



HAL
open science

Localisation d'une source sonore par un réseau de microphones

Adel Thaljaoui, Damien Brulin, Thierry Val, Nejah Nasri

► **To cite this version:**

Adel Thaljaoui, Damien Brulin, Thierry Val, Nejah Nasri. Localisation d'une source sonore par un réseau de microphones. Journées Nationales des Communications Terrestres - JNCT 2014, May 2014, Toulouse, France. pp. 95-109. hal-01146396

HAL Id: hal-01146396

<https://hal.science/hal-01146396>

Submitted on 28 Apr 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Open Archive TOULOUSE Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : <http://oatao.univ-toulouse.fr/>
Eprints ID : 13060

To cite this version : Thaljaoui, Adel and Brulin, Damien and Val, Thierry and Nasri, Nejah *Localisation d'une source sonore par un réseau de microphones*. (2014) In: Journées Nationales des Communications Terrestres - JNCT 2014, 22 May 2014 - 23 May 2014 (Toulouse, France).

Any correspondence concerning this service should be sent to the repository administrator: staff-oatao@listes-diff.inp-toulouse.fr

Localisation d'une source sonore par un réseau de microphones

Adel THALJAOUÏ¹, Damien BRULIN², Thierry VAL³, Nejah NASRI⁴

^{1,3}Université de Toulouse-UT2, Laboratoire CNRS-IRIT équipe IRT

²Université de Toulouse-UT2, Laboratoire LAAS-CNRS équipe N2IS

⁴Ecole Nationale des Ingénieurs de Sfax, Laboratoire LETI équipe ISCSF

¹Adel.Thaljaoui@irit.fr, ²damien.brulin@laas.fr, ³val@irit.fr,
⁴nejah.nasri@iseecs.rnu.tn

Résumé : L'assistance à domicile d'une personne âgée, notamment la connaissance de sa position géographique en tout instant, est devenue actuellement l'une des problématiques les plus urgentes. L'exploitation de l'information audio captée par un réseau de capteurs munis de microphones constitue un axe de recherche prometteur qui pourrait contribuer à une meilleure localisation dans le cadre des maisons intelligentes. Nous introduisons, dans cet article, nos premiers travaux sur la localisation audio en présentant un système de localisation sonore par un ensemble de deux microphones qui se base sur l'estimation de la différence de temps d'arrivée (TDOA). Les résultats montrent qu'un couple de microphones peut localiser une source sonore dans un rayon de 3m et avec une précision de moins de 3 degrés.

Mots clés : Assistance à domicile, réseaux de capteurs sans fil, localisation, audio, microphone, TDOA

1. INTRODUCTION

L'oreille est l'un des organes les plus complexes dans le corps humain. En effet, de part sa nature, l'oreille est composée d'une multitude de cellules spécialisées permettant l'analyse du spectre sonore dans une gamme de fréquences allant de 22 Hz à 22kHz. Cette analyse assure une perception assez précise de l'espace sonore qui nous entoure en permettant non seulement d'analyser les sons qu'elle capte mais aussi de localiser la source émettrice de ces sons, même si elle ne se trouve pas dans notre champ de vision.

Localiser une source sonore revient à déterminer sa direction d'arrivée ainsi que sa distance. Selon une étude réalisée par [1], la précision de cette localisation peut atteindre $\pm 1^\circ$ pour certains mammifères.

La modélisation du fonctionnement de l'oreille a suscité, ces dernières années, un intérêt particulier notamment dans le domaine de la robotique. En effet, plusieurs équipes se sont attachées, durant cette dernière décennie, à produire des robots pouvant assurer les mêmes fonctions élémentaires qu'un être humain peut exécuter à l'instar de la vision (par des caméras), l'ouïe (par des microphones) et même la sensibilité de la peau humaine (par des capteurs) [2] [3] [4] [5] [6]. Mais les applications de la localisation audio dans les réseaux de capteurs sans fils dans le cadre de l'assistance à domicile sont encore à leur balbutiement [7].

L'objectif de nos recherches est de concevoir un système de localisation hybride qui permet de combiner l'information audio à celle des capteurs radio afin de localiser une personne âgée dans un habitat intelligent.

Nos premiers travaux concernent ainsi la localisation d'une source sonore par un ensemble de deux microphones en estimant sa direction d'arrivée.

Dans la littérature, il existe plusieurs méthodes pour la détermination de la direction d'arrivée d'une onde sonore. Les deux les plus utilisées sont l'ILD (*Interaural Level Differences*) et le TDOA (*Time Difference of Arrival*) connue aussi sous le nom d'ITD (*Interaural Time Difference*)(Cf. Figure).

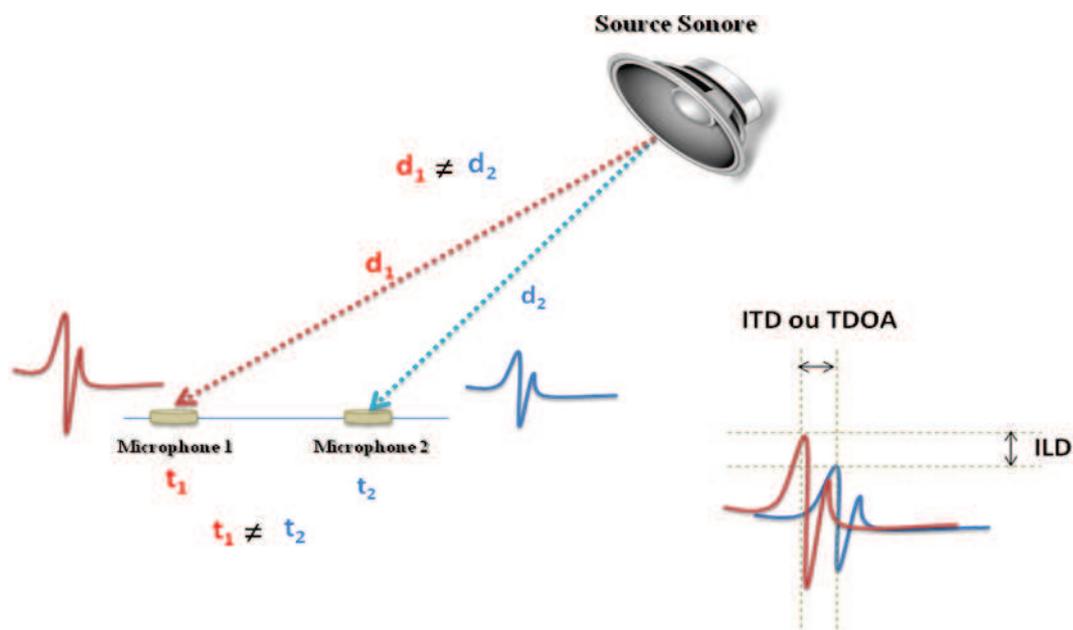


Figure 1: Méthode de calcul de la direction d'une source sonore

2. CHOIX DE LA METHODE DE CALCUL DE LA DIRECTION D'ARRIVEE

2.1. ILD

L'ILD (*Interaural Level Differences*) mesure la différence d'intensité sonore entre les signaux reçus par deux récepteurs (microphones ou oreilles).

En fait, l'intensité d'un signal est généralement considérée comme l'un des indices principaux de perception de l'emplacement d'une source sonore. En effet, l'appareil auditif humain est très sensible à toute variation de cet indice [8].

Selon la loi de l'inverse du carré (*inverse-square law*) [9], l'intensité d'un son dans un champ libre est inversement proportionnelle au carré de la distance entre la source et le récepteur:

$$I = \frac{W}{4\pi d^2} \quad (1)$$

où I est l'intensité sonore en watts/m², W est l'intensité sonore de la source en watts et d est la distance de la source en mètres.

Cependant, cette loi ne s'applique que pour une source relativement proche. En effet, la différence d'intensité devient négligeable en fonction de la distance réduite du récepteur [10]. En outre, la connaissance à priori du niveau de la source sonore à l'émission est indispensable pour sa localisation du fait que le niveau du son à la réception est relatif à celui du son émis.

2.2. TDOA

L'une des techniques les plus utilisées est la technique de la différence du temps d'arrivée (TDOA) qui permet d'obtenir de bons résultats de localisation [6].

Si l'on considère le cas d'un réseau de deux capteurs (microphones) (Figure 2), le principe de l'estimation de la direction d'arrivée d'une onde sonore repose sur le calcul de ses temps de capture par chaque microphone puis en déduire le retard (τ) ou la différence des temps d'arrivée (TDOA) enregistrée entre ces deux microphones.

Ainsi, à l'instar du système auditif humain, l'exploitation du retard que met un son pour arriver à deux capteurs (microphones) pourrait nous conduire à déterminer la direction d'arrivée de ce son.

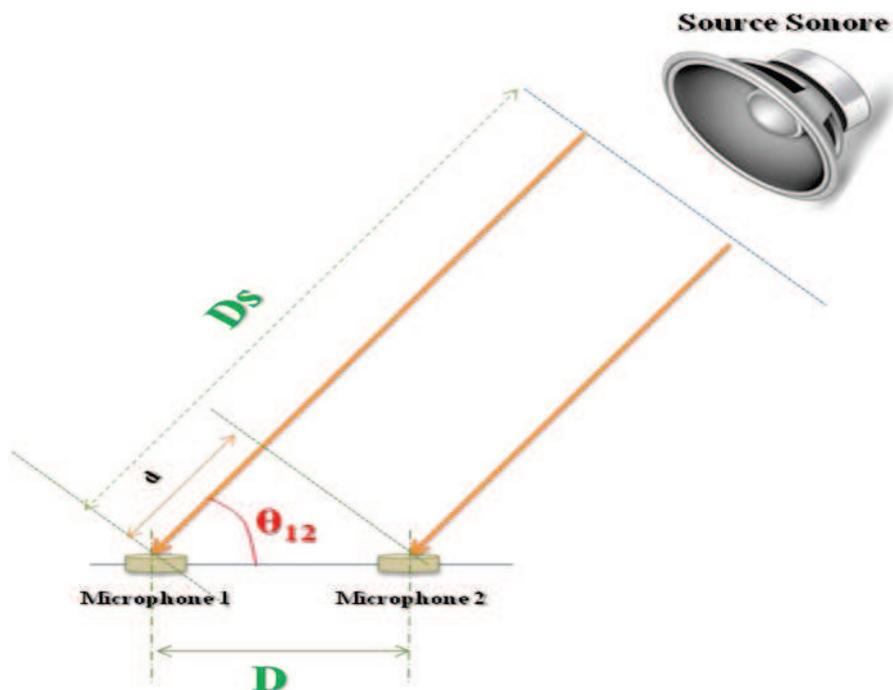


Figure 2: Exemple d'un réseau de deux microphones

Cette méthode est beaucoup plus précise et robuste dans le cas d'une source sonore assez éloignée (cas des ondes planes) ce qui correspond aux objectifs que nous avons fixés. C'est pour cette raison que la méthode TDOA a été retenue pour le reste de nos calculs.

3. METHODE PROPOSEE POUR LE CALCUL DE LA DIRECTION D'ARRIVEE D'UNE ONDE SONORE

Nous considérons un système de deux capteurs et une source sonore placée en dehors du plan perpendiculaire à la droite passant par ces deux capteurs et passant par le milieu de leur entraxe. Le front d'onde de cette source sonore atteindra alors les deux capteurs avec une légère différence de temps ou délai, dépendant de l'angle d'arrivée de la source (Figure 2). La valeur maximale que peut prendre ce délai renvoie à une source sonore se trouvant sur la même droite passant par les deux capteurs (à 0° ou 180°).

Le réseau que nous utilisons est composé de deux microphones pouvant être disposés comme illustré par la Figure .

3.1. PRINCIPE DE LA METHODE

Notre objectif est de déterminer l'angle θ correspondant à la direction d'arrivée de l'onde produite par la source sonore (cf. Figure). Nous supposons que $\theta \in]0, \pi[$.

Nous avons :

- l = distance entre S et Capteur 2
- D_s = distance entre S et Capteur 1
- v : vitesse de propagation du son dans l'air donnée par la formule de Laplace suivante :
 $v = (331 + 0.6 \times T) \text{ ms}^{-1}$; avec T est la température de l'air (supposée proche de 20°C).
- TDOA (différence de temps d'arrivée) notée aussi τ .

Afin de déterminer l'angle θ , nous appliquons les formules de la trigonométrie pour un triangle rectangle de coté d et d'hypoténuse D (entraxe des deux capteurs) :

$$\cos \theta = \frac{d}{D} \quad (2)$$

Or pour une distance d (Fig.2), la vitesse de l'onde sonore est donnée par:

$$v = \frac{d}{\tau} \Rightarrow d = \tau \times v \quad (3)$$

(2) devient alors:

$$\cos \theta = \frac{\tau \times v}{D}$$

D'où: $\theta = \cos^{-1} \frac{\tau \times v}{D} \quad (4)$

La méthode décrite dans cet article permet de calculer l'angle θ selon la méthodologie suivante:

- Acquisition du signal sonore $s(t)$, sa discrétisation selon une fréquence d'échantillonnage F (obtention de $s(n)$ et récupération des signaux correspondant à chacun des deux microphones ($s_1(n)$ et $s_2(n)$))
- Calcul de l'intercorrélacion entre ces deux signaux $s_1(n)$ et $s_2(n)$
- Recherche de la position temporelle du maximum de l'intercorrélacion et détermination du retard τ
- Calcul de la Direction d'Arrivée (angle θ)

3.2. ACQUISITION DU SIGNAL SONORE

Dans cette première étape, nous enregistrons le signal sonore capté par le module à deux microphones que nous avons conçu (Figure). Ce signal est ensuite échantillonné à une fréquence F_e .

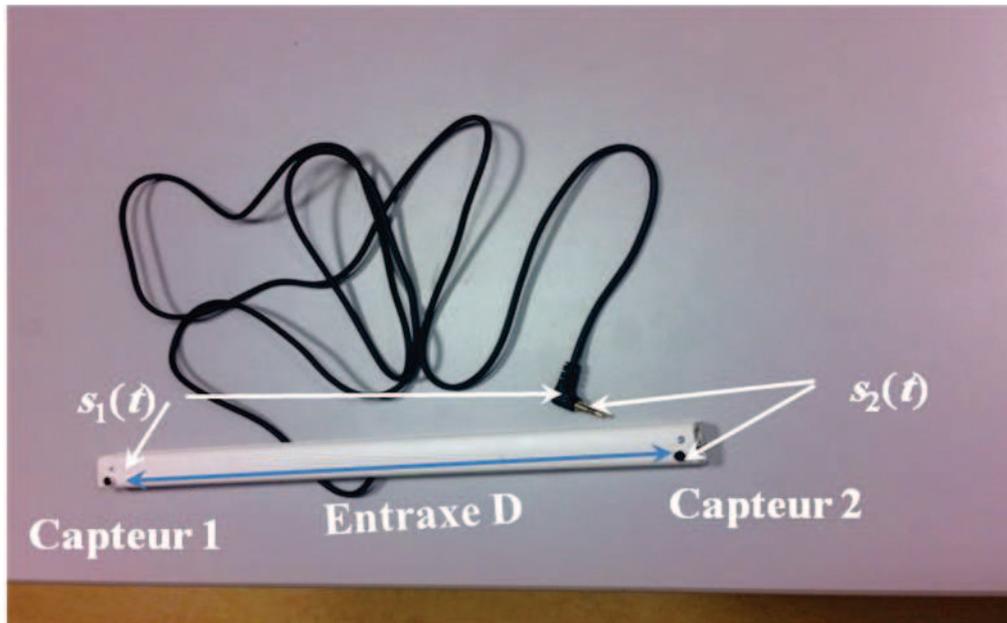


Figure 3: Barre à deux microphones conçue

Le principe de la méthode est donné par la Figure :

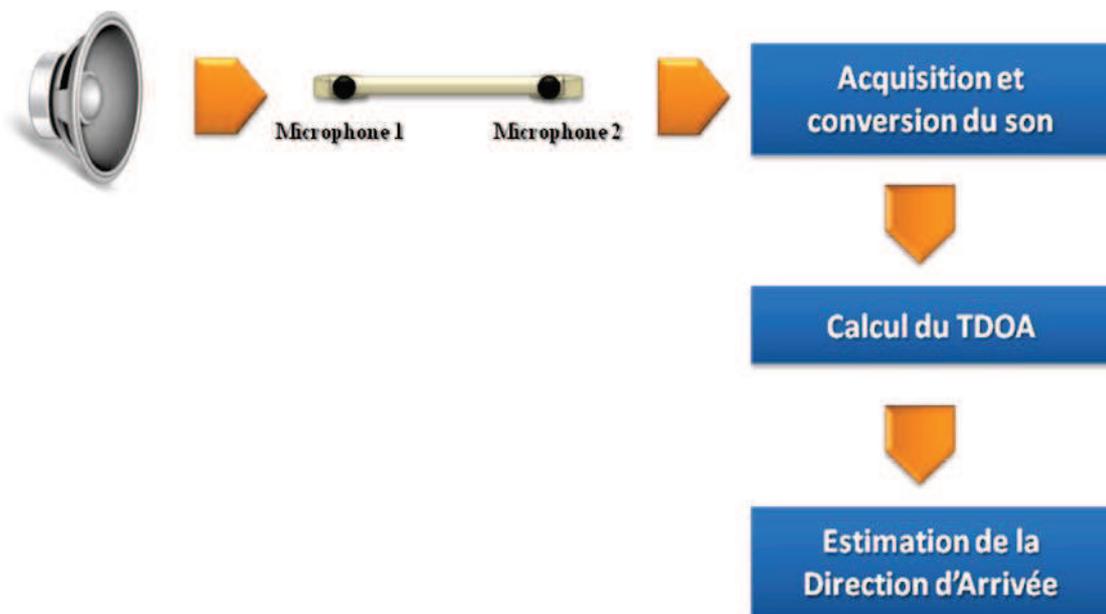


Figure 4: Localisation d'une sonore par un réseau de 2 microphones

Les microphones sont reliés à l'ordinateur par un câble jack stéréo où la voie de gauche et de droite correspondent respectivement aux microphones 1 et 2. Les deux signaux relatifs à ces deux microphones sont ensuite restitués ($s_1(n)$ et $s_2(n)$).

3.3. INTERCORRELATION ET ESTIMATION DU TDOA (τ)

Afin de pouvoir appliquer l'équation (4), il nous faut déterminer le retard τ ou TDOA (Figure). L'une des méthodes les plus simples et les plus utilisées pour la détection du délai temporel entre deux signaux consiste à déterminer le maximum de la fonction d'intercorrélacion entre eux.

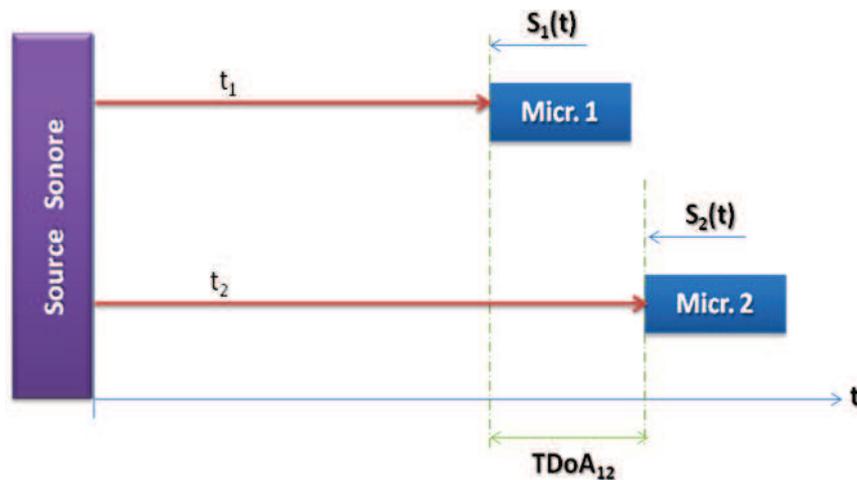


Figure 5: Différence du temps d'arrivé TDoA

La fonction d'intercorrélacion (équation 5) permet de donner une mesure du taux de similitude entre deux signaux. En fait, le signal s_2 reçu par le capteur 2 n'est autre qu'une version décalée dans le temps du signal s_1 reçu par le capteur 1 (selon le signe de τ).

Dans le cas de signaux continus, la fonction d'intercorrélacion est donnée par:

$$C_{s_1s_2}(\tau) = \int_{-T}^{+T} s_1(t)s_2(t - \tau)dt \quad (5)$$

Dans notre cas, les signaux que nous manipulons sont des signaux discrets (échantillonnés). La formule que nous utilisons est la suivante:

$$C_{s_1s_2}(\tau) = \sum_{-T}^{+T} s_1(n)s_2(n - \tau) \quad (6)$$

Il s'agit donc de comparer deux signaux de même longueur l'un décalé de l'autre d'un retard τ . Le principe de cette technique est de balayer le premier signal selon un axe de temps i croissant alors que le second signal balaie l'axe du temps dans le sens contraire ($-i$).

La fonction d'intercorrélacion prend comme paramètres d'entrée deux vecteurs **A** et **B** représentant respectivement les deux signaux audio captés par les deux microphones (dans notre cas, les signaux $s_1(n)$ et $s_2(n)$).

Notons que **longueur (A) = longueur (B)=L**.

Les deux signaux $s_1(n)$ et $s_2(n)$ sont alors "glissés" l'un à travers l'autre dans tous les points pour aboutir enfin à un vecteur **C** de longueur donnée par [11]:

longueur (C) = $2 \times L - 1$.

La valeur maximale que peut prendre le vecteur obtenu **C** représente en fait la position de la similarité maximale entre les deux signaux $s_1(n)$ et $s_2(n)$ moyennant un délai τ .

A titre d'exemple et afin de tester cette méthode, nous avons capté un son « claquement des doigts » (Figure). La figure 7 montre le maximum d'intercorrélacion entre les deux signaux captés par les deux microphones du dispositif expérimental.

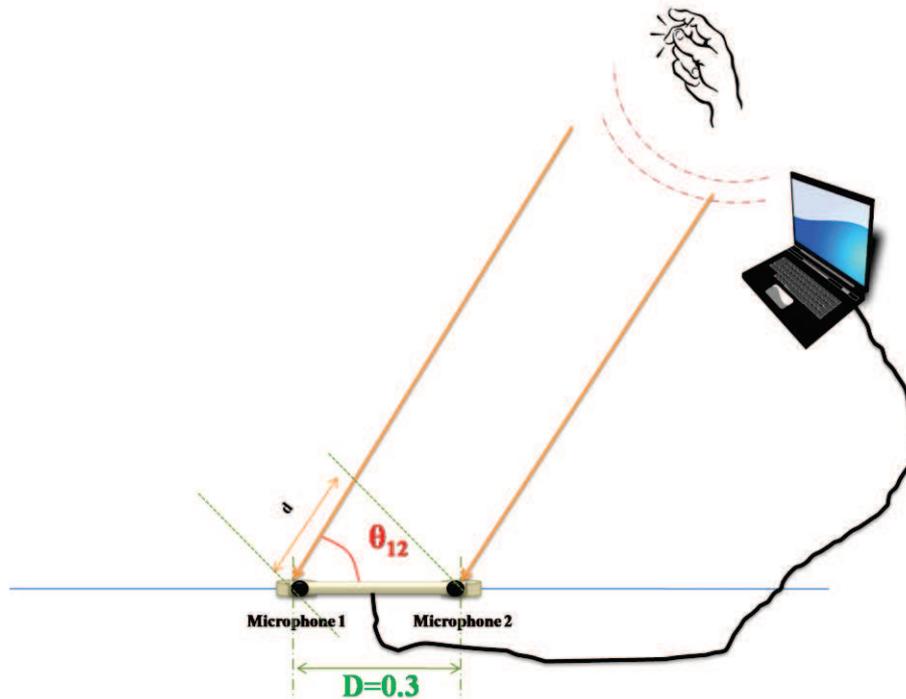


Figure 6: Exemple d'expérimentation de la méthode développée

Une fois que le maximum de l'intercorrélacion entre s_1 et s_2 est déterminé (retard τ entre ces deux signaux), nous pouvons déduire l'endroit de ce maximum en appliquant la fonction d'argument du maximum:

$$\tau = \text{Argmax}(C_{s_1s_2}) \quad (7)$$

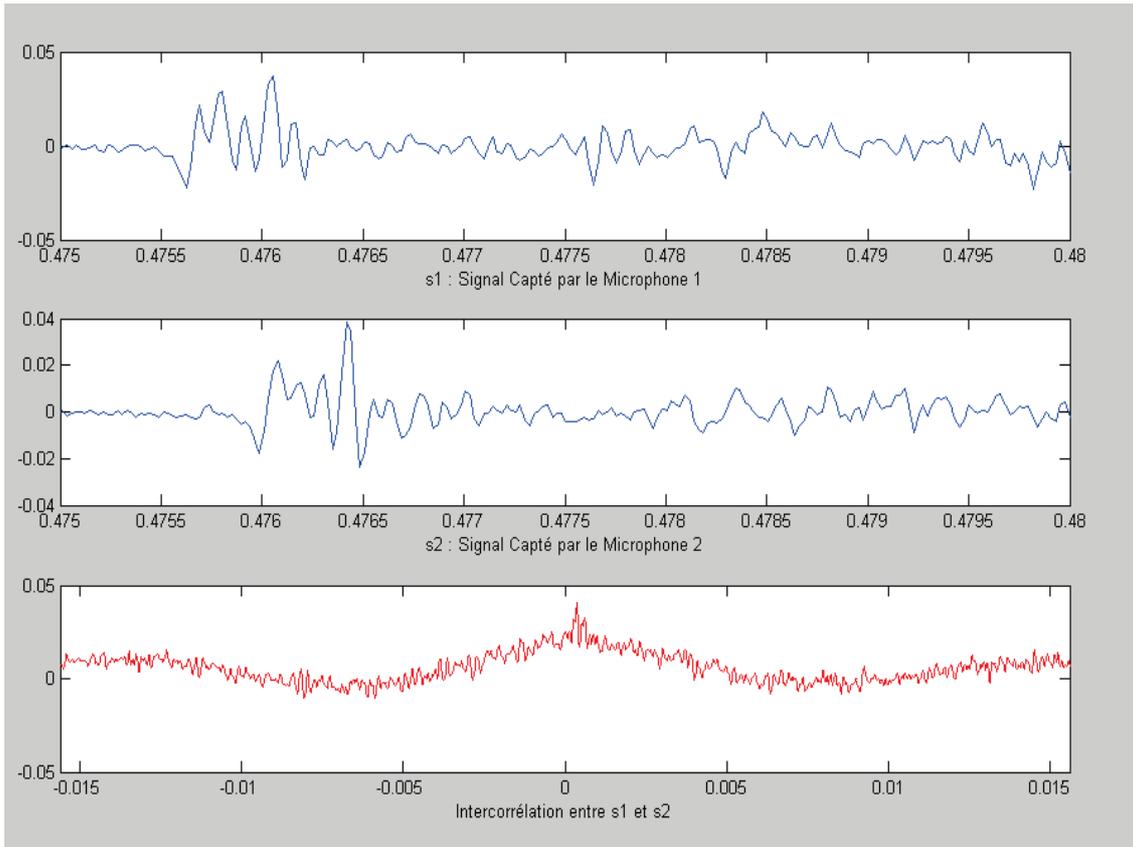


Figure 7: Intercorrélation entre deux signaux captés par les deux microphones du dispositif expérimental

3.4.PERFORMANCES DE LA METHODE PROPOSEE

Les performances de la méthode d'estimation de θ proposée ont été analysées en évaluant l'erreur E commise sur l'estimation de θ .

$$E = |\theta - \hat{\theta}| \quad (8)$$

Cette erreur est calculée en faisant varier la variance du bruit blanc additif gaussien BBAG de 10 dB à -10dB par pas de -1 dB.

Pour que les estimations que nous réalisons soient correctes, nous avons moyenné E sur $N_{est}=200$ estimations de différents θ . Les N_{est} retards sont obtenus à partir de N_{est} différentes valeurs de θ tirées de façon aléatoire dans l'intervalle $]0, \pi[$.

Le rapport signal sur bruit (RSB) est un indicateur sur la qualité de transmission d'une information. Il représente le rapport des puissances entre le signal d'origine et le bruit additif :

$$RSB = 10 \times \log_{10} \left(\frac{P_{signal}}{P_{bruit}} \right) \quad (9)$$

Avec P_{signal} = puissance du signal original et P_{bbg} = puissance du bruit additif

Soit bbg_{init} un bruit blanc gaussien initial de moyenne nulle et de variance 1.

$$bbg_{init} = \text{randn}(1, N_{est}) \quad (10)$$

Nous voulons bruite le signal original avec un bruit blanc gaussien bbg de moyenne nulle et de variance σ^2 tel que : $bbg = \sigma \times bbg_{ir}$

$$P_{bbg} = \sigma^2 \times P_{bbg_{ir}},$$

$$\text{Or } RSB = 10 \times \log_{10} \left(\frac{P_{signal}}{P_{bbg}} \right)$$

$$\text{D'où : } \sigma^2 = \frac{P_{signal}}{P_{bbg_{init}}} \times 10^{-\frac{RSB}{10}}$$

$$\sigma = \sqrt{\frac{P_{signal}}{P_{bbg_{init}}} \times 10^{-\frac{RSB}{10}}} \quad (11)$$

La variation de l'erreur pour différentes valeurs du RSB est donnée par la courbe suivante :

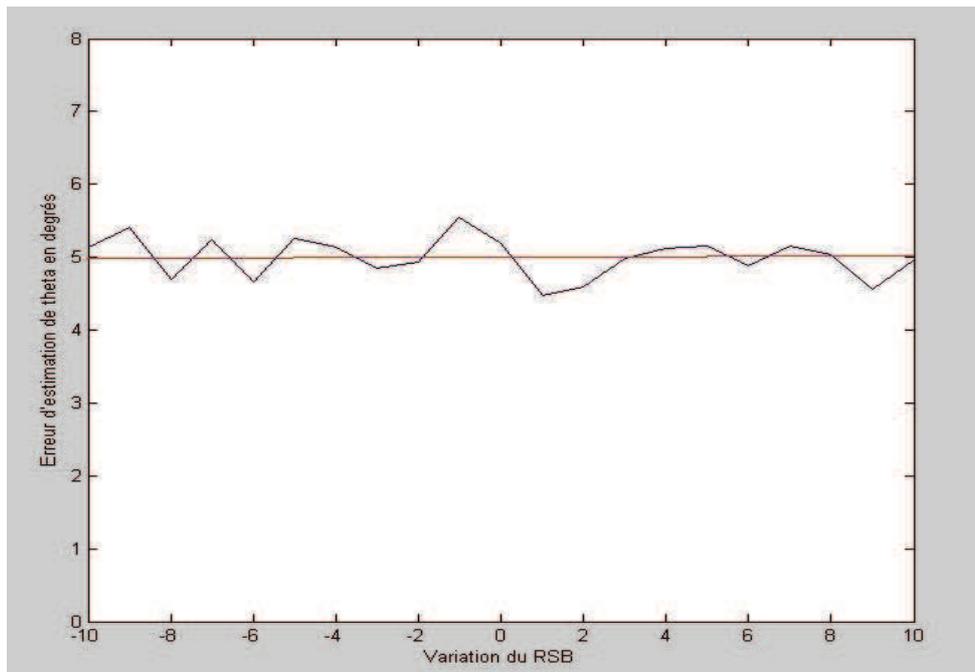


Figure 8: Variation de l'erreur d'estimation de θ en fonction du RSB (en degrés)

Comme nous pouvons le remarquer, l'erreur moyenne théorique d'estimation est de 5° ce qui est considérée comme une bonne estimation pour notre cas.

4. EXPERIMENTATION

Afin de valider la méthode que nous avons proposée, nous avons conçu une barre équipée de deux microphones séparés par une distance $D=30\text{cm}$.

Nous avons ensuite développé une application MATLAB (fig 9) qui pour différentes valeurs de D_s (distance entre la source sonore et le réseau de capteurs) et de directions d'arrivée théoriques, calcule l'angle θ .

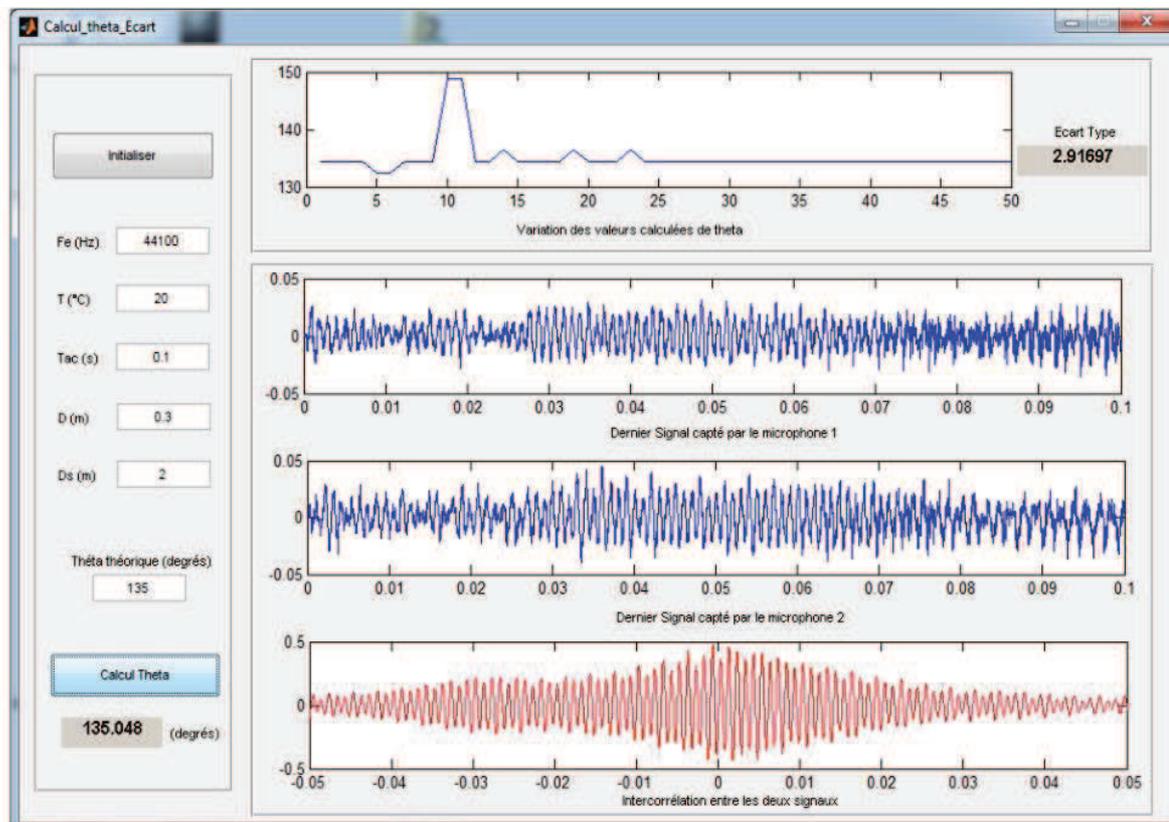


Figure 9: Calcul de la direction d'arrivée d'une source sonore

Pour que les calculs que nous réalisons soient corrects, nous avons moyenné l'angle θ sur 50 itérations (Fig. 10). Pour chacune des 50 estimations de l'angle θ , nous avons calculé l'écart-type (Cf. Tableau 2).

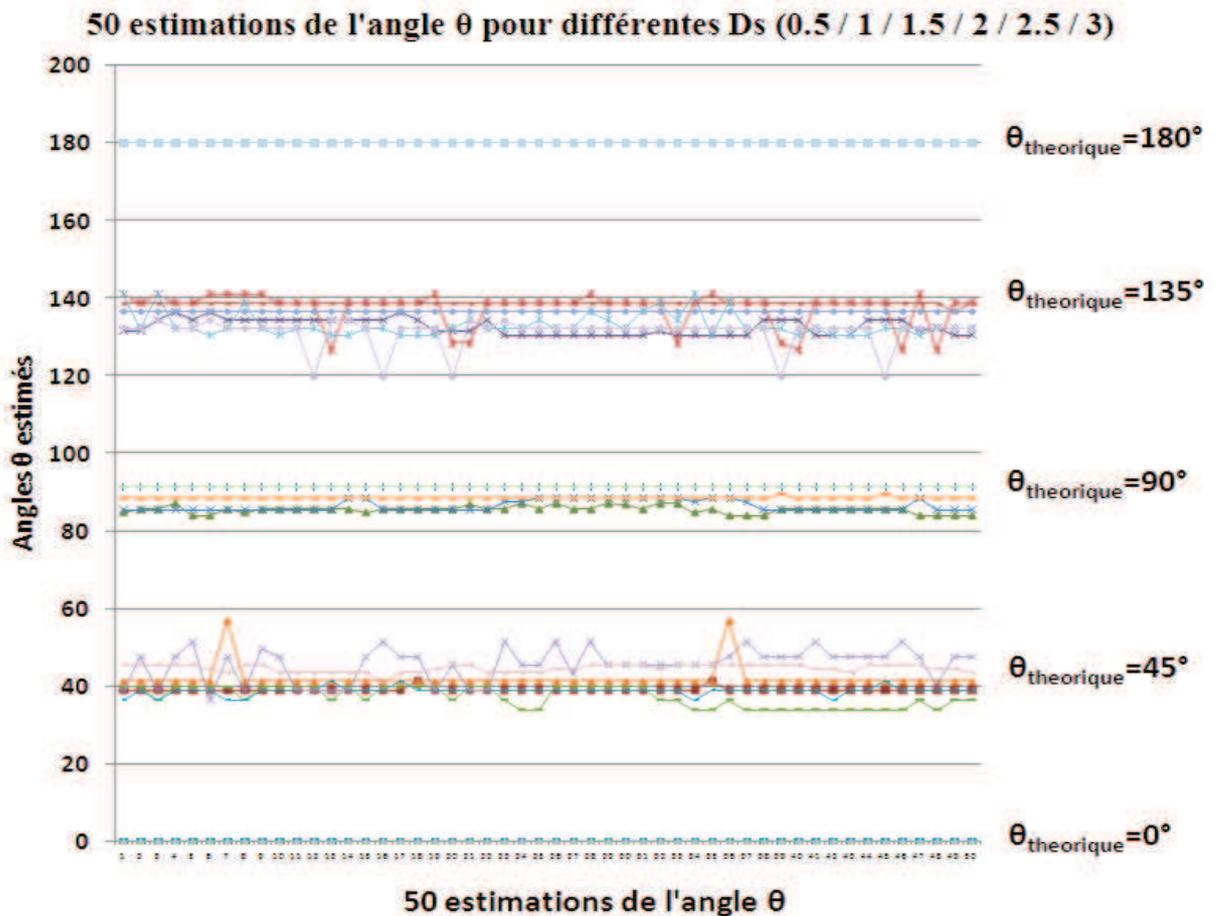


Figure 10: 50 estimation de l'angle θ pour différentes valeurs de D_s

Nous avons choisi $F_e=44100\text{Hz}$. En effet, selon le théorème d'échantillonnage de Nyquist-Shannon [12], l'échantillonnage d'un signal exige une fréquence d'échantillonnage supérieure au double de l'écart entre les fréquences minimale et maximale qu'il contient ($22\text{KHz} - 22\text{Hz}$).

Nous avons fait varier D_s ($0.5 - 1 - 1.5 - 2 - 2.5 - 3$ m). Pour cinq positions théoriques de l'angle θ ($0^\circ - 45^\circ - 90^\circ - 135^\circ - 180^\circ$), nous avons obtenu une erreur moyenne d'estimation de la direction d'arrivée de 3° , ce qui est considéré comme bon.

Les résultats obtenus sont donnés dans le tableau suivant (pour $f_c = 44100\text{Hz}$ et $D=0.3\text{m}$) :

Ds	Angle théo.	Angle estimé	Erreur (degrés)	Ecart type des 50 estimations de l'angle θ
0.5	0	0	0	0
	45	39,03	5,97	4,57E-01
	90	85,40	4,60	9,34E-01
	135	132,59	2,41	2,08E+00
	180	180	0	0
1	0	0	0	0
	90	91,49	1,49	4,31E-14
	135	138,71	3,71	3,12E-01
	45	37,32	7,68	2,62E+00
	180	180	0	0
1.5	0	0	0	0
	45	41,88	3,12	3,12E+00
	90	86,52	3,48	1,37E+00
	135	137,42	2,42	4,46E+00
	180	180	0	0
2	0	0	0	0
	45	38,79	6,21	1,01E+00
	90	88,55	1,45	1,98E-01
	135	136,55	1,55	1,72E-13
	180	180	0	0
2.5	0	0	0	0
	45	45,68	0,68	4,34E+00
	135	133,10	1,90	2,94E+00
	180	180	0	0
3	0	0	0	0
	45	44,66	0,34	1,06E+00
	90	91,49	1,49	4,31E-14
	135	131,34	3,66	4,04E+00
	180	180	0	0

Tableau 2: Résultats des expérimentations

5. CONCLUSION ET PERSPECTIVES

Dans ce papier, un système de localisation d'une source sonore par un réseau de deux capteurs a été introduit. Ce système fait appel à la méthode du calcul du TDOA des signaux sonores reçus par chaque paire de microphones. Un premier travail a été réalisé sur une expérimentation à partir d'un module groupant 2 microphones que nous avons réalisée, associée à un algorithme de traitement développé sous Matlab. Les résultats obtenus sont jugés satisfaisants.

Un second travail est en cours de préparation. La localisation de la source sonore sera réalisée par un réseau de deux couples de microphones. L'objectif à atteindre est de pouvoir estimer non seulement la direction d'arrivée de la source sonore mais aussi sa position géographique (cf figure 11).

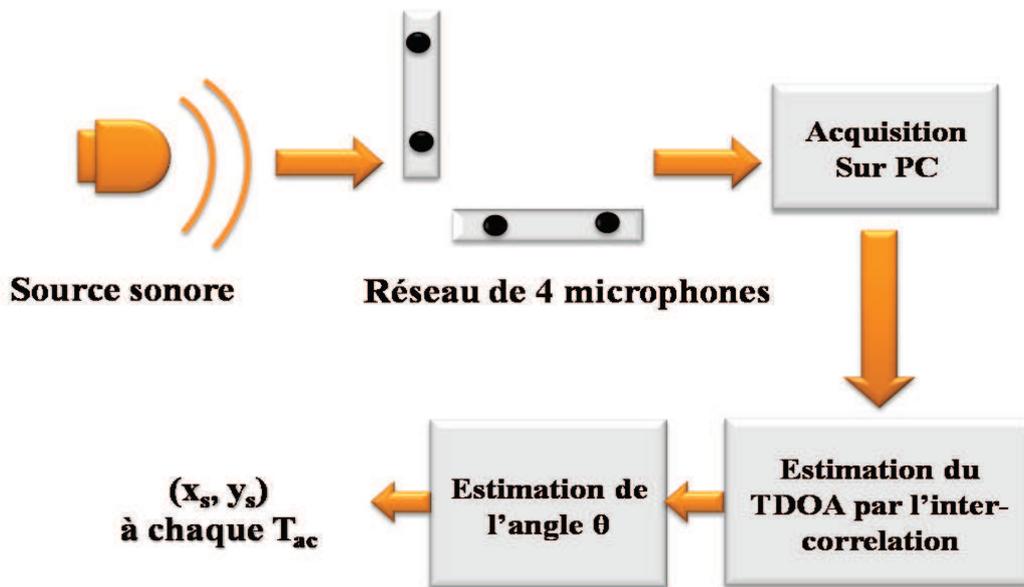


Figure 11: Localisation d'une sonore par un réseau de 4 microphones

Un autre axe de recherche est en cours d'étude : l'ajout d'un module qui assure le débruitage de la source sonore après son acquisition. En effet, les signaux sonores que nous traitons sont des signaux bruités, ce qui peut affecter la précision de notre méthode. Outre cela, dans notre étude, nous avons supposé que les signaux n'ont pas subi d'effet Doppler ce qui n'est pas toujours le cas. Enfin, nous envisageons de travailler de façon conjointe sur une localisation radio, afin de proposer un système de localisation hybride radio et audio tirant avantage des deux systèmes en fonction de la position de la personne.

6. REFERENCES

- [1] HAROLD et al., Harold L. Hawkins, Teresa A McMullen, Arthur N. Popper, Richard R. Fay. Auditory Computation. Springer Handbook of Auditory Research. Springer 1996, pages 334-336., 1996.
- [2] VALIN et al., Valin, J.-M.; Michaud, F.; Hadjou, B.; Rouat, J "Localization of Simultaneous Moving Sound Sources for Mobile Robot using a Frequency-Domain Steered Beamformer Approach, ICRA '04. 2004 IEEE International Conference, Volume: 1., 2004.
- [3] YEOUN et al., Ji-Yeoun Lee; Su-young Chi; Jae-Yeun Lee; Minsoo Hahn;

- Young-Jo Cho, "Real-time sound localization using time difference for human robot interaction", World Congress, Volume # 16 | Part# 1., 2005.
- [4] HÖRNTEIN et al., Jonas Hörnstein, Manuel Lopes, José Santos-Victor, "Sound Localization for Humanoid Robots -Building Audio-Motor Maps based on the HRTF", International Conference on Intelligent Robots and Systems, China., 2006.
- [5] LIU et al., H. Liu, M. Shen Continuous Sound Source Localization Based on Microphone Array for Mobile Robots. IEEE International Conference on IROS, Taipei, Taiwan, 2010 :4332-4339., 2010.
- [6] LI et al., Xiaofei Li, Miao Shen, Wenmin Wang and Hong Liu, "Real-time Sound Source Localization for a Mobile Robot Based on the Guided Spectral-Temporal Position Method", International Journal of Advanced Robotic Systems., 2012.
- [7] VACHER et al., M. VACHER, F. PORTET, A. FLEURY et N. NOURY : Challenges in the processing of audio channels for ambient assisted living. In 12th International Conference on E-Health Networking, Applications and Services, Lyon, France, 1-3 Jul. 2010., 2010.
- [8] Coleman, Coleman P. An analysis of cues to auditory depth perception in free space. Psychol. Bull. 60:302–315., 1963.
- [9] BALLOU et al., G. Ballou, Handbook for Sound Engineers: The New Audio Cyclopedia (Howard Sams, Indianapolis, IN), pp. 14-15., 1987.
- [10] HELLER, François HELLER, "Écoute de contrôle au casque en prise de son multicanal", Mémoire de fin d'études, École Nationale Supérieure Louis Lumière., 2013.
- [11] GU et al., Hung-Yan Gu, Shan-Siang Yang, A sound-source localization system using three-microphone array and crosspower spectrum phase, Machine Learning and Cybernetics (ICMLC), 2012 International Conference on (Volume:5)., 2012.
- [12] BERCHER, Jean-François Bercher, TF, Dirac, convolution, et tutti quanti, École Supérieure d'Ingénieurs en Électrotechnique et Électronique., 2001.