personnes malentendantes utilisent également ces indices [4], et il semblerait qu'aucun autre indice de localisation n'ait été développé par leur système auditif [2]. Lorenzi *et al.* [5] ont étudié la capacité d'auditeurs malentendants à localiser des sources sonores dans le bruit. Ils ont observé que les sujets malentendants voyaient leurs performances se réduire à des niveaux de bruits inférieurs à ceux obtenus avec des sujets normo-entendants, traduisant ainsi une plus forte sensibilité aux bruits ambiants.

Les prothèses auditives actuelles n'apportent pas d'aide significative pour la localisation des sons [6]. Ceci a notamment été rapporté par Köbler et Rosenhall [7], qui ont conclu que l'amplification fournie par les audioprothèses n'améliore pas la capacité de localisation sonore dans le plan horizontal. Les microphones intégrés à la prothèse sont le plus souvent placés derrière l'oreille, ce qui empêche une bonne reproduction des indices spectraux, principalement dus au filtrage du pavillon. De plus, le traitement du signal (réduction du bruit, compression...) réalisé de façon asynchrone et indépendante dans les deux appareils génère des temps de latence et des niveaux de sortie différents d'une oreille à l'autre, ce qui brouille les indices interauraux (ITD et ILD). Ainsi, Van den Bogaert *et al.* [4, 8] ont remarqué que des sujets malentendants présentaient de meilleurs résultats de localisation lorsque leurs audioprothèses étaient éteintes. Cela a été récemment confirmé par Mueller *et al.* [9] qui ont évalué l'impact de différents algorithmes standards de

réduction de bruit dans les prothèses auditives en simulant des environnements sonores réels. Les résultats de leur étude montrent que ces algorithmes tendent à distordre la dimension spatiale du son dans le plan horizontal.

La perception spatiale est une composante importante du système auditif humain qui permet une représentation précise de l'environnement sonore, une immersion dans celui-ci et une meilleure compréhension de la parole dans le bruit [10]. Puisque la plupart des personnes malentendantes sont capables d'extraire les indices spatiaux nécessaires à la localisation sonore, il est pertinent de mieux prendre en compte les informations spatiales présentes dans le signal audio traité par les audioprothèses. Ceci pourrait être réalisé avec l'aide d'un algorithme de localisation binaurale (ALB) capable de déterminer la position de la source sonore principale relative à l'auditeur (par exemple un locuteur s'exprimant avec un malentendant appareillé).

Dans cet article, nous étudions les possibilités d'implémentation d'un tel algorithme sur une paire d'audioprothèses. Dans une première partie, nous nous intéressons aux caractéristiques et au fonctionnement des ALB rapportés dans la littérature. Nous présentons ensuite les contraintes spécifiques imposées par les prothèses auditives, avant de proposer une solution tenant compte de ces spécifications. Enfin, nous exposons nos conclusions.

## 2 Etat de l'art

La littérature des deux dernières décennies propose un nombre important d'algorithmes utilisant le signal audio capté au niveau des oreilles d'un auditeur ou d'une tête artificielle dans le but de déduire la position de la source sonore incidente principale. Ces ALB trouvent leurs applications dans différents domaines que sont notamment la robotique et les technologies de l'audition. La plupart de ces algorithmes reposent sur l'estimation des indices interauraux (ITD et ILD), mais certains ont aussi recours aux indices spectraux monauraux via l'utilisation de fonctions de transfert relatives à la tête (HRTFs en anglais). Les ALB les plus simples font appel à une approche purement calculatoire, d'autres incluent également un modèle statistique.

### 2.1 Approches calculatoires

Lim et Duda [11] ont proposé un ALB qui calcule l'ITD et l'ILD sur 45 bandes de fréquences afin de

modéliser le filtrage de la membrane basilaire. L'ITD est estimé en déterminant le délai qui maximise la fonction d'intercorrélation entre les signaux gauche et droite alors que l'ILD est obtenu en calculant le logarithme du rapport des amplitudes de chaque côté. Ces observations sont ensuite comparées à des valeurs de référence provenant d'une base de données de HRTFs, ce qui permet de trouver la direction d'incidence la plus probable. La déviation absolue moyenne (DAM) obtenue en condition anéchoïque avec une seule source émettant une impulsion est de  $0.8^\circ$  en azimuth.

L'utilisation de HRTFs se retrouve aussi dans la méthode du filtre inverse (HRTF inverse filter) [12] ou de l'algorithme d'annulation de source (source cancellation algorithm) [13]. Dans le premier cas, l'objectif est de déterminer la fonction de transfert inverse à appliquer aux signaux gauche et droite afin de les rendre les plus similaires possible. Dans le second cas, les signaux gauche et droite sont divisés l'un par l'autre, puis l'algorithme calcule la fonction d'intercorrélation entre ce quotient et une série de rapports de HRTFs correspondant à plusieurs directions d'incidence. Pour les deux algorithmes, le couple de HRTFs qui donne la corrélation la plus élevée permet d'estimer la direction d'incidence du son. Des résultats de simulation donnent 96% de réussite pour le premier algorithme, alors que le second a été testé en environnement réel et présente un taux de réussite de 79%.

L'algorithme développé par MacDonald [14] procède au filtrage du signal gauche avec un jeu de HRTF provenant du côté droit, et inversement, jusqu'à ce que les signaux filtrés obtenus soient les plus proches. A nouveau, la paire de HRTFs satisfaisant ce critère donne accès à la direction d'incidence la plus probable. Ce procédé a été évalué avec des sons provenant d'environnements réels pour différents niveaux de bruits ambiants. La DAM est de  $2^\circ$  pour un rapport signal-à-bruit (SNR en anglais) de 10 dB.

Raspaud *et. al* [15] ont recours à une transformée de Fourier à fenêtre glissante pour estimer l'ITD et l'ILD sur des trames successives par bandes de fréquence. Ils utilisent alors un modèle mathématique permettant de calculer des valeurs théoriques de ces mêmes indices en fonction de la direction d'incidence puis de les confronter aux observations. Ce modèle a été ajusté via des facteurs de correction calculés par bandes de fréquence et déterminés à partir de mesures réalisées sur 45 sujets différents. La DAM la plus importante obtenue en simulation est de  $4.5^\circ$  pour une source latérale et de  $2.5^\circ$  pour une source frontale.

## 2.2 Approches statistiques

Des modèles statistiques sont aussi utilisés dans certains algorithmes de localisation. C'est le cas de l'ALB de Li et Levinson [16] qui consiste en l'estimation des indices spatiaux monauraux et interauraux couplé à un système de décision basé sur la règle de Bayes. Le calcul de l'ITD s'effectue à partir d'une fonction d'intercorrélation entre les signaux gauche et droite, alors que l'ILD et les indices spectraux sont déterminés à partir des différences interaurales et intra-aurales sur les spectres obtenus en sortie d'un modèle cochléaire. L'algorithme a été entraîné avec des données expérimentales et présente de très faibles erreurs de localisation en simulation.

Nix et Hohmann [17] ont caractérisé les variations des indices interauraux dans des environnements bruités au

moyen d'une série de densités de probabilité. Ils ont collecté des histogrammes d'ITD et d'ILD pour plusieurs types de bruits (bruit routier, bruit de cafétéria...) et différents SNR, ce qui donne lieu à 27 configurations acoustiques différentes. Ils ont ensuite développé un ALB qui utilise à la fois un modèle de perception auditive et une classification bayésienne. La précision de l'algorithme a été évaluée en termes d'erreur de confusion avant/arrière. Le pourcentage d'erreur passe de 9.5% dans le silence (SNR supérieur à 40 dB) à 41.2% pour des SNR inférieurs à -5 dB.

L'algorithme proposé par Willert *et. al* [18] est basé sur un modèle biologique utilisant une représentation des signaux gauche et droite sous forme de cochléogrammes. Leur système a été entraîné en environnement anéchoïque avec des voix de femmes et d'hommes afin d'obtenir des cochléogrammes de référence pour les indices d'ITD et d'ILD. Cet ALB atteint un taux de réussite de 98.9% en chambre anéchoïque, pour une résolution spatiale de  $15^\circ$ . Ce taux chute à 68.7 % dans un environnement faiblement réverbérant ( $T_{60} \approx 0.43$  s).

## 2.3 Effets du bruit et de la réverbération

Les algorithmes précédemment rapportés ne sont que très rarement évalués dans des conditions acoustiques réelles. Peu d'auteurs ont cherché à contrôler la robustesse de leur ALB vis à vis du bruit et de la réverbération, et à développer ainsi des techniques permettant de se prémunir de leurs effets néfastes sur la précision des résultats.

La démarche la plus approfondie pour contrecarrer les interférences dus au bruit ambiant semble être celle de Nix et Hohmann, mais elle demande une estimation du SNR *a priori*. Huang *et. al* ont quant à eux tenté de réduire la dégradation dus aux réflexions du son en implémentant un algorithme qui modélise l'effet de précedence. Leur objectif est d'exécuter la procédure de localisation seulement dans les trames présentant un fort rapport du son direct sur le son réverbéré. La DAM maximum rapportée est de  $4^\circ$  en chambre anéchoïque, et seulement de  $5^\circ$  en environnement réverbérant, pour une source frontale et une localisation dans le plan horizontal. Vieira et Almeida [19] ont utilisé le même principe en introduisant un détecteur d'attaque pour extraire les trames sur lesquelles la fonction d'intercorrélacion est supposée donner les meilleurs résultats. Cependant, cette méthode demande la donnée de paramètres *a priori* (temps de réverbération, délai avant la première réflexion...), ce qui en constitue la principale limitation.

## 3 Contraintes spécifiques au cas des audioprothèses

Dans cette partie, nous présentons les contraintes imposées par l'utilisation de prothèses auditives comme support de l'algorithme de localisation. Ces contraintes peuvent être classées en trois catégories : contraintes acoustiques, contraintes matérielles et contraintes logicielles. Pour la suite, nous nous focaliserons sur le problème spécifique de la localisation d'un locuteur dans un environnement réel (salle de téléconférence, salle de classe...).

### 3.1 Contraintes acoustiques

Les environnements de fonctionnement considérés ne sont ni anéchoïque, ni silencieux, à la différence de ceux souvent rapportés dans la littérature. Des erreurs de localisation importantes peuvent être observées en cas de réverbération, si l'algorithme détecte la direction d'incidence des réflexions au lieu de celle du son direct. C'est une des raisons qui explique les mauvais résultats de la fonction d'intercorrélacion en environnements réels.

Les bruits ambiants peuvent être classés en deux types de bruit : les bruits stationnaires (générés par exemple par la climatisation ou les vidéoprojecteurs) et les bruits non stationnaires (comme la voix d'un autre locuteur, un bruit de toux, ...). Le premier type de bruit tend à réduire la dynamique de l'ILD mais son effet sur la précision des résultats est limité, tant que le SNR reste positif et que le bruit est diffus. En revanche les bruits non stationnaires entraînent une dégradation de l'estimation de l'ITD et de l'ILD en fonction de leur niveau, de leur durée et de leur spectre.

### 3.2 Contraintes matérielles

Les prothèses auditives disposent de capacités de calcul et de stockage en mémoire particulièrement réduites. Il est de ce fait impossible d'envisager l'utilisation de bases de données conséquentes telles que celles constituées de HRTFs [12, 13, 14] ou de données statistiques importantes [17, 18]. Le coût de calcul de l'algorithme doit également être limité au maximum, et ce pour deux raisons principales. Tout d'abord, un nombre d'opérations successives excessif empêcherait le fonctionnement de la procédure en temps réel. De plus, cela entraînerait une décharge trop rapide des piles des appareils.

Le fait que les microphones de chaque prothèse ne soit pas appariés est un autre point sensible. En effet, des différences significatives peuvent être observées entre les transducteurs gauche et droit en termes de réponse en fréquence et de sensibilité, résultant en une fausse estimation de l'ILD.

### 3.3 Contraintes logicielles

L'algorithme doit gérer le fait que les deux appareils auditifs sont indépendants, disposant chacun de leur propre horloge de fonctionnement. Cela se traduit par des fréquences d'échantillonnage pouvant être légèrement différentes d'un côté à l'autre. Par ailleurs, il n'est pas possible d'échanger des flux audio entre les deux appareils, ce qui s'avère être particulièrement critique pour notre application. Si bon nombre de prothèses auditives dispose aujourd'hui d'un lien sans fil entre les deux appareils, celui-ci engendre une considérable consommation de courant et ne peut donc être utilisé que pour communiquer des données isolées et non l'ensemble des échantillons d'une trame audio. Cela rend de fait impossible le calcul d'une fonction d'intercorrélacion entre les signaux droite et gauche pour estimer l'ITD. En outre, le fréquence d'échantillonnage standard des audioprothèses est en moyenne de 20 kHz (entre 16 kHz et 22.05 kHz pour la plupart), ce qui correspond à une résolution temporelle de 50  $\mu$ s. En se basant sur la formule de Hornbostel [20] (connue sous le nom de "sine law" en anglais), cela revient à une résolution spatiale de

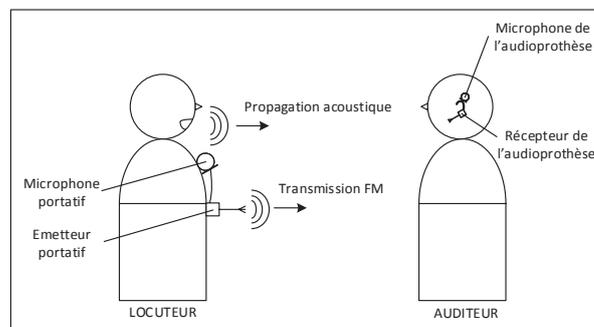


FIGURE 1 – Système d'audioprothèse comportant une transmission FM. En plus de la propagation acoustique captée par le microphone de la prothèse, la voix du locuteur est également capturée par le microphone portatif et transmise au récepteur connecté à l'appareil de l'auditeur.

4.7° dans le plan horizontal<sup>1</sup>. En effet, on ne peut calculer de délai plus petit que la période d'échantillonnage avec les méthodes habituellement utilisées pour estimer un retard temporel.

### 3.4 Discussion

L'implémentation d'un ALB sur une paire d'audioprothèse se révèle être une tâche complexe et délicate. Les méthodes traditionnellement utilisées pour se prémunir des effets du bruit et de la réverbération ne sont pas adaptées aux contraintes matérielles et logicielles imposées par les appareils auditifs actuels. L'estimation des indices de localisation interauraux est également compliquée : l'ILD peut être faussé par les différences existantes entre les microphones de gauche et de droite, et le calcul d'une fonction d'intercorrélacion interaurale est impossible. Enfin, la fréquence d'échantillonnage relativement basse des prothèses auditives nous empêche de disposer d'une haute résolution spatiale.

Dans la partie suivante, nous montrons comment surmonter ces difficultés en se focalisant sur un certain type d'audioprothèses.

## 4 Algorithme de localisation proposé

Certains systèmes d'audioprothèses actuellement commercialisés propose une connexion sans fil supplémentaire, qui permet de transmettre le son de la voix d'un locuteur via une communication FM. Pour cela, un microphone portatif est placé sur les vêtements du locuteur afin de capter le signal de parole optimal (fort SNR et peu de réverbération) et donc hautement intelligible. Ce signal est ensuite modulé et transmis de l'émetteur portatif au récepteur connecté à l'appareil auditif de l'auditeur, comme cela est représenté sur la figure 1. Ainsi, chaque prothèse doit gérer deux signaux : le signal audio démodulé provenant de l'émetteur (noté RX pour la suite) et le signal acoustique capté par le microphone de l'audioprothèse (noté AU).

1. La formule de Hornbostel est calculé est fixant la vitesse du son à 343 m/s et la distance entre les deux microphones modélisant les deux oreilles à 21 cm.

Dans la suite de l'article, nous proposons de profiter de cette fonctionnalité pour effectuer la localisation binaurale. Nous supposons que la source à localiser est le locuteur disposant du microphone portatif et nous montrons comment il est possible d'utiliser ce signal audio supplémentaire pour contourner les problèmes préalablement décrits. Nous faisons également l'hypothèse que le locuteur se situe dans le plan horizontal et dans le champ de vision de l'auditeur.

## 4.1 Algorithme

La transmission FM donne accès à un signal propre et identique sur les deux prothèses auditives. Il est donc possible de l'utiliser comme référence de temps pour estimer l'ITD. Au lieu de calculer une fonction de corrélation interaurale, on peut évaluer le délai entre les signaux RX et AU de chaque côté. Il convient ensuite d'échanger la valeur de ce délai d'un côté à l'autre via la connexion sans fil disponible entre les deux appareils et de soustraire ces deux délais pour remonter à l'ITD. Cela peut être écrit de la façon suivante :

$$\begin{cases} \tau_G = \arg \max_{\tau} R_G(\tau) \\ \tau_D = \arg \max_{\tau} R_D(\tau) \end{cases} \quad (1)$$

où  $\tau_G$  et  $\tau_D$  sont les délais à gauche et à droite qui maximisent la fonction d'intercorrélation entre les signaux RX et AU de chaque côté :

$$\begin{cases} R_G(\tau) = \mathbb{E}[\text{RX}(t)\text{AU}_G(t - \tau)] \\ R_D(\tau) = \mathbb{E}[\text{RX}(t)\text{AU}_D(t - \tau)] \end{cases} \quad (2)$$

avec  $\mathbb{E}$  l'opérateur d'espérance mathématique.

Ensuite, l'ITD se déduit de la différence entre les délais gauche et droite :

$$\Delta_{\tau} = \tau_L - \tau_R. \quad (3)$$

Cette méthode vient résoudre le problème du calcul de la fonction d'intercorrélation interaurale en permettant de n'échanger qu'une seule valeur entre les deux appareils lors de chaque trame. Les effets néfastes du bruit et de la réverbération sont également réduits puisqu'un des deux signaux utilisé est propre. Une fois l'ITD estimé, il faut l'associer à un angle d'incidence dans le plan frontal horizontal. Pour cela, nous utilisons la formule de Hornbostel :

$$\Delta_{\tau} = \frac{a}{c} \sin \theta, \quad (4)$$

où  $a$  est la distance entre les deux microphones modélisant les oreilles,  $c$  est la vitesse du son et  $\theta$  l'angle d'incidence, avec  $\theta \in [-\pi : \pi]$ . Il faut noter que  $a$  doit être fixé à une valeur supérieure au rayon moyen de la tête car il ne tient pas compte de la trajectoire circulaire du son autour de la tête. Il convient ensuite d'isoler l'angle  $\theta$  en fonction de l'observation courante de l'ITD :

$$\theta = \arcsin\left(\Delta_{\tau} \frac{c}{a}\right). \quad (5)$$

La valeur donnée au paramètre  $a$  est de 25.2 cm. Il s'agit de la valeur fournissant la meilleure correspondance au sens des moindres carrés entre le modèle mathématique et les résultats de mesures.

En résumé, la procédure de l'ALB présenté dans cet article est la suivante :

1. Acquisition des trames de 512 échantillons sur les signaux AU et RX de chaque côté (échantillonnage à 16 kHz),
2. Calcul de la fonction d'intercorrélation entre les signaux AU et RX sur les deux appareils, puis déduction de l'ITD,
3. Estimation d'un angle d'incidence à partir de l'observation courante (en utilisant l'équation 5).

## 4.2 Mesures

L'acquisition d'une base de données enregistrée dans différentes configurations acoustiques a été nécessaire à l'évaluation de notre ALB. L'étude s'est portée sur l'influence de différents facteurs (SNR, type de bruit ambiant, réverbération, et effet combiné du bruit et de la réverbération) sur la précision de notre algorithme. Le set-up de mesure et le protocole ont été réalisés de telle façon que ces effets puissent être, autant que possible, évalués de façon indépendante.

Les enregistrements ont été effectués dans une chambre anéchoïque et dans une salle de classe (270 m<sup>3</sup>) de l'EPFL. Le set-up était composé des éléments suivants :

- Un HATS 4128 (Head And Torso Simulator, B&K) utilisé comme locuteur (les stimuli acoustiques étaient émis via sa bouche artificielle),
- Un KEMAR (Knowles Electronic Manikin for Acoustic Research, GRAS) utilisé comme auditeur et monté sur une table tournante. Deux microphones de mesures (PCB Piezotronics ICP 130D20) étaient situés derrière chaque oreille afin de simuler les microphones des prothèses auditives. Un troisième microphone de mesure se trouvait au dessus de sa tête et était utilisé comme capteur de référence pour le contrôle du niveau de pression acoustique.
- Un ensemble de sept haut-parleurs (Genelec 1029A) disposés aux quatre coins de la salle et au milieu des trois murs, aucun haut-parleur n'ayant été placé au niveau du mur derrière le KEMAR, afin d'éviter la présence de sources sonores coïncidentes. Ces haut-parleurs étaient destinés à diffuser du bruit ambiant au niveau désiré.

Le signal de parole utilisé était un extrait de 20 secondes de l'ISTS V1.0<sup>2</sup> qui est constitué d'un mélange de 21 locutrices s'exprimant dans six langues différentes (allemand, anglais américain, arabe, chinois, espagnol, français). Deux types de bruits ambiants ont été utilisés : du bruit blanc et du bruit "cocktail party". La fréquence d'échantillonnage était fixée à 25.6 kHz afin de satisfaire aux besoins de notre appareil de mesure (01dB-Metraviv NetdB), puis réduite à 16 kHz en post-traitement.

Les mesures en chambre anéchoïque ont permis d'étudier l'influence du SNR et du type de bruit sur l'ALB, sans réverbération. L'auditeur (KEMAR) se trouvait à une distance de 2 m du locuteur (HATS) et tourné de -80° à 80° par pas de 10°. Le SNR le plus haut s'élevait à 40

2. International Speech Test Signal, de l'European Hearing Industry Manufacturers Organization, <http://www.ehima.com/>.

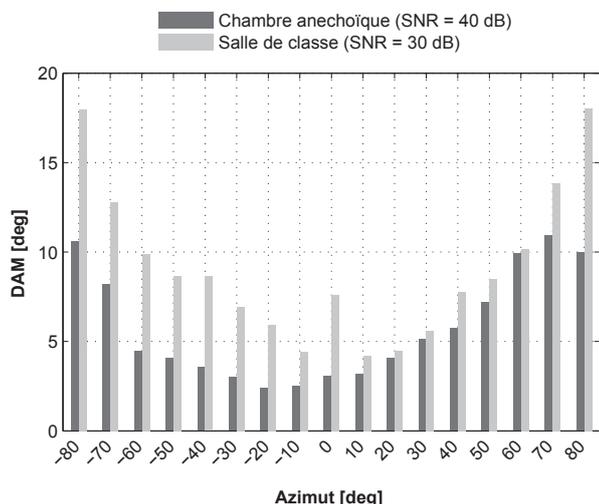


FIGURE 2 – Deviation absolue moyenne en fonction de l'angle d'incidence de la source, obtenue dans le silence en chambre anéchoïque et dans la salle de classe.

dB dans le silence, puis le niveau des haut-parleurs était progressivement augmenté pour atteindre 6 SNRs différents : 25, 18, 12, 6, 0 et -6 dB. Cela a été fait pour les deux types de bruits, résultant en un total de 221 configurations. Les mesures dans la salle de classe ont été utilisées pour évaluer l'impact de la réverbération et de l'effet combiné du bruit et de la réverbération. La distance entre le locuteur et l'auditeur était aussi fixée à 2 m, avec une rotation similaire de la table tournante. Le SNR dans le silence était de 30 dB, puis quatre autres niveaux ont été testés : 12, 6, 0 et -6 dB, résultant en un total de 153 configurations.

### 4.3 Résultats et discussion

Les DAMs obtenues dans le silence en chambre anéchoïque (40 dB de SNR) et dans la salle de classe (30 dB de SNR) sont visibles sur la figure 2, en fonction de l'angle d'incidence. L'algorithme a été appliqué en post-traitement par ordinateur sur les signaux enregistrés. On peut remarquer que notre algorithme présente une bonne corrélation entre les angles d'incidence réels et estimés. A noter une légère asymétrie des résultats dans la chambre anéchoïque, pouvant être attribué à une asymétrie du set-up. Dans la salle de classe, les différents matériaux constituant le mur de droite (mur rigide) et de gauche (fenêtres) peuvent aussi être à l'origine de cette observation. Dans les deux environnements, la DAM est faible quand la source est frontale et augmente à mesure qu'elle devient latérale. Ceci est conforme aux résultats obtenus avec d'autres algorithmes se basant également sur l'ITD (voir par exemple Raspaud *et al.* [15]). L'effet néfaste de la réverbération est aussi clairement visible, ce qui confirme que la localisation basée sur la fonction d'intercorrélacion n'est pas robuste vis à vis des réflexions acoustiques.

La figure 3 montre les DAMs moyennées sur tous les angles d'incidence (de  $-80^\circ$  à  $80^\circ$ ) en fonction du SNR pour les deux types d'environnements et de bruits ambiants. Il apparaît clairement que la réverbération a un effet bien plus prononcé que le bruit sur les résultats de notre algorithme : les courbes représentant la situation en salle de classe traduisent une erreur bien plus importante

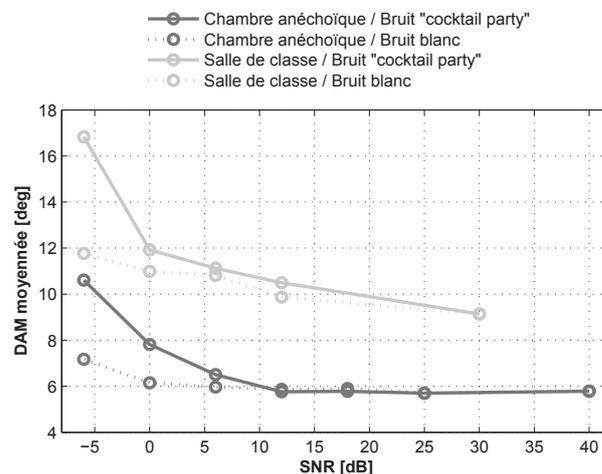


FIGURE 3 – Deviation absolue moyenne moyennée sur tous les angles entre  $-80^\circ$  et  $80^\circ$ , en fonction du SNR. Les résultats dans les deux environnements et pour les deux types de bruits sont représentés.

que celle obtenue en chambre anéchoïque. En outre, l'erreur moyenne observée dans le silence dans la salle de classe est supérieure à l'erreur résultant d'un SNR de 0 dB dans la chambre anéchoïque. Comme attendu, l'algorithme fournit de meilleurs résultats lorsque le bruit ambiant est stationnaire (bruit blanc), ce qui est en accord avec les informations rapportées à la section 3.1. En environnement anéchoïque, les deux types de bruits ambiants ne montrent pas d'effet notable tant que le SNR est supérieur à 10 dB. Néanmoins cet effet apparaît à des SNRs plus élevés en environnement réverbérant et de façon plus importante. Ceci indique que la réverbération tend à accentuer l'influence néfaste du bruit ambiant sur la localisation.

D'une façon générale, nous pouvons affirmer que notre algorithme est robuste vis à vis du bruit ambiant stationnaire pour des SNR positifs. Les résultats sont également satisfaisants en considérant du bruit non-stationnaire, tant que le SNR est au moins de 10 dB. Cependant, notre ALB est très sensible à la réverbération qui affecte la précision des estimations de façon significative. Afin de ne pas lancer la procédure de localisation sur des trames de mauvaises qualité acoustique, il serait possible de s'assurer que le signal AU est suffisamment similaire au signal RX en vérifiant la valeur du maximum de l'intercorrélacion.

## 5 Conclusion

De nombreuses études rapportées dans la littérature ont souligné la capacité d'auditeurs malentendants à localiser le son dans l'espace, malgré d'importantes difficultés dans les environnements sonores complexes. Les prothèses auditives actuellement disponibles sur le marché n'apportent pas d'aide particulière pour la localisation du son et peuvent même réduire les performances des utilisateurs, en raison de la distorsion des indices spatiaux. L'implémentation d'un ALB embarqué sur une paire d'audioprothèses permettrait de mieux prendre en compte le contenu spatial présent dans les signaux traités.

Les algorithmes de localisation précédemment développés ne sont pas compatibles avec les contraintes

spécifiques des prothèses auditives, telles que l'impossibilité d'échanger de larges données audio d'un appareil à l'autre et les faibles capacités de calcul et de stockage en mémoire. C'est pourquoi nous avons proposé une autre stratégie pouvant fonctionner sur les systèmes d'audioprothèses qui incluent une connexion sans fil, permettant de communiquer la voix d'un locuteur via une transmission FM. Nous avons évalué notre algorithme dans des conditions acoustiques anéchoïques et réverbérantes. L'ALB rapporté dans cet article se révèle être robuste vis à vis du bruit ambiant, mais souffre d'une sensibilité excessive à la réverbération. La prochaine étape consiste à optimiser cet algorithme afin de réduire son coût de calcul et de stockage en mémoire.

La version complète [21] de cet article (en anglais) a été soumise à la 136<sup>ème</sup> convention de l'Audio Engineering Society (AES) à Berlin en avril 2014.

## Remerciements

Cette étude est financée par la Commission pour la Technologie et l'Innovation (CTI) du département fédéral de l'économie, de la formation et de la recherche (DEFR), en collaboration avec la société Phonak Communications AG.

## Références

- [1] N. I. Durlach, C. L. Thompson, and H. S. Colburn, "Binaural interaction in impaired listeners : A review of past research," *International Journal of Audiology*, vol. 20, no. 3, p. 181–211, 1981.
- [2] W. Noble, D. Byrne, and B. Lepage, "Effects on sound localization of configuration and type of hearing impairment," *The Journal of the Acoustical Society of America*, vol. 95, p. 992, 1994.
- [3] R. Plomp, "Auditory handicap of hearing impairment and the limited benefit of hearing aids," *The Journal of the Acoustical Society of America*, vol. 63, p. 533, 1978.
- [4] T. Van den Bogaert, T. J. Klasen, M. Moonen, L. Van Deun, and J. Wouters, "Horizontal localization with bilateral hearing aids : without is better than with," *The Journal of the Acoustical Society of America*, vol. 119, p. 515, 2006.
- [5] C. Lorenzi, S. Gatehouse, and C. Lever, "Sound localization in noise in hearing-impaired listeners," *The Journal of the Acoustical Society of America*, vol. 105, p. 3454, 1999.
- [6] W. Noble and D. Byrne, "A comparison of different binaural hearing aid systems for sound localization in the horizontal and vertical planes," *British Journal of Audiology*, vol. 24, no. 5, p. 335–346, 1990.
- [7] S. Köbler and U. Rosenhall, "Horizontal localization and speech intelligibility with bilateral and unilateral hearing aid amplification," *International Journal of Audiology*, vol. 41, no. 7, pp. 395–400, 2002.
- [8] T. Van den Bogaert, E. Carette, and J. Wouters, "Sound localization with and without hearing aids," in *NAG-DAGA International Conference on Acoustics*, pp. 1–4, 2009.
- [9] M. F. Mueller, A. Kegel, S. M. Schimmel, N. Dillier, and M. Hofbauer, "Localization of virtual sound sources with bilateral hearing aids in realistic acoustical scenes," *The Journal of the Acoustical Society of America*, vol. 131, p. 4732, 2012.
- [10] T. Van den Bogaert, E. Carette, and J. Wouters, "Sound source localization using hearing aids with microphones placed behind-the-ear, in-the-canal, and in-the-pinna," *International Journal of Audiology*, vol. 50, no. 3, pp. 164–176, 2011.
- [11] C. Lim and R. O. Duda, "Estimating the azimuth and elevation of a sound source from the output of a cochlear model," in *Signals, Systems and Computers. Conference Record of the Twenty-Eighth Asilomar Conference on*, vol. 1, pp. 399–403, 1994.
- [12] F. Keyrouz, Y. Naous, and K. Diepold, "A new method for binaural 3-d localization based on HRTFs," in *Acoustics, Speech and Signal Processing. IEEE International Conference on*, vol. 5, p. 341–344, 2006.
- [13] F. Keyrouz and K. Diepold, "An enhanced binaural 3D sound localization algorithm," in *Signal Processing and Information Technology, IEEE International Symposium on*, pp. 662–665, 2006.
- [14] J. A. MacDonald, "A localization algorithm based on head-related transfer functions," *The Journal of the Acoustical Society of America*, vol. 123, no. 6, pp. 4290–4296, 2008.
- [15] M. Raspaud, H. Viste, and G. Evangelista, "Binaural source localization by joint estimation of ILD and ITD," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 18, no. 1, pp. 68–77, 2010.
- [16] D. Li and S. E. Levinson, "A bayes-rule based hierarchical system for binaural sound source localization," in *Acoustics, Speech, and Signal Processing. IEEE International Conference on*, vol. 5, p. 521, 2003.
- [17] J. Nix and V. Hohmann, "Sound source localization in real sound fields based on empirical statistics of interaural parameters," *The Journal of the Acoustical Society of America*, vol. 119, no. 1, pp. 463–479, 2006.
- [18] V. Willert, J. Eggert, J. Adamy, R. Stahl, and E. Korner, "A probabilistic model for binaural sound localization," *Systems, Man, and Cybernetics, Part B : Cybernetics, IEEE Transactions on*, vol. 36, no. 5, pp. 982–994, 2006.
- [19] J. Vieira and L. Almeida, "A sound localizer robust to reverberation," in *Audio Engineering Society Convention 115*, 2003.
- [20] E. v. Hornbostel and M. Wertheimer, "Über die wahrnehmung der schallrichtung," *Sitzungsberichte der preussischen Akademie der Wissenschaften*, vol. 388, p. 396, 1920.
- [21] G. Courtois, P. Marmaroli, H. Lissek, Y. Oesch, and W. Balande, "Implementation of a binaural localization algorithm in hearing aids : specifications and achievable solutions," in *Audio Engineering Society Convention 136*, 2014.