



**HAL**  
open science

# Techniques de spatialisation binaurale pour le guidage de sportifs non-voyants

Sylvain Ferrand

► **To cite this version:**

Sylvain Ferrand. Techniques de spatialisation binaurale pour le guidage de sportifs non-voyants. Traitement des images [eess.IV]. Institut Polytechnique de Paris, 2021. Français. NNT: 2021IP-PAX026 . tel-03297189

**HAL Id: tel-03297189**

**<https://theses.hal.science/tel-03297189v1>**

Submitted on 23 Jul 2021

**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.



INSTITUT  
POLYTECHNIQUE  
DE PARIS

NNT : 2021IPPAX026

Thèse de doctorat



# Techniques de spatialisation binaurale pour le guidage de sportifs non-voyants

Thèse de doctorat de l'Institut Polytechnique de Paris  
préparée à École polytechnique

École doctorale n°626 École doctorale de l'Institut Polytechnique de Paris (EDIPP)  
Spécialité de doctorat : Signal, Image, Automatique et Robotique

Thèse présentée et soutenue à Palaiseau, le 15 juin 2021, par

**SYLVAIN FERRAND**

Renaud Séguier Professeur, Centrale Supélec	Président
René Farcy Professeur, Université Paris-Saclay	Rapporteur
Véronique Zimpfer Chargée de recherche, Institut Saint Louis	Rapportrice
Guillaume Andéol Médecin en chef, Institut de Recherche Biomédicale des Armées	Examineur
Thierry Bodineau Directeur de recherche, École Polytechnique (CMAP)	Examineur
Caroline Cohen Maître de conférences, École Polytechnique (LadHyX)	Examinatrice
François Alouges Professeur, École Polytechnique (CMAP)	directeur de thèse
Matthieu Aussal Ingénieur de Recherche, École Polytechnique (CMAP)	co-directeur de thèse

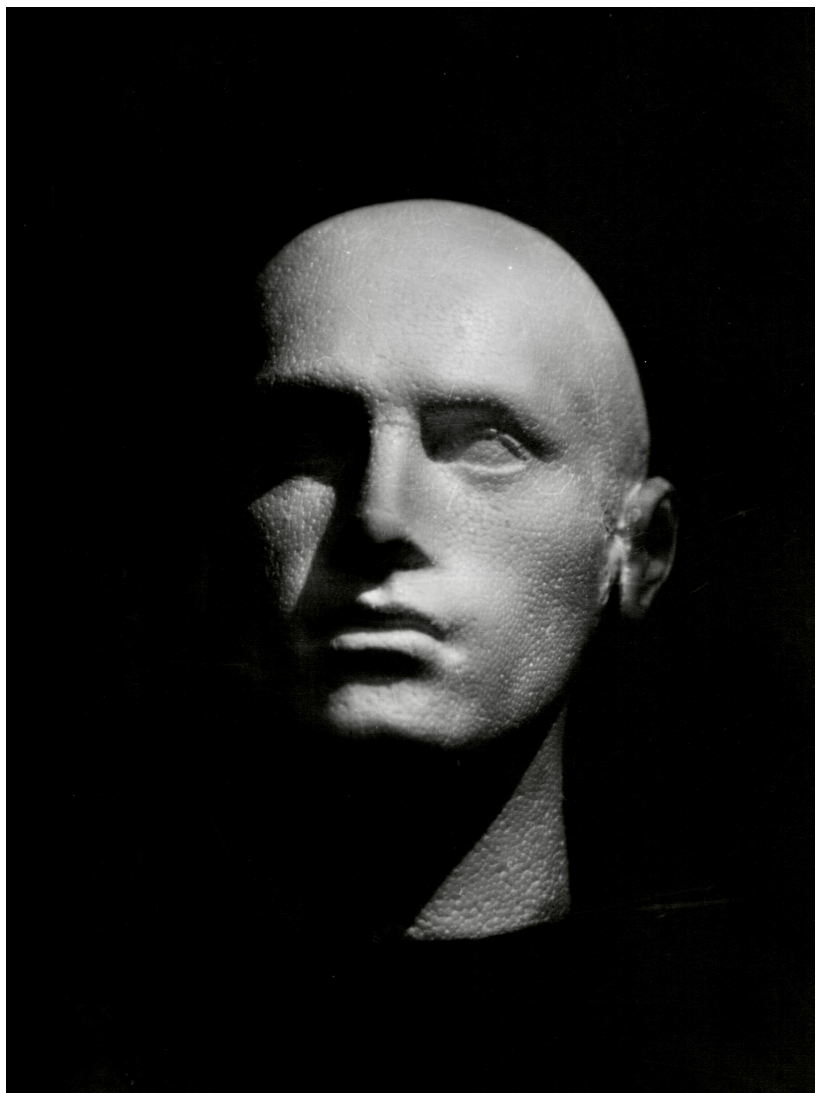








À ma femme Barbara,  
à nos enfants Marianna, David, Natalia et Raphaël.



Tête de captation binaurale artisanale.  
Chambre 13x18, film radiographique



# Remerciements

En premier lieu je tiens à remercier mes directeurs de thèse sans qui ce travail n'aurait jamais pu voir le jour. François, tu m'as fait ce cadeau incroyable et inespéré en acceptant d'encadrer cette thèse. Alors que j'étais ingénieur, sans expérience de recherche, tu m'as fait confiance, dès le début du projet et tout au long de ces cinq années. Matthieu, tu m'as permis d'avancer rapidement en partageant généreusement ton expertise, ton savoir faire, tes codes et travaux. Messieurs, merci pour votre disponibilité et votre bienveillance constante. Alors que le plaisir d'apprendre et d'explorer de nouveaux sujets était ma principale motivation, vous avez su me laisser toute la liberté pour explorer des thématiques variées et parfois très éloignées de vos domaines habituels et de vos centres d'intérêt. J'ai pris un réel plaisir à travailler avec vous !

Merci aux rapporteurs de cette thèse, Véronique Zimpfer et René Farcy, pour la relecture attentive de ce manuscrit et pour les appréciations et remarques toujours constructives qu'ils ont formulées. Elles me pousseront certainement vers des recherches complémentaires. Je remercie l'ensemble des membres du jury, qui m'ont fait l'honneur de bien vouloir étudier avec attention mon travail. Caroline Cohen pour son expertise Sport & Sciences avec qui j'avais eu le plaisir d'échanger autour du Cécifoot ; Guillaume Andéol dont l'aide a été précieuse pour monter des expériences de localisation sonore ; Renaud Séguier pour son expertise en signal et enfin Thierry Bodineau qui m'a grandement soutenu tout au long ce travail de thèse en tant que directeur du CMAP. Merci également aux membres de mon comité de suivi : Amandine Veber, Grégoire Allaire et Vincent Giovangigli pour leurs conseils avisés. Je tiens également à témoigner ma très grande reconnaissance à Anne de Bouard, directrice du CMAP au commencement de ce travail, pour sa bienveillance et son soutien dans ma démarche. D'une manière générale j'aimerais remercier l'ensemble du personnel du CMAP qui a su s'adapter lorsque j'étais moins disponible sur les questions informatiques. En particulier, je tiens à remercier très spécialement Pierre Straebler sans qui je ne m'en serais probablement pas sorti. Pierre je te dois beaucoup, je sais que tu as beaucoup pris sur toi, mais je n'en attendais pas moins d'un ex Flem'ard ! Pas de débat non plus pour décerner une mention spéciale à Laurent malgré son scandaleux rejet des traditions flamandes séculaires. Je souhaite aussi exprimer ma gratitude à Jean-François pour m'avoir fait profiter de son expertise remarquable en trains électriques... compétence fort utile bien que trop méconnue. Merci à Aldjia et Nasséra qui ont su dénouer bien des problèmes aussi qu'à toute l'équipe administrative pour son aide et sa bonne humeur. Merci également aux équipes de l'IRBA et en particulier Jean-Christophe Bouy qui m'a consacré tant de temps. Je suis également très reconnaissant envers Christian Graff pour ses précieux conseils en évaluation perceptive.

Ce travail doit aussi énormément aux sportifs déficients visuels et aux associations avec qui nous avons travaillé. L'association MCV, première à nous suivre sur le projet, a joué un rôle majeur dans ce projet. Merci Stéphane pour ton enthousiasme et ta confiance, y compris lors des premiers balbutiements. Baudouin tu as tout de suite compris l'intérêt de ces travaux et tu nous a fait profiter de ta précieuse expérience comme guide et coach. Jacques, tu as su entraîner l'équipe avec ton dynamisme habituel. D'autres associations et sportifs amateurs nous ont fait profiter de leurs conseils

et de leur enthousiasme : Martin et Guillaume les intrépides d'Accessi'Méca Sport, l'incroyable équipe des Miro Volants, Aziz à la bonne humeur communicative, la fine équipe de Courir en duo (Sophie, Caroline), Janina qui fait un travail formidable à Savigny Athlé, les athlètes de la JAM Montrouge, Ludovic notre triathlète infatigable, Benoit, Nicolas et tous les anonymes coureurs d'un jour qui ont participé à nos essais.

# Table des matières

<b>I Localisation en temps réel</b>	<b>17</b>
<b>1 Principes et méthodes en localisation</b>	<b>21</b>
1.1 Proximité	21
1.2 Latération et angulation	21
1.2.1 Latération : approche circulaire	22
1.2.2 Latération : approche hyperbolique	24
1.3 Reconnaissance de motifs et vision par ordinateur	25
1.4 Navigation à l'estime (Dead reckoning)	25
<b>2 Technologies</b>	<b>27</b>
2.1 GNSS	27
2.1.1 Constellations :	28
2.1.2 Précision et source d'erreurs :	28
2.1.3 GNSS augmenté	29
2.1.4 Discussion : GNSS et guidage piéton	31
2.2 Technologies pour la mesure de distances	33
2.2.1 Télémétrie	33
2.2.2 Technologies radiofréquence : Wi-fi, ZigBee, BlueTooth	33
2.2.3 UWB	34
2.3 Centrale inertielle, capteur d'orientation	38
2.3.1 Les différents capteurs : modélisation et calibration	38
L'accéléromètre	38
Le gyroscope	39
Le magnétomètre	39
2.3.2 Estimation de l'orientation	42
2.4 Autres techniques, empreintes	42
<b>3 Vers un système de localisation pour le guidage des sportifs non voyants</b>	<b>43</b>
3.1 Localisation par réseau de capteurs UWB	43
3.1.1 Positionnement temps réel par filtre de Kalman et calibration par méthode des moindres carrés (LSE - <i>Least square Estimation</i> )	44
3.1.2 Calibration	45

3.1.3	Filtrage temps réel . . . . .	46
	Mise en œuvre . . . . .	48
	Algorithme : . . . . .	49
	L'exactitude . . . . .	50
	Robustesse : . . . . .	50
	Nombre minimal de balises : . . . . .	51
3.1.4	Conclusion . . . . .	51
3.2	Localisation par réseau de capteurs radar Doppler hyperfréquence . . . . .	52
3.2.1	Estimation de la vitesse et de la distance . . . . .	53
3.2.2	Estimation de la distance utilisant le RSS . . . . .	55
3.2.3	Filtre particulière pour la localisation en temps réel . . . . .	56
3.2.4	Mise en oeuvre et tests . . . . .	59
3.2.5	Conclusion . . . . .	61
3.3	Flot optique et localisation piétonne . . . . .	61
3.3.1	Odométrie piétonne par flot optique . . . . .	62
	Capteur de flot optique : . . . . .	62
	Déplacement d'un capteur dans le plan parfaitement parallèle au sol : . . . . .	63
	Capteur en conditions réelles : . . . . .	64
3.3.2	Exemples de localisation par flot optique . . . . .	66
	Fusion AHRS et flot optique . . . . .	66
3.3.3	Conclusion . . . . .	67

## II Synthèse Binaurale

71

### 4 Écoute spatialisée et localisation sonore

75

4.1	Système auditif . . . . .	75
4.2	Référentiel . . . . .	76
4.3	Indices de localisation . . . . .	76
4.3.1	Différence Interaurale de temps ( <i>Interaural Time Difference</i> ou <i>ITD</i> ) . . . . .	76
4.3.2	Différence Interaurale de niveau ( <i>Interaural Level Difference</i> ou <i>ILD</i> ) . . . . .	78
4.3.3	Cônes de confusion . . . . .	78
4.3.4	Indices dynamiques et spectraux . . . . .	79
4.4	Capacités humaines de localisation sonore . . . . .	79
4.4.1	Perception de la direction de la source sonore . . . . .	79
4.4.2	Perception de la distance . . . . .	80
4.5	Perception sonore chez les personnes aveugles . . . . .	80
4.5.1	Perception de l'intensité et de la fréquence . . . . .	80
4.5.2	Localisation sonore chez la personne Déficiante Visuelle (DV) . . . . .	81
4.5.3	Echolocation . . . . .	81
4.5.4	Des performances supérieures aux personnes valides ? . . . . .	81

<b>5</b>	<b>HRTF et son spatialisé</b>	<b>83</b>
5.1	HRTF	83
5.2	Décomposition ITD + filtre à minimum de phase	85
5.3	Son spatialisé : Enregistrement et synthèse binaurale	87
<b>6</b>	<b>Filtrage binaural</b>	<b>89</b>
6.1	Implémentation	89
6.1.1	Algorithme	89
6.1.2	Choix technologiques	90
6.2	Latence d'un moteur audio	91
6.2.1	Latence audio	91
6.2.2	Temps de traitement des buffers audio	92
6.2.3	Mesure des latences système	92
6.2.4	Latence du head-tracker	94
6.3	Auralisation temps réel pour le guidage audio	95
	Implémentation et performances :	98
6.4	Conclusion	99
<b>7</b>	<b>Évaluation de filtres pour le guidage</b>	<b>101</b>
7.1	Filtres	101
7.1.1	HRTF	101
7.1.2	Panning	102
7.2	Méthode	103
	Identifications des expériences :	103
7.3	Expériences <b>N.Stat.</b> et <b>N.Dyn.</b> : Localisation statique et dynamique de sons naturels	104
	L'expérience N.Stat. (source fixe)	104
	L'expérience N.Dyn. (source mobile)	105
7.3.1	Notes concernant le dispositif expérimental	105
7.3.2	Résultats	105
7.4	Expériences <b>V.Stat.</b> et <b>V.Dyn.</b> : Localisation statique et dynamique de sons spatialisés	106
	L'expérience <b>V.Stat.</b> (source fixe)	107
	L'expérience <b>V.Dyn.</b> (source mobile)	107
7.4.1	Résultats	107
	Tests statiques <b>V.Stat.</b> :	107
	Tests dynamiques :	107
7.4.2	Discussions	108
7.5	Conclusion des expériences	109
<b>III</b>	<b>Vers un dispositif de guidage pour la pratique sportive</b>	<b>113</b>
<b>8</b>	<b>Navigation et locomotion des personnes DV</b>	<b>117</b>
8.1	Cognition spatiale	117



8.1.1	Sources d'informations lors d'un déplacement . . . . .	117
8.1.2	Représentation de l'espace et référentiels . . . . .	118
8.2	Navigation globale et fine, locomotion . . . . .	118
8.3	Pratique sportive chez les déficients visuels . . . . .	119
8.3.1	Les sports pratiqués par les personnes déficientes visuelles . . . . .	119
8.3.2	Compétitions : . . . . .	121
8.3.3	Guidage . . . . .	121
8.4	Substitution et augmentation sensorielle . . . . .	122
<b>9</b>	<b>Les dispositifs d'aide à la mobilité des personnes déficientes visuelles</b>	<b>125</b>
9.1	Dispositifs classiques . . . . .	125
9.2	Dispositif électroniques . . . . .	125
9.2.1	Historique et état de l'art . . . . .	125
9.2.1.1	Les premiers développements . . . . .	125
9.2.1.2	L'approche actuelle . . . . .	126
9.2.1.3	Tendance . . . . .	127
9.3	Le son spatialisé comme outil de guidage . . . . .	127
9.3.1	Le suivi sonore comme tâche écologique pour la personne déficiente visuelle . . . . .	129
9.3.2	Guidage sonore et pratique sportive . . . . .	129
<b>10</b>	<b>Un dispositif de guidage par son spatialisé pour la pratique sportive</b>	<b>131</b>
10.1	Conception générale du système . . . . .	131
10.2	Architecture . . . . .	133
10.2.1	Conception logicielle . . . . .	133
10.2.2	Conception matérielle . . . . .	134
	Encombrement : . . . . .	135
10.3	Restitution audio . . . . .	136
	Stimuli : . . . . .	137
10.4	Gestion du parcours . . . . .	139
10.4.1	Enregistrement du parcours . . . . .	139
10.4.2	Suivi du parcours . . . . .	140
10.4.3	Courbure . . . . .	140
10.5	Stratégie de suivi du parcours . . . . .	140
10.5.1	Preuve de concept, suivre un parcours inconnu . . . . .	140
10.5.2	Contrôle en boucle fermée . . . . .	142
10.6	Tests du prototype . . . . .	142
	Méthodologie pour les pré-tests : . . . . .	143
10.6.1	Expériences de guidage en condition réelles . . . . .	143
10.6.1.1	Ligne droite . . . . .	143
10.6.1.2	Parcours ayant une géométrie connue : apprentissage et performance . . . . .	145
10.6.2	Techniques de guidage avancées . . . . .	147
10.6.2.1	Parcours sinueux avec courbes et portions rectilignes . . . . .	147

<i>TABLE DES MATIÈRES</i>	11
10.6.2.2 Signalisation des zones de courbure . . . . .	148
10.6.3 Retour des utilisateurs . . . . .	149
10.7 Conclusion de la troisième partie . . . . .	150
<b>IV Conclusion et Perspectives</b>	<b>151</b>
<b>A Communications, publications</b>	<b>159</b>
<b>B Filtrage, méthodes d'estimation d'états, fusion de données</b>	<b>161</b>
B.1 Filtrage bayésien . . . . .	161
B.2 Filtre de Kalman . . . . .	162
B.3 Filtres particulaires (PF) . . . . .	163
B.3.1 Introduction aux méthodes de Monte-Carlo . . . . .	163
B.3.2 Échantillonnage préférentiel . . . . .	164
B.3.3 Filtrage particulière . . . . .	165
<b>C Exemple de fusion de donnée flot optique et UWB</b>	<b>167</b>
<b>D Echolocalisation et guidage : expériences exploratoires</b>	<b>169</b>



# Introduction

## La déficience visuelle, causes et types

La déficience visuelle est une forme de handicap qui affecte fortement l'autonomie et la vie quotidienne des personnes qui en sont atteintes. Selon l'OMS [177], on comptait en 2010 : 39 Millions de personnes aveugles à travers le monde (2.7M en Europe) et 246M présentant une déficience visuelle sévère (25M en Europe).

Les causes de cécité sont diverses. La cataracte (opacification du cristalin) est à elle seule la cause de 51% des cas de cécité à travers le monde (principalement dans les pays en voie de développement) alors qu'en Europe elle est principalement due aux maladies de la rétine [146].

L'OMS distingue 4 niveaux de déficience visuelle qui vont de la déficience légère à la cécité totale. En France, la cécité légale est définie par une acuité visuelle inférieure à 1/20 pour le meilleur œil après correction. Pour autant l'importance du handicap subi n'est pas complètement corrélée avec l'importance de la déficience. En effet d'autres facteurs jouent des rôles importants, en particulier l'ancienneté du handicap, l'apprentissage de techniques compensatoires de locomotion ou le rôle des proches.

## Augmentation sensorielle

De nombreux dispositifs visant à pallier les difficultés liées aux handicaps visuels ont été développés en lien avec les évolutions technologiques, et aujourd'hui, les applications de synthèse vocale sur téléphone portable font partie du quotidien des personnes aveugles autant que la canne blanche. Les chercheurs ont cherché à suppléer à la déficience visuelle pour les différentes tâches impactées par le handicap, qu'il s'agisse des déplacements, de la reconnaissance d'objets ou de la communication écrite. En parallèle aux évolutions technologiques, les recherches en sciences cognitives et en psychologie expérimentale ont conduit à développer le concept de substitution sensorielle [16] qui vise à remplacer un sens par un autre (par exemple, substituer la vue par le toucher ou l'audition) en utilisant la capacité des personnes à s'adapter à ces nouvelles modalités. Cette approche, qui a conduit à de nombreuses études et développements, a également montré ses limites [141] et aujourd'hui peu de dispositifs d'aides aux personnes déficientes visuelles utilisant ce principe ont été adoptées par les utilisateurs. Une approche différente, dite d'*augmentation sensorielle* consiste à extraire l'information utile et à la traduire avant de la transmettre à l'utilisateur. L'information peut alors être captée par toute forme de technologies (caméra, capteurs de distances, GPS, etc.), traitée, analysée et transmise via une modalité adaptée.

## Guidage sonore

Cette approche d'augmentation sensorielle connaît aujourd'hui un développement important, mais la question du mode de transmission de l'information reste fondamentale. Dans cette thèse, nous faisons l'hypothèse que les indices de localisation sonore constituent un moyen efficace pour fournir des informations de direction adaptées aux tâches de guidage. En effet, les personnes déficientes visuelles utilisent régulièrement leurs capacités à localiser et à suivre des sources sonores dans la vie quotidienne. C'est le cas, lorsqu'elles doivent s'orienter vers une personne pour lui parler, ou lorsqu'il s'agit de suivre les bruits de pas d'une personne pour se déplacer dans la rue. C'est donc une tâche pour laquelle elles sont, a priori, entraînées. Cette capacité est d'ailleurs déjà utilisée pour la pratique sportive puisque des sportifs non voyants pratiquent déjà le roller ou la course à pied avec un guidage sonore. L'athlète suit alors son guide en s'orientant grâce au bruit de ses pas, de ses rollers, ou via un haut-parleur. Dans un dispositif d'augmentation sensorielle pour le guidage, on peut alors envisager de reproduire de tels sons en les spatialisant par filtrage binaural. Il s'agit ainsi de reproduire les propriétés de l'écoute naturelle par filtrage numérique, ce qui permet à l'utilisateur de percevoir des sources sonores virtuelles dans l'espace, à l'aide d'un casque audio stéréo conventionnel. Le son spatialisé présente alors le double avantage de correspondre à une pratique habituelle, limitant le besoin d'apprentissage, et de fortement limiter la latence de transmission de l'information. Cette technique est donc proche d'un contexte écologique préservant les boucles perception-action naturelle et proposant une représentation spatiale égo-centrée. Les techniques de guidage par son binaural semblent alors bien adaptées à la pratique sportive où la recherche de performance autant que le confort sont des considérations prépondérantes.

## Sport et handicap

En France, en 2017, sur 28 684 licenciés à la Fédération Française Handisport, 1 841 étaient des Déficiants Visuels (DV) [75] (à comparer aux 932 000 français atteint de déficience visuelle sévère). Le taux de pratique sportive en club est donc très faible chez les personnes DV. Ces chiffres peuvent en partie être expliqués par la difficulté à trouver des clubs permettant la pratique sportive et la nécessité d'être accompagné individuellement. En effet, la plupart des sports nécessitent un guide (athlétisme, roller, etc.) pour être pratiqués par les personnes DV. D'autres sports nécessitent des aménagements ou des équipements spécifiques (vélo, tandem, tir). En outre certains sports ont été spécifiquement développés pour les pratiquants DV comme le Goalball, le Torball, le Showdown ou encore le cécifoot.

L'OMS recommande [178] pour tous les adultes de moins de 65 ans de pratiquer au moins 150 minutes d'activité d'endurance d'intensité modéré par semaine. Les personnes qui pratiquent davantage d'activité physique ont une espérance de vie supérieure par rapport aux personnes moins actives. Si les bienfaits du sport d'un point de vue physique, psychologique et sociaux sont bien connus, on peut également mettre en avant un intérêt spécifique pour les personnes déficientes visuelles. En effet le sport permet de stimuler la sensibilité proprioceptive et en conséquence de développer la motricité et la perception des membres dans l'espace lors des mouvements, ce qui est susceptible d'avoir un impact direct sur l'autonomie.

Dans le sport de compétition, les personnes DV sont désignées par la lettre B (Blind) et classés en 3 catégories distinctes, suivant l'acuité visuelle et le champ visuel de : B1 (considéré comme non-voyant) à B3 (mal voyant). Aux JO 2024, les disciplines suivantes seront ouvertes aux personnes DV : Cécifoot, GoalBall, Athlétisme, Aviron, Cyclisme, Judo, Natation et Triathlon. Là encore, la question du guide est primordiale. Le guide doit être capable de jouer son rôle sans ralentir l'athlète, ce qui implique une très bonne coordination (et donc un entraînement en commun) et des

capacités sportives adaptées. De ce fait, au plus haut niveau, les guides sont également des athlètes de premier plan, dont les disponibilités sont forcément très limitées.

Dans le sport de loisir, comme en sport de compétition, la question de la disponibilité du guide est donc primordiale et constitue souvent un frein au développement de ces activités.

## Objectifs

Aujourd'hui, mises à part les applications GPS sur smartphone, très peu d'appareils sont réellement utilisés par les personnes déficientes visuelles. Sur le terrain, de nombreux projets de laboratoire perdent leur intérêt par manque d'ergonomie ou parce qu'ils interfèrent avec les autres activités de l'utilisateur. Ils peuvent nécessiter une formation longue ou augmenter sensiblement la charge cognitive. Sans surprise, les appareils les plus populaires actuellement sont des accessoires spécifiques, dédiés à des tâches bien identifiées, comme par exemple la détection d'obstacles proches (on pense à la canne blanche électronique développée par l'équipe de R. Farcy [74]). Par conséquent, l'objectif du dispositif électronique pour le guidage des personnes déficientes visuelles que nous cherchons à développer n'est pas de remplacer la vision dans son ensemble. Nous nous concentrons plutôt sur l'amélioration de la qualité de l'aide pour une tâche spécifique, à savoir orienter une personne sur un parcours en se focalisant sur la pratique sportive en environnement maîtrisé. Le son binaural comme indication d'orientation est utilisé pour fournir uniquement l'information essentielle selon un principe de parcimonie. La thématique *binaural/Son 3D* fut introduite au CMAP il y a plusieurs années par les travaux de François Alouges et Matthieu Aussal et a conduit à la conception d'outils de spatialisation sonore temps-réel efficaces sur lesquels ce travail s'appuie largement.

Toutefois, dans le but de pouvoir guider une personne grâce à un son spatialisé situé en amont et lui indiquant le chemin à suivre, nous devons d'abord *localiser précisément* l'utilisateur dans l'espace et créer la ou les sources virtuelles spatialisées *en temps réel* qui lui serviront de guide. Le dispositif devra enfin être *extrêmement réactif* pour pouvoir être utilisé en contexte sportif. Enfin, le dispositif complet doit pouvoir être *embarqué* par l'athlète.

Ce travail s'articule donc autour de trois axes, qui constituent les trois parties distinctes du manuscrit :

- le développement de techniques et d'outils pour la localisation précise ;
- le développement des moteurs binauraux embarqués et l'étude de la perception des sources sonores en mouvement ou pas ;
- l'intégration de ces éléments dans des prototypes utilisables pour l'expérimentation et le guidage.

## Partenariats

Pour développer des méthodes et des technologies réellement adaptées aux personnes déficientes visuelles, nous cherchons à adopter une approche centrée sur les besoins utilisateurs plutôt que sur la technologie. Pour cela, nous nous sommes rapprochés d'associations sportives de personnes déficientes visuelles, en particulier Mix handi'Cap sur la Vie [2] (roller, sport partagé, activités culturelles), Accessi Meca Sport [1] (sports mécaniques accessibles aux personnes DV) et Savigny Athlétisme 91 [3] (athlétisme et sport adapté). Riches de ces partenariats, et dans une démarche itérative, nous avons cherché à mieux comprendre les techniques utilisées au quotidien par les personnes aveugles pour appréhender l'espace afin d'en tirer le meilleur profit pour réaliser un guidage efficace.



**Première partie**

**Localisation en temps réel**





# Introduction

Les systèmes de localisation en temps réel (en anglais *Real Time Localisation Systems* ou *RTLS*) sont des dispositifs destinés à suivre les déplacements d'objets ou de personnes en temps réel<sup>1</sup>. Certains de ces systèmes sont adaptés pour une utilisation en intérieur, alors que d'autres fonctionnent uniquement en extérieur.

Nous verrons qu'un système destiné au guidage par audio binaural devra être la fois suffisamment précis et réactif pour pouvoir prendre en compte chaque mouvement du sujet à guider. Cela impliquera le développement et la mise en œuvre de systèmes adaptés.

Cette partie est organisée en trois chapitres : dans un premier temps nous décrivons les quatre principaux paradigmes de localisation, puis nous présenterons un panorama des algorithmes de mesures de distances avant de traiter des technologies matérielles sur lesquelles un tel système peut s'appuyer. Après cet état de l'art, nous décrivons nos travaux sur les méthodes et les technologies de localisation que nous avons développées pour répondre à la problématique de guidage.

Il convient dans un premier temps de comprendre les spécificités d'un dispositif de localisation, qui permettront de les évaluer et de les qualifier. On définit généralement les principales qualités que l'on peut attendre de ce type de systèmes en considérant les caractéristiques suivantes :

L'**exactitude** ou la **justesse** (accuracy) d'un système de localisation est souvent définie comme la moyenne de la distance euclidienne entre la position estimée et la position réelle [148], autrement dit, il s'agit de l'estimation du biais statistique. On distingue souvent l'exactitude de la **précision** qui correspond alors à la dispersion des mesures. La précision pourra alors typiquement être exprimée sous la forme de l'écart type de l'erreur. Cependant, le terme de précision est parfois utilisé pour désigner à la fois les erreurs aléatoires (variabilité) et systématiques (biais) ce qui peut rendre l'utilisation de ces termes délicates.

La **robustesse** ou la **disponibilité** pourront être exprimées sous la forme d'un pourcentage de temps où une mesure fiable est disponible. En effet, pour certaines technologies, un signal peut être absent ou trop faible pour être exploitable.

La **zone de couverture** définit la surface ou le volume où le système est disponible et utilisable. Il peut être très variable suivant les techniques utilisées. Pour étendre cette couverture, certains systèmes peuvent être densifiés ou couplés entre eux. Cette capacité à **passer à l'échelle** est alors un critère de choix important.

La **fréquence de mesure**, qui déterminera en particulier l'aspect *temps réel du système* est la fréquence déterminée par l'intervalle de temps entre deux mesures, d'où découle également la **latence** du système.

Enfin, le **coût** du système est évidemment un critère important. Il devra être envisagé de manière globale, en considérant à la fois le coût de déploiement ainsi que les coûts d'équipement des usagers.

---

1. Temps réel doit s'entendre ici comme *au fil du temps* en introduisant le moins de latence possible entre la localisation et le déroulé réel de l'événement. Dans la pratique les latences de ce type de systèmes peuvent être de l'ordre de la milliseconde à la seconde.



# Chapitre 1

## Principes et méthodes en localisation

Nous commencerons dans un premier temps à décrire les grands paradigmes utilisés en localisation. De manière non exhaustive, on décrira brièvement les quatre principales approches : *l'identification directe d'une position (ou détection de proximité)*, *les techniques de latération ou de triangulation*, *la reconnaissance de motif* et *la navigation à l'estime*. Ces approches étant bien entendu non exclusives, les applications pratiques s'appuient souvent sur plusieurs d'entre elles comme nous le verrons dans les applications mises en œuvre dans ce document.

### 1.1 Proximité

La détection de proximité est conceptuellement la méthode la plus simple puisqu'elle consiste à établir la proximité avec des balises placées dans l'environnement. L'information est essentiellement binaire (proche/non-proche) et la précision dépend de la densité des balises utilisées. On peut alors distinguer deux modes de fonctionnement, soit la balise fournit directement une information sur sa position, soit le système va identifier la balise en se référant à une table. Ces dispositifs, généralement simples à mettre en œuvre peuvent par exemple être inclus dans une canne [53] ou une chaussure [232], cependant cette technique n'offre qu'une précision limitée et pose des difficultés de passage à l'échelle.

### 1.2 Latération et angulation

Ces méthodes consistent à utiliser des mesures de distances (latération) ou d'angles (triangulation ou angulation) par rapport à des repères dont les positions sont connues. Ces approches qui figurent parmi les plus anciennes (par exemple, le sextant est un instrument de triangulation ancien) sont aujourd'hui toujours très largement utilisées. Elles sont par exemple à la base des systèmes de localisation par satellites. Pour la latération, il existe deux approches : circulaire et hyperbolique. Dans le premier cas on utilise les distances entre l'objet mobile et un ensemble de points de référence, et dans le second on utilise les différences entre les distances des objets.

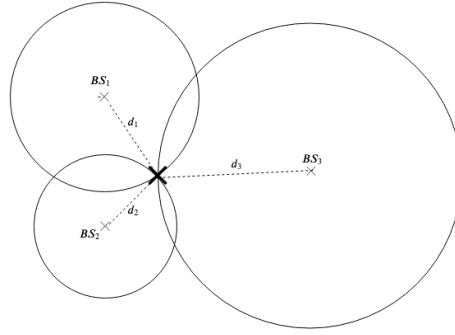


FIGURE 1.1 – Un objet Mobile à localiser (OM) et trois balises fixes (BS).

### 1.2.1 Latération : approche circulaire

Résoudre un problème de latération peut être vu comme la recherche d'un point d'intersection de cercles en 2D ou de sphères en 3D (voir Fig. 1.1). Estimer la position d'un objet sans ambiguïté sur deux dimensions nécessite alors de connaître au moins les distances à trois points non alignés, contre au moins 4 points non coplanaires en trois dimensions.

Ainsi, en dimension 2, si on considère le point à déterminer  $M(x, y)$  et trois balises fixes  $BS_i$  aux positions  $(x_i, y_i)$ , et en appelant  $r_i$  la distance entre  $M$  et  $BS_i$ , on obtient les trois équations :

$$\begin{cases} (x - x_1)^2 + (y - y_1)^2 = r_1^2 \\ (x - x_2)^2 + (y - y_2)^2 = r_2^2 \\ (x - x_3)^2 + (y - y_3)^2 = r_3^2 \end{cases} \quad (1.1)$$

que l'on peut écrire sous la forme du système suivant :

$$\begin{aligned} \underbrace{(-2x_1 + 2x_2)}_A x + \underbrace{(-2y_1 + 2y_2)}_B y &= \underbrace{r_1^2 - r_2^2 - x_1^2 + x_2^2 - y_1^2 + y_2^2}_C, \\ \underbrace{(-2x_2 + 2x_3)}_D x + \underbrace{(-2y_2 + 2y_3)}_E y &= \underbrace{r_2^2 - r_3^2 - x_2^2 + x_3^2 - y_2^2 + y_3^2}_F. \end{aligned} \quad (1.2)$$

La résolution de ce système aboutit à l'expression analytique de la solution :

$$\begin{aligned} x &= \frac{CE - FB}{EA - BD} \\ y &= \frac{CD - AF}{BD - AE}. \end{aligned} \quad (1.3)$$

La solution précédente (1.3) du problème de trilatération ne tient pas compte du bruit de mesure présent lors de mesures réelles et l'intersection des cercles peut ne pas exister. De façon classique, on peut alors considérer une approche par moindres carrés qui consistera alors à trouver la position  $(x, y)$  qui minimise l'expression :

$$\sum_{i=1}^n |r_i^2 - ((x - x_i)^2 + (y - y_i)^2)|. \quad (1.4)$$

Cette minimisation peut être effectuée par des méthodes numériques classiques (Gauss-Newton, Levenberg–Marquardt, etc.). En outre, la méthode permet de multiplier les mesures (les systèmes sont alors surdéterminés) en équilibrant la

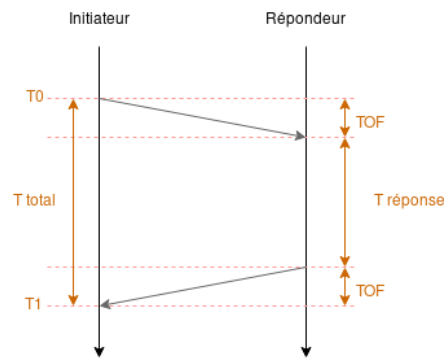


FIGURE 1.2 – Principe du protocole Two-Way Ranging.

solution entre les mesures.

En ce qui concerne les méthodes utilisées pour mesurer des distances, il y a principalement les deux approches suivantes :

- Le *Time of Arrival* ou *Time of Flight* qui est par définition le temps parcouru par un signal entre un émetteur  $A$  et un récepteur  $B$ . L'algorithme *One Way Ranging* [63] consiste à évaluer directement ce temps de trajet. Un émetteur  $A$  envoie un signal horodaté qui est ensuite reçu par un récepteur et qui va alors calculer le temps écoulé lors de la transmission. Cette méthode a peu d'intérêt pratique car elle nécessite une parfaite synchronisation de l'émetteur et du récepteur, qui peut être difficile à mettre en œuvre en pratique. La méthode *Two Way Ranging* est beaucoup largement plus répandue (voir Fig. 1.2). Elle cherche à mesurer le temps d'aller-retour d'un signal. Le signal est envoyé par l'émetteur-récepteur  $A$ , puis reçu par l'émetteur-récepteur  $B$  et ré-émis par celui-ci après un temps connu.  $A$  va alors recevoir un signal de retour. Le temps de vol correspond alors à la moitié de la durée écoulée, à laquelle on aura préalablement soustrait le temps de réponse. Cette technique ne nécessite pas de synchronisation d'horloge, elle est simple à mettre en œuvre en particulier dans les environnements comportant peu de balises fixes et/ou de mobiles à localiser.
- Le *Received Signal Strength* (RSS). Cette technique consiste à déduire une distance en considérant que la puissance d'un signal reçu est fonction de la distance parcourue.

Différents modèles ont été proposés pour représenter l'évolution de la puissance d'un signal avec la distance en tenant compte de l'environnement.

Si on considère que l'émetteur et le récepteur sont en champ libre, en l'absence d'obstacles, sans réflexion ni absorption, en 3D on peut utiliser le modèle suivant :

$$P_r(d) = C_f \frac{P_t}{R^2} \quad (1.5)$$

où  $P_t$  et  $P_r$  sont respectivement les puissances émise et reçue,  $C_f$  est une constante dépendante des *transceivers* et  $R$  la distance.

Dans le cas d'une transmission par onde électromagnétique, on utilise généralement l'équation de Friis [86] (dites *Équation des télécommunications*) :

$$P_r = P_t G_t G_r \left( \frac{\lambda}{4\pi R} \right)^2, \quad (1.6)$$

où  $G_t G_r$  est le produit des gains des antennes d'émission et de réception,  $\lambda$  est la longueur d'onde et  $R$  est la distance entre les antennes.

Des modèles ont été proposés pour tenir compte des réflexions sur le sol et sur d'éventuels obstacles. Dérivés de méthodes analytiques et empiriques, l'évolution de la puissance reçue peut être formulée sous la forme [200], [26]

$$\frac{P_r(d_o)}{P_r(d)} = \left( \frac{d}{d_o} \right)^\alpha, \quad (1.7)$$

où le paramètre  $\alpha$ , souvent appelé dans la littérature *gradient de distance-puissance*, prend une valeur comprise entre 2 et 6 suivant l'environnement. A titre indicatif, des exposants typiques sont données en Table 1.1.

Environnement	$\alpha$
Champ libre	2
Env. urbain extérieur	2.7 à 5
Intérieur en vision directe	1.6 à 1.8
Intérieur avec vision non directe	4 à 6

TABLE 1.1 – Exposants  $\alpha$  typiques (d'après [123], [180]).

## 1.2.2 Latération : approche hyperbolique

Dans cette méthode on va utiliser les différences de distance entre l'objet mobile et les différentes balises de référence. Le lieu des points pour lesquels la différence des distances à deux points fixes (balises) est constante étant une hyperbole, c'est l'intersection de ces hyperboles, obtenues en faisant varier le couple de balises considérées, qui permettra de trouver la position de l'objet.

Soit  $N$  stations de bases positionnées en  $BS_n = (x_n, y_n)$  pour  $n = 1, \dots, N$  et un objet mobile de coordonnées  $OM = (x_s, y_s)$ . On peut alors écrire la distance de l'objet mobile à la  $n$ -ème station de base sous la forme :

$$d_n = \sqrt{(x_s - x_n)^2 + (y_s - y_n)^2}. \quad (1.8)$$

La différence des distances entre l'objet mobile et les  $n$ -ème et  $(n + 1)$ -ème stations peut s'écrire :

$$d_{n+1,n} = \sqrt{(x_s - x_{n+1})^2 + (y_s - y_{n+1})^2} - \sqrt{(x_s - x_n)^2 + (y_s - y_n)^2}. \quad (1.9)$$

Dans ce cas, la méthode utilisée s'appelle *Time Difference of Arrival (TDOA)* [63]. Cette technique met en œuvre des balises de référence déployées dans l'environnement à des positions connues et dotées d'horloges parfaitement synchronisées. Le mobile émet des trames par intermittence, et lorsqu'une ancre (balise) reçoit cette trame, elle l'horodate. Les temps d'arrivée aux différentes balises sont alors enregistrés.

La localisation par TDOA consiste alors à résoudre des systèmes d'équations hyperboliques de la forme (1.9) à partir de ces mesures.

Cette méthode présente certains avantages car en limitant les messages de synchronisation entre les balises et l'objet mobile elle peut permettre de localiser plus facilement plusieurs objets (on dit qu'elle passe à l'échelle ou est *scalable*).

## 1.3 Reconnaissance de motifs et vision par ordinateur

La **reconnaissance de motifs** consiste à comparer des mesures effectuées par les capteurs à des mesures effectuées dans l'environnement au préalable. On part alors du principe que les caractéristiques d'un signal mesuré sont dépendantes du lieu. Pratiquement, on va donc dans un premier temps chercher à enregistrer une carte de la distribution du signal dans l'environnement, pour ensuite identifier des zones lors de l'étape de localisation. Plusieurs algorithmes ont été proposés, comme par exemple les  $k$ -proches voisins [18], les méthodes d'inférence Bayésiennes [135], les réseaux de neurones [72] ou les supports vecteurs machines [36].

Paradigme assez comparable, le **la vision par ordinateur** est une technique qui consiste à identifier des marqueurs, lignes, objets ou toutes caractéristiques visuelles de l'environnement à partir d'images vidéo collectées en temps réel par une caméra. La vision par ordinateur permet d'analyser dynamiquement l'environnement, elle est donc particulièrement intéressante pour l'évitement d'obstacle et les technique de *localisation et cartographie simultanées* (SLAM, Simultaneous Localization And Mapping) permettant d'évoluer dans un environnement inconnu ou mal maîtrisé. Les méthodes de vision par ordinateur utilisent souvent la reconnaissance d'images pré-enregistrées [246], de formes ou de caractéristiques. Ces techniques ont trouvé plusieurs applications dans l'aide à la mobilité des personnes aveugles par reconnaissance d'objets (portes, chaises) [106] ou en utilisant la reconstruction 3D par stéréoscopie [42] pour l'évitement d'obstacles.

Dans sa version la plus simple, la vision par ordinateur peut être utilisée comme outil **d'odométrie par analyse du flot optique**. Cette technique consiste à estimer le mouvement apparent des objets filmés par une caméra en comparant les séquences d'images successives.

On fait l'hypothèse d'une illumination constante dans le temps. Si on considère un point de l'image d'intensité  $I(x, y, t)$  qui se déplace de  $\Delta x$ ,  $\Delta y$  en  $\Delta t$ , l'hypothèse d'illumination constante peut s'écrire :

$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t). \quad (1.10)$$

Le flot optique au temps  $t$  et au point  $(x(t), y(t))$  est défini par la vitesse du point image  $\mathbf{V} = (u, v) = (\frac{\Delta x}{\Delta t}, \frac{\Delta y}{\Delta t})$

Le flot optique est donc contraint par l'équation suivante :

$$I_x \cdot u + I_y \cdot v + I_t = \nabla I^T \cdot (u, v, 1) = 0 \quad (1.11)$$

où  $I_x$ ,  $I_y$  et  $I_t$  représentent les composantes horizontales, verticales et temporelles du gradient d'intensité  $I$ . En mesurant le gradient d'intensité, on peut, grâce à (1.11), obtenir une composante de la vitesse de déplacement du point image. Si on souhaite déterminer complètement la vitesse de déplacement, cette technique est habituellement complétée par des méthodes locales qui vont prendre en compte des hypothèses supplémentaires sur un domaine de taille réduite. Par exemple, la méthode de Lucas & Kanade [151] suppose que le flot est constant dans un voisinage local du pixel considéré.

## 1.4 Navigation à l'estime (Dead reckoning)

La technique consiste à estimer une position à partir d'un état précédent connu ou estimé. La navigation à l'estime utilise généralement des données d'odométrie qui permettent d'estimer un déplacement à chaque pas de temps (distance et/ou orientation). Elle a pour avantage de ne nécessiter aucune infrastructure de mesure dans la surface



d'évolution. Par nature c'est une technique itérative qui nécessite une estimation initiale de position et conduit à l'accumulation de l'erreur avec le temps, elle est donc adaptée pour les temps courts. En général, pour des résultats fiables sur des temps longs, il est nécessaire d'incorporer régulièrement une correction à l'aide d'une autre méthode. Ainsi, elles peuvent facilement être associées aux techniques de *proximité de balises* qui permettent de re-synchroniser les positions mesurées avec des positions de référence, ou avec des techniques de triangulation.

Ces techniques, très populaires en robotique sont plus complexes à mettre en œuvre pour la localisation des personnes en raison du manque de capteurs d'odométrie disponibles. À titre d'exemple, des algorithmes d'odométrie spécifiques ont été développés pour la localisation des personnes en utilisant l'analyse des enjambées [211], [84]. D'une manière générale des méthodes de navigation à l'estime s'appuient sur des méthodes d'estimation d'état comme le filtrage de Kalman ou les filtres particuliers (décrits en annexe B).

# Chapitre 2

## Technologies

Les méthodes de localisation décrites au chapitre 1 peuvent être implémentées avec différentes technologies physiques dont on propose de faire un panorama rapide. On peut distinguer les technologies passives, cherchant à analyser l'environnement (empreinte de champ magnétique, gravité, vision) et les technologies actives qui vont interagir avec l'environnement (e.g. émission/réception d'un signal).

De façon pratique, les technologies mises en oeuvre en environnement extérieur sont souvent différentes de celles utilisées en intérieur. Dans cet exposé, on présentera dans un premier temps la géolocalisation par satellites qui est une technologie spécifique pour l'extérieur puis on donnera un aperçu des principales technologies utilisables en intérieur en mettant l'accent sur les outils que nous utiliserons par la suite (Centrale inertielle, Ultra-Wide Band, Flot-optique et radar Doppler).

### 2.1 GNSS (Géolocalisation et Navigation par un Système de Satellites/Global Navigation Satellite System)

Popularisé par le système américain GPS, cette technologie est probablement la plus connue pour localiser, avec plus ou moins de précision, un opérateur en extérieur muni d'un récepteur adéquat. Cette technologie équipe maintenant tous les téléphones cellulaires et est très utilisée par les automobilistes par exemple.

Ces systèmes utilisent des informations temporelles fournies par des satellites pour déterminer la position des récepteurs. Chaque satellite transmet des informations concernant son orbite (éphéméride) et sa position ainsi qu'une horodate de précision synchronisée par une horloge atomique. Les récepteurs comparent alors les signaux reçus par plusieurs satellites, en déduisent les temps de vols et peuvent alors évaluer leur position.

En appelant  $t_r$  le temps où le signal est reçu et  $t_i$  le temps où il est émis par le satellite  $i$ , la mesure de la *pseudo-distance*  $D_i$  entre le récepteur de coordonnées  $(U_x, U_y, U_z)$  et le satellite  $i$  de coordonnées connues  $(X_i, Y_i, Z_i)$ , peut s'écrire :

$$D_i = c(t_r - t_i) = \sqrt{(U_x - X_i)^2 + (U_y - Y_i)^2 + (U_z - Z_i)^2}. \quad (2.1)$$

En réalité les récepteurs disposent d'une horloge imparfaite qui n'est pas parfaitement synchronisée avec le temps des satellites et toutes les mesures sont réalisées avec un décalage d'horloge (offset). C'est pourquoi on parle de mesures de *pseudo-distances*. En revanche, les satellites de la constellation sont supposés parfaitement synchrones

entre eux et l'offset peut être considéré comme étant identique quel que soit le satellite. En introduisant l'offset  $\Delta t$  comme inconnue supplémentaire l'équation (2.1) devient :

$$D_i = \sqrt{(U_x - X_i)^2 + (U_y - Y_i)^2 + (U_z - Z_i)^2} + c\Delta t. \quad (2.2)$$

Ainsi les mesures d'un minimum de quatre pseudo-distances sont nécessaires pour établir une position et un temps en 3D (en éliminant la solution qui n'est pas sur la surface terrestre) et un minimum de trois mesures pour déterminer une position et un temps si l'altitude est connue.

### 2.1.1 Constellations :

Plusieurs constellations de satellites sont actuellement en service. Le système américain **GPS** (Global Positioning System) a été le premier mis en service vers la fin des années 70, et est aujourd'hui constitué d'une trentaine de satellites en orbite moyenne (qui effectuent une révolution en 12H) répartis sur 6 plans. Il fut rapidement suivi par le système russe **GLONASS** (déployé à partir des années 80) qui est aujourd'hui constitué d'environ 25 satellites. Depuis 2011, cette dernière constellation couvre l'ensemble de la surface terrestre. Le système Chinois **BEIDU** a, quant à lui, été développé dans les années 2000 et est aujourd'hui composé d'une trentaine de satellites. Cependant, sa couverture hors d'Asie reste partielle. Dernier en date, le système européen **Galileo**, dont le déploiement est en cours d'achèvement, est lui constitué d'une trentaine de satellites (dont 6 de secours) offrant une couverture mondiale. Les satellites Galileo sont répartis dans trois plans de façon à ce qu'il y ait 8 satellites opérationnels par plan orbital. Enfin, on peut également citer des constellations dont la couverture n'est pas globale, comme le système **NAVIC** (NAVigation with Indian Constellation) indien ou le **Quasi-Zenith Satellite System** (QZSS) japonais.

### 2.1.2 Précision et source d'erreurs :

La précision des systèmes GNSS dépend de deux facteurs indépendants : l'erreur sur les pseudo-mesures et la dilution géométrique de la précision liée aux angles d'observation des satellites.

L'erreur sur les pseudo-mesures est principalement liée [68, 111] aux variations du temps d'arrivée des signaux dans l'ionosphère et la troposphère (nous noterons l'écart type de l'erreur induite  $\sigma_{IT}$ ), au biais de l'horloge du satellite, aux erreurs d'éphéméride ( $\sigma_{\text{spatial}}$ ), à un ensemble d'erreurs  $\sigma_{\epsilon_p}$  liés aux trajets multi-chemins (particulièrement importants en environnement urbain) et aux bruits du récepteur (y compris l'horloge).

Pour quantifier cette erreur, on considère généralement l'*UERE* (*User Equivalent Range Error*), qui correspond à la contribution de toutes les sources d'erreurs supposées indépendantes sur la mesure des pseudo-distances :

$$\sigma_{\text{UERE}} = \sqrt{\sigma_{\text{spatial}}^2 + \sigma_{IT}^2 + \sigma_{\epsilon_p}^2}. \quad (2.3)$$

La précision avec laquelle les positions peuvent être déterminées n'est pas seulement fonction de la précision des mesures. Elle est également fonction de la configuration spatiale des satellites et du récepteur. L'effet de la géométrie de la configuration des satellites est généralement exprimée par le facteur de *dilution géométrique de la précision* (GDOP) [139]. C'est une grandeur adimensionnée, proportionnelle à l'inverse du volume du polyèdre formé par le récepteur et les satellites visibles.

La GDOP est le rapport entre la précision du positionnement et la précision de la mesure (UERE). La précision à

1- $\sigma$  de la mesure peut ainsi s'écrire :

$$\text{Precision} = \text{GDOP} * \text{UERE} . \quad (2.4)$$

Ainsi une valeur de GDOP élevée va amplifier les erreurs de mesure des pseudo-distances. La GDOP varie en fonction du nombre de satellites considérés. Elle est inférieure à l'unité quand l'information est redondante, ce qui est souvent le cas en pratique pour les récepteurs multi-constellations qui donnent accès à un plus grand nombre de satellites.

Globalement pour un système GNSS utilisant une seule bande de fréquence - seule possibilité offerte aux applications grand public jusqu'aux années 2010 - l'erreur pour le positionnement vertical est inférieure à 8m dans 95% des mesures pour des cas d'utilisation typiques [213].

Les **GNSS multi-bandes** utilisent les différents canaux émis par les satellites. Pour le système GPS, les cinq canaux disponibles sont nommés de L1 à L5. Pour les applications civiles, le canal L1 (1575.42 MHz) historique a été complété par un canal L2 (1176.45 MHz) sur les satellites mis en orbite depuis 2010. Avec Galileo, comme sur le GPS, un canal haute fréquence (E1, 1575.42MHz) et un basse fréquence (E5, 1176.45MHz) sont dédiés aux applications de localisation pour le grand public.

La réception de ces deux fréquences permet de pallier partiellement à deux grandes sources d'erreurs des systèmes GNSS, à savoir les erreurs liées à l'ionosphère (1) et les multi-chemins (2).

1. Le retard apparent  $I_r^s$  lié à l'ionosphère peut s'exprimer empiriquement (d'après [130, 168]) :

$$I_r^s = \frac{40.3}{cf^2} \text{TEC} , \quad (2.5)$$

où  $f$  est la fréquence du signal et TEC la quantité d'électrons libres par unité de surface.

Ce retard est donc dépendant de la fréquence. La différence de retard entre les deux canaux permet alors d'estimer le retard de groupe et ainsi d'effectuer une correction des erreurs susceptible d'augmenter la précision de localisation.

2. Alors que le canal L1 transmet un débit binaire de 1.023MHz (modulation par changement de phase à 1.023MHz ou BPSK(1)), les canaux L2/E5 offrent des débits binaires 10 fois supérieurs (10.23MHz en modulation BPSK(10)). Cette différence affecte les pics d'autocorrélation, qui, plus étroits pour les canaux L2/E5, peuvent être plus facilement distingués (cf. Fig. 2.1). Les récepteurs peuvent alors plus facilement éliminer les échos indésirables des trajets multichemins.

Des puces GNSS grand public sont disponibles sur le marché depuis 2018 (ex : Broadcom BCM47755, U-Blox ZED-F9) et équipent les dernières générations de téléphones portables. Selon [70] ou [195], et nos propres expériences, ce type de récepteur permet d'atteindre une précision de l'ordre du mètre lors de tests en mouvements ciel visible.

### 2.1.3 GNSS augmenté

Les techniques de **GNSS différentiels**, ou D-GNSS, utilisent le fait que pour des positions proches, les erreurs de mesures sont très comparables. On va alors utiliser une station de référence proche qui connaît précisément sa position et transmettra les erreurs de positionnement au récepteur mobile. Ce dernier pourra alors corriger sa propre mesure. Ce type de système peut permettre d'atteindre une précision décimétrique pour les meilleurs implémentations.

La station de référence peut être un deuxième récepteur, placé dans les meilleures conditions de réception possibles à une distance de quelques mètres à quelques kilomètres et capable de transmettre ses informations par radio. Il peut également s'agir d'un réseau de bases géré par un tiers privé (en France par exemple les sociétés Teria ou

### How does the L5 signal help?

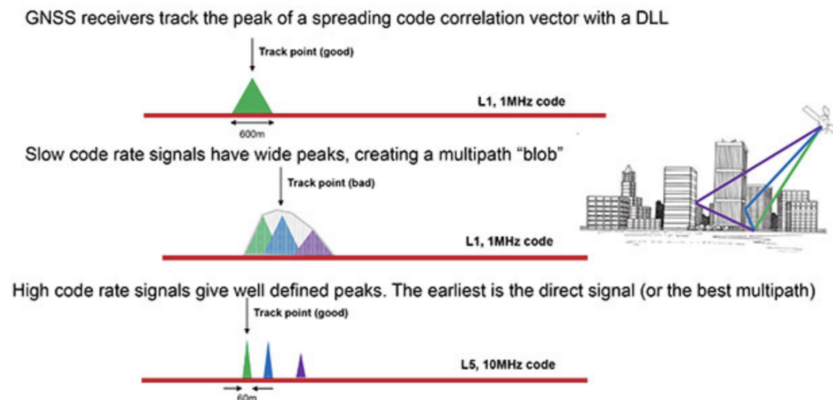


FIGURE 2.1 – Pic d'autocorrélation des signaux L1 vs L2 en environnement multi-chemin. La figure est extraite de [171]. Les pics d'autocorrélation, plus étroits peuvent être plus facilement distingués.

Orphéon proposent de tels réseaux de bases). Dans ce cas, le maillage sera moins fin<sup>1</sup> et les communications seront typiquement assurées par le réseau téléphonique cellulaire. Le coût des récepteurs D-GNSS a récemment beaucoup chuté ce qui rend cette technologie accessible<sup>2</sup>.

Le système **RTK** (Real Time Kinematic) est une amélioration du D-GNSS. Un système GNSS classique reçoit le signal numérique des satellites (i.e. le signal modulé) et cherche à l'aligner avec un signal de référence. Dans un système RTK, l'alignement se fait directement sur la phase du signal analogique reçu (i.e. la porteuse) ce qui permet d'améliorer la précision. Dans la pratique il peut se produire des erreurs d'un ou plusieurs cycles<sup>3</sup> difficiles à éliminer. Dans la situation idéale, sans aucun décalage de cycle, on parle dans la littérature de mesures *RTK fixées*, sinon de mesures *RTK flottantes*. Les erreurs de positionnement en D-GPS sont typiquement de l'ordre du mètre, de 20-30cm en RTK flottant et de l'ordre de quelques centimètres en RTK fixé [179].

Les systèmes *d'augmentation par satellite* (SBAS ou Satellite-Based Augmentation Systems) permettent une couverture différentielle à l'échelle régionale ou continentale en utilisant les mesures de stations au sol. Des messages de correction sont alors envoyés à un ou plusieurs satellites géostationnaires pour être diffusés aux utilisateurs. La couverture de ces systèmes se limite à certaines régions du monde (Europe, Amérique du nord, Russie, Asie du sud-est). Le continent européen est lui couvert dans sa totalité par le système EGNOS (European Geostationary Navigation Overlay Service) d'accès gratuit et développé et mis en œuvre par l'Agence Spatiale Européenne et EUROCONTROL. Selon l'ESA [54], la précision obtenue à  $2\sigma$  est alors de l'ordre de 1 à 3 mètres en horizontal et de 2 à 4 mètres en vertical. Les satellites EGNOS, en position géostatique sur le plan équatorial sont typiquement visibles de nos latitudes à environ  $30^\circ$  au dessus de l'horizon. Pour des applications au sol, en milieu urbain, ils sont donc peu visibles et leur usage reste donc limité dans le cadre que nous envisageons.

De façon comparable, mais à l'échelle locale, les systèmes *Ground-Based Augmentation System* (GBAS) offrent eux une couverture visant à améliorer les niveaux de service GNSS en termes de précision, d'intégrité et de sécurité

1. jusqu'à 100km entre les bases

2. le coût est maintenant inférieur à 1000 euros pour un système complet station de base + récepteur mobile par exemple : Ublox dev kit NEO-M8P, Drotek NEO M8P, ArduSimple (Ublox F9P). Il est également possible de développer des récepteurs D-GNSS en utilisant un matériel à bas coût et associé à la bibliothèque Open-source RTKLib [211].

3. Chaque erreur de cycle produit une erreur d'une longueur d'onde, soit environ 19cm aux fréquences qui sont en jeu.

pour l'aviation pendant les phases d'approche, d'atterrissage et de départ. Ces systèmes normalisés dans l'aviation sont généralement disponibles à proximité des aéroports.

### 2.1.4 Discussion : GNSS et guidage piéton

Les récepteurs GPS, principalement sur téléphone, sont utilisés depuis plusieurs années pour des applications de navigation piétonne pour personnes déficientes visuelles. Cependant la précision du système GPS sans correction peut conduire à des erreurs de positionnement importantes, empêchant des vitesses élevées, voire pouvant rendre le système dangereux. Le système SBAS apporte une précision supplémentaire (c.f. Table 2.1) qui reste néanmoins limitée. Comme nous l'avons dit, SBAS n'est par ailleurs pas toujours disponible, en particulier en milieu urbain. Les systèmes GNSS multibandes, qui commencent tout juste à apparaître sur le marché offrent quant à eux une précision nettement améliorée. À titre d'exemple, nous avons comparé la précision des deux systèmes en Figure 2.2. Celle-ci montre les résultats obtenus sur un même lieu (milieu semi-urbain, ciel partiellement masqué) pendant 6 heures de mesures stationnaires avec deux types de récepteurs, un supportant GPS+EGNOS, l'autre étant un récepteur multibande GPS L1/L2 (U-Blox-ZED-F9P) sans support EGNOS. Le récepteur multibande permet alors d'atteindre une précision sub-métrique, avec une erreur maximale de 1.4 m.

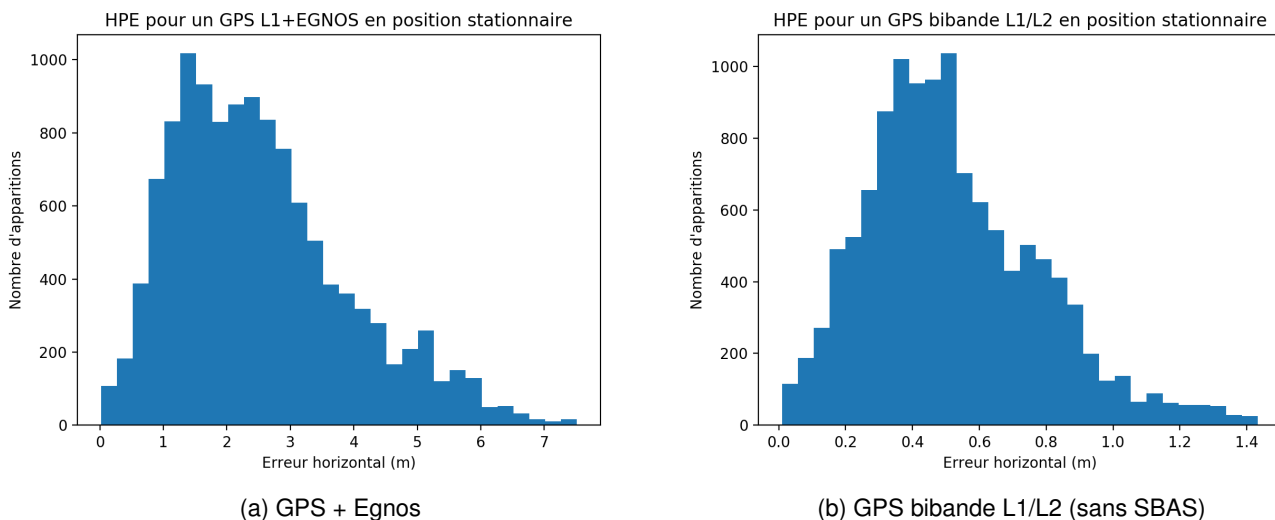
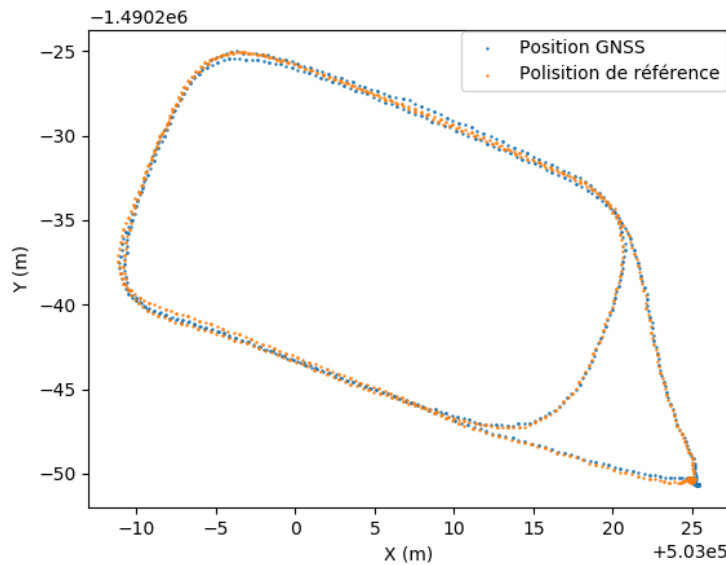


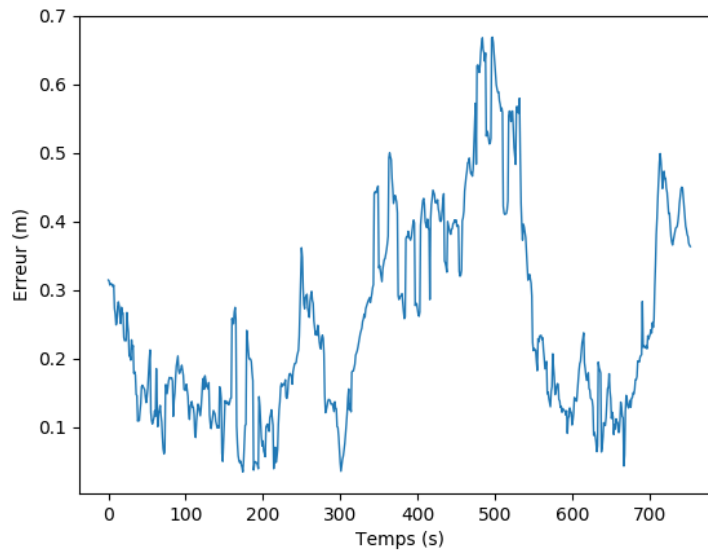
FIGURE 2.2 – Exemples d'erreurs de Précision Horizontale pour deux récepteurs en position stationnaire en milieu semi-urbain (Breuillet 91). L'erreur de précision (différence par rapport à la position moyennée sur un temps de plusieurs heures) est en moyenne très nettement inférieure sur la solution bi-bande. Par ailleurs l'erreur maximum est supérieure à 7m avec le GPS monobande, pour 1.4m avec le GPS en mode bi-bande.

Nous avons également testé le récepteur ZED-F9P dans le cas de la course à pied en terrain dégagé. Pour cela nous avons équipé le coureur de deux récepteurs : un récepteur GNSS-RTK, fournissant une trajectoire de référence et un récepteur GNSS-multibande (ZED-F9P). Les résultats sont visibles sur la Figure 2.3.

Dans ce cas encore nous avons obtenu des erreurs très raisonnables pour notre cas d'usage puisque qu'elles n'excèdent jamais ici 0.65m. Cela est certainement dû aux conditions favorables de l'expérience, en particulier le ciel qui est dégagé.



(a) Double boucle réalisée en conditions *open-sky* avec un récepteur GNSS dual-band et un récepteur RTK utilisé comme référence.



(b) Erreur (distance) en m lors du parcours ci-dessus.

FIGURE 2.3 – Erreur obtenue avec un récepteur GNSS multibande. Le parcours de référence est obtenu avec un GNSS RTK (solution fixée). Le parcours est effectué à la vitesse de la course lente (8 à 10km/h). Les mesures GNSS sont légèrement lissées pour gommer l'asynchronisme des mesures entre les deux récepteurs (moyenne mobile sur 0.5s). Les deux antennes, positionnées sur la tête du coureur sont éloignées de 10cm.

Compte tenu des performances obtenues, les technologies RTK-GNSS et GNSS multibandes nous semblent donc être des technologies tout à fait appropriées pour notre cas d'usage lorsqu'un signal GNSS de qualité est disponible (c'est-à-dire en *open-sky*). Malheureusement ces erreurs croissent de façon importante en environnement multichemin

ou lorsque le ciel est masqué, ce qui peut limiter l'utilisation de ces technologies pour le guidage fin en dehors des endroits dégagés (comme une piste d'athlétisme).

Type d'erreur	GPS	GPS + EGNOS
Orbite et synchronisation d'horloge	1 m	0,5 m
Erreur troposphérique	0,25 m	0,25 m
Erreur ionosphérique	2 m	0,3 m
UERE (Somme quadratique des erreurs - $1\sigma$ )	2,31 m	0,83 m
Erreur horizontal( $1\sigma$ )	2,54 m	0,92 m
Erreur horizontal( $2\sigma$ , 95 %)	5,08 m	1,84 m

TABLE 2.1 – Précision de systèmes de localisation par satellites (ordres de grandeurs typiques d'après [54] (CNES/ESA)). Ce tableau rapporte des erreurs typiques, qui peuvent varier sensiblement selon la situation. L'erreur horizontale est donnée pour un HDOP typique de 1.1 non modifié par le système SBAS. Les erreurs liées au multitrajét et au bruit du récepteur ne sont pas modifiées par le système SBAS.

## 2.2 Technologies pour la mesure de distances

Comme décrit précédemment, beaucoup de systèmes de localisation s'appuient sur des mesures de distances qui seront ensuite utilisés avec des méthodes de latéralisation. Ces techniques sont largement utilisées pour la localisation en intérieur. On peut distinguer des méthodes mesurant la distance à des objets dont l'emplacement est connu (télémétrie) et les méthodes mesurant des distances entre balises.

### 2.2.1 Télémétrie

Ces techniques utilisent généralement le *temps de vol* d'ondes acoustiques (sonar), lumineuses (Lidar) ou électromagnétiques (Radar) en utilisant la réflexion sur l'objet à localiser. Associés à des systèmes de balayage, elles permettent de scanner tout ou partie de l'espace sur deux ou trois dimensions. Plus que pour la latéralisation, elles sont souvent utilisées pour la détection d'obstacles, par exemple dans la *canne blanche laser* [74] ou dans le guidage avec évitement par sonar [12].

### 2.2.2 Technologies radiofréquence : Wi-fi, ZigBee, Bluetooth

Les réseaux sans fils WiFi (norme IEEE 802.11) sont largement présents dans nos environnements. Ils sont simples à déployer et peu coûteux. Il est donc séduisant d'utiliser des informations issues de ces réseaux pour se localiser. Les informations fournies sont à la fois de nature logique (identifiant de réseau) ou physique (niveau de puissance reçue). Les fréquences utilisées sont soit sur la bande des 2.4GHz ou sur celle de 5GHz et la portée typiquement de quelques dizaines de mètres. Lors de l'émergence de cette technologie au début des années 2000, plusieurs techniques de localisation utilisant la puissance du signal reçu comme indicateur de distance [133, 135] ont été développées. Compte tenu du manque de fiabilité des modèles RSS, des techniques utilisant le "fingerprinting" ont ensuite été envisagées [72, 238]. Plus récemment, on a vu apparaître des projets utilisant la mesure de l'effet Doppler induit par les déplacements entre les bornes [51].



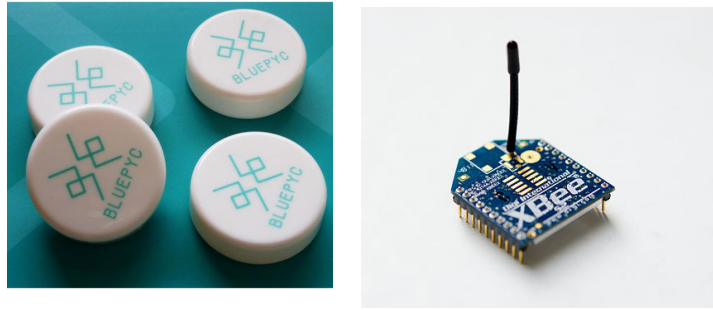


FIGURE 2.4 – Exemple de balises : (gauche) Bluetooth Low Energie et (droite) module ZigBee XBee.

D'autres technologies radiofréquences très comparables au WiFi peuvent être employées. Certains projets [27, 73, 237] ont utilisé ZigBee, un protocole radio à faible portée et très faible consommation que l'on envisage habituellement pour *l'internet des objets - IoT* ou la domotique.

La technologie *Bluetooth* (IEEE 802.15.1) est également très répandue et à très faible coût. Bluetooth utilise toujours les fréquences de la bande des 2.4GHz, mais a connu plusieurs évolutions majeures. La norme Bluetooth 4 (2010) a introduit une version à faible consommation (Bluetooth Low Energie, BLE) qui permet entre autre, de créer des balises (*Beacon*) diffusant des données à intervalles réguliers et alimentées par de simples piles. Ces balises sont largement utilisées dans des applications commerciales (Eddystone, AltBeacon, iBeacon, etc.) pour la localisation par proximité. Comme pour les autres technologies radio, l'intensité du signal reçu peut également être utilisée pour le positionnement temps réel [76, 236].

L'API Bluetooth des récepteurs permet généralement de récupérer la valeur en dBm du RSSI souvent codée sur un entier signé de 8-bit.

Si on considère le RSSI exprimé en dB, l'équation 1.7 peut s'écrire (modèle LNSM pour *Log-Normal Shadowing radio propagation Model*) :

$$\text{RSSI} = \left[ \frac{P_r(d_o)}{P_r(d)} \right]_{dB} = -10\alpha \log \frac{d}{d_o}. \quad (2.6)$$

Les figures 2.5 montrent deux types de mesures de RSSI en fonction de la distance que nous avons effectuées. Ces mesures ont été réalisées avec un utilisateur portant un émetteur Bluetooth en intérieur. Dans un premier cas, l'utilisateur se déplace sur un arc de cercle à distance constante du récepteur et faisant plusieurs fois demi-tour. L'émetteur est porté sur l'épaule et l'effet d'ombrage du corps et/ou les variations de polarisation des antennes sont clairement visibles. Dans le deuxième exemple, l'utilisateur effectue deux aller-retours s'éloignant et se rapprochant du récepteur à une vitesse la plus régulière possible.

Comme on peut le constater, la force du signal reçu contient un bruit complexe qui n'a généralement pas de distribution gaussienne et qui est susceptible d'affecter sérieusement la précision du positionnement. D'une manière générale, le RSSI est difficile à modéliser et très sensible à l'environnement (effet d'ombrage, réflexions, défaut de polarisation des antennes) ce qui rend son exploitation délicate pour une localisation précise.

### 2.2.3 UWB

L'Ultra Wide Band est une technologie radio qui utilise la transmission d'impulsions radio de très courte durée, la bande passante utilisée est alors beaucoup plus large que pour les techniques de modulation usuelles. En pratique,

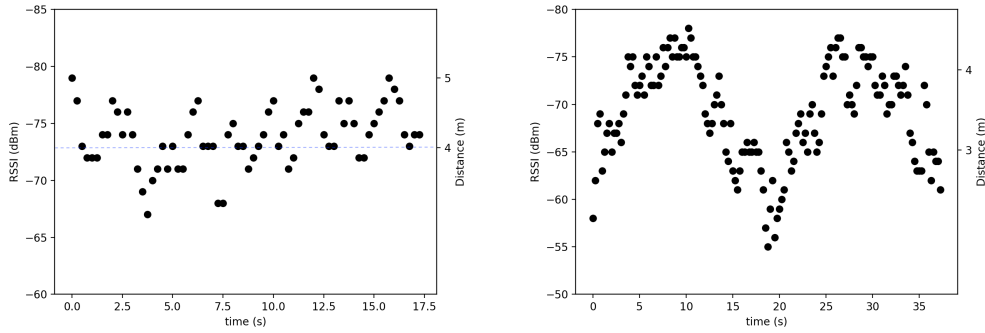


FIGURE 2.5 – Mesures RSSI Bluetooth. À gauche : déplacement latéraux en conservant une distance constante (4m). À droite : l'utilisateur fait deux aller-retours, partant à 1.5m du récepteur et s'en éloignant jusqu'à 5m, en conservant une vitesse approximativement constante.

lorsque l'on considère les fréquences basse et haute  $f_l$  et  $f_h$  du spectre à -10dB d'un signal radio, et  $f_c$  sa fréquence centrale, la largeur de bande relative (FBR pour *Fractional Bandwidth*) d'un signal UWB sera supérieure à 20% de la fréquence centrale.

$$FBR_{UWB} = \frac{f_h - f_l}{f_c} > 0.2. \tag{2.7}$$

Ceci distingue les signaux UWB des signaux *classiques*, à bande étroite (Narrowband), pour lesquels la FBR est généralement inférieure à 0.1%.

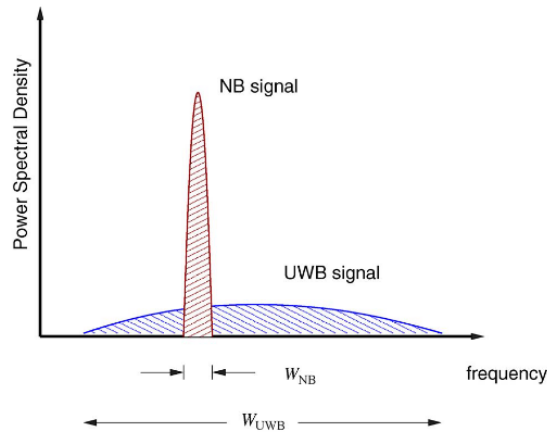


FIGURE 2.6 – Spectre de fréquence UWB vs NarrowBand (NB) (Extrait de [52]).

À la différence des systèmes de transmission radio conventionnels qui transmettent l'information par modulation d'intensité, de fréquence ou de phase, la technologie UWB transmet l'information en envoyant des impulsions à différents intervalles de temps, on parle alors de *Modulation en Position d'Impulsion* (PPM pour *pulse-position modulation*).

Ainsi, l'UWB permet un débit de données important avec une puissance d'émission faible [145]. Cette propriété peut s'expliquer grâce au théorème de Shannon-Hartley (c.f. eq (2.8)) qui montre que la capacité d'un canal de communi-

tion analogique noté  $C$  soumis à un bruit blanc gaussien additif de puissance  $N$  augmente d'une manière linéaire avec la largeur de bande  $B$  alors qu'elle augmente d'une manière logarithmique avec la puissance  $S$

$$C = B \log_2 \left( 1 + \frac{S}{N} \right) . \tag{2.8}$$

L'UWB est particulièrement bien adapté aux mesures des distances par *temps de vol*. En effet, l'UWB permet une grande précision temporelle en étant moins sensible aux multi-trajets. Avec ses impulsions très courtes, cette technologie permet une meilleure séparation temporelle des différents trajets qui sont alors détectables et identifiables. Par ailleurs, avec la transmission sur un large spectre de fréquences on peut supposer que certaines fréquences seront moins affectées par les obstacles et auront même une trajectoire en visibilité directe. Dans les applications de localisation, les principaux algorithmes utilisés sont alors TDoA et TWR.

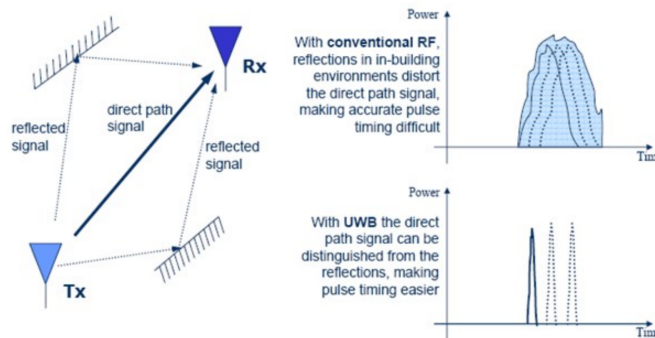


FIGURE 2.7 – Illustration de la séparation temporelle de signaux UWB (extrait de [245]).

Les technologies UWB sont utilisées depuis les années 60 mais étaient restées cantonnées aux applications radar pour des raisons réglementaires (utilisation de larges bandes de fréquences) et des difficultés techniques. Cette technologie connaît un développement très important depuis que le régulateur américain des télécom (FCC, Federal Communications Commission) à permis en 2002 l'utilisation des bandes de fréquences 3.1 - 10.6 GHz sans licence [218]. En outre les évolutions technologiques ont permis une baisse sensible des coûts et même la miniaturisation sous forme de puces uniques incluant toutes les fonctions d'un "transceiver" UWB (émetteur/récepteur, horloge de précision, modulation, interface numérique). Un exemple est visible Fig. 2.8.

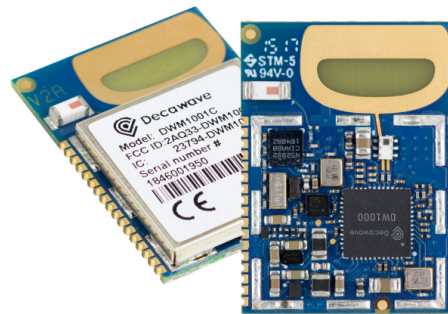


FIGURE 2.8 – Module transceiver UWB modèle DWM1001C (Decawave). On peut apercevoir l'antenne sous forme de piste de cuivre, sa forme atypique lui permet d'être efficace sur une large bande de fréquence.

La mesure de la distance  $d_{AT}$  entre une paire de balises UWB peut se modéliser sous la forme :

$$d_{AT} = \hat{d} + b_{AT} + \eta \quad (2.9)$$

où  $\hat{d}$  est la distance réelle,  $b_{AT}$  est un biais de mesure dépendant de la paire de transceivers utilisée et  $\eta$  est un bruit gaussien centré en zéro.

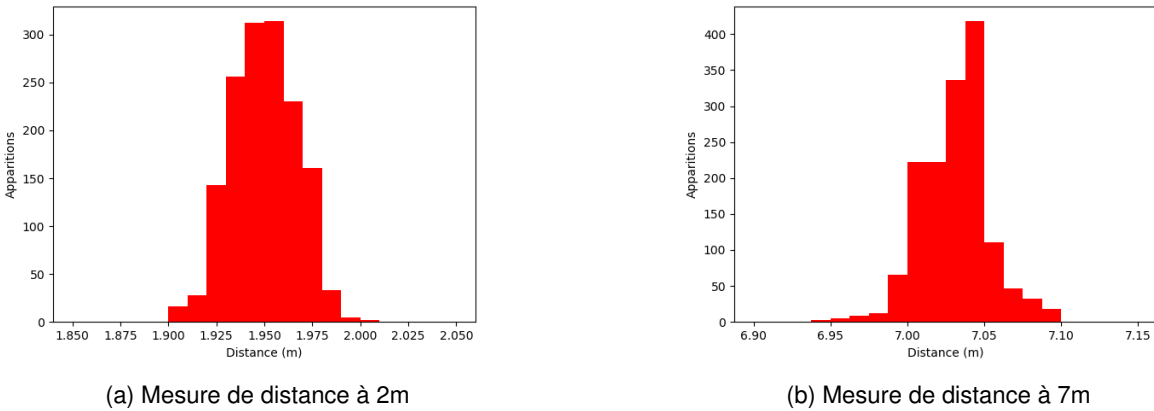


FIGURE 2.9 – Bruit de mesure pour une paire de balises éloignées de 2m et 7m.

Le biais, lié au matériel (en particulier délais d'antennes), est propre à la paire de transceivers. Il peut être supposé stable dans le temps et peu dépendant de la distance (dans certaines limites de distances d'utilisation). Une phase de calibration pourra alors permettre de l'estimer pour l'éliminer. Le bruit lui est peu dépendant de la distance contrairement aux technologies RSS [105]. Les ordres de grandeurs des biais sont de l'ordre de 10-30cm sur les modules que nous utilisons ce qui est de nature à dégrader la précision de localisation. Plusieurs procédures de calibration ont été proposées pour les estimer :

*Decawave* propose une méthode [60] où l'on place trois balises équidistantes et on cherche à minimiser la norme des différences entre les matrices de distances euclidiennes réelles ( $EDM_{réelle}$ ) et mesurée ( $EDM_{mesurée}$ )

$$\min \|EDM_{réelle} - EDM_{mesurée}\|. \quad (2.10)$$

*Time Domain* (un autre constructeur de matériel UWB) propose une méthode de calibration comparable [214] consistant à déterminer les délais d'antenne par méthode des moindres carrés.

La figure 2.9 représente une expérience de mesures que nous avons menée entre une paire de transceivers UWB *Decawave DW1000* distantes de 2m et de 7m respectivement (en vue directe en intérieur). On donne l'histogramme de répartition de 800 mesures qui ont été effectuées.

On constate que le bruit de mesure reste faible :  $\sigma = 1.7cm$  à 2m (biais  $-5, 5cm$ ) et  $\sigma = 2.3cm$  (biais  $3.1cm$ ) à 7m et si on élimine (facilement) quelques mesures aberrantes, 100% des mesures étaient dans un intervalle de 16cm.

## 2.3 Centrale inertielle, capteur d'orientation

Les centrales inertielles, souvent désignées par l'acronyme anglais IMU (Inertial Motion Unit), sont des dispositifs électroniques qui permettent à la fois de mesurer l'accélération linéaire, la vitesse angulaire et souvent le vecteur champ magnétique local. Ces capteurs sont utilisés de longue date pour la navigation à l'estime dans les domaines naval et aérospatial (le missile balistique V2 de la 2ème guerre mondiale possédait déjà un gyroscope). Il existe plusieurs classes d'IMU correspondant à différents types d'usage, leurs coûts variant de 3 à plusieurs centaines de milliers d'euros suivant la catégorie (c.f. Table 2.2).

Type	Biais de l'accéléromètre (mg)	Biais du gyro (deg/s)	Coût (euro)
Marine	0.01	0.001	1 000 000
Aviation	0.03-0.1	0.01	100 000
Pro	1-10	1-100	1000
low-cost/smartphone	>10	>100	3-100

TABLE 2.2 – Caractéristiques des différentes catégories d'IMUs (d'après [94])

Dans les applications portables pour le guidage des personnes, nous nous intéressons uniquement aux IMU utilisant des capteurs MEMS (Micro-Electro-Mechanical-Systems). Ils sont constitués de micro-structures mécaniques encapsulées dans une ou plusieurs puces électroniques de très petite taille. Ces capteurs sont très petits, légers (leur poids est inférieur au gramme), et de faible coût. Ils offrent cependant qu'une précision limitée.

Les IMU à base de MEMS peuvent être utilisées pour les applications de navigation à l'estime pour les piétons, par exemple pour compter ou estimer la longueur de pas [186, 189, 207, 223], mais ils sont également très utiles pour l'estimation de l'orientation (attitude). Dans ce cas il peuvent être associés à d'autres technologies d'odométrie dans des dispositifs de positionnement [77, 136].

### 2.3.1 Les différents capteurs : modélisation et calibration

**L'accéléromètre** mesure l'accélération subie par le capteur ainsi que la gravité (un champ de gravitation est localement équivalent à une accélération uniforme). La mesure de la gravité permet de connaître l'orientation du plan du capteur par rapport à la surface de la terre (Pitch et Roll).

Les accéléromètres MEMS sont des capteurs qui fournissent des données à un débit élevé (de l'ordre du KHz). Ils sont par contre affectés d'un bruit de mesure  $n_a$  et d'un biais de mesure  $b_a$  lié à la fois au capteur et aux conditions extérieures. Les erreurs dans la mesure sont multiples et peuvent être décomposées ainsi :

1. les facteurs d'échelle des différences de sensibilité entre les axes. On peut représenter ces différences par une matrice diagonale  $S = \text{diag}(S_x, S_y, S_z)$ ,
2. les défaut d'orthogonalité des axes du capteur  $N = \begin{bmatrix} n_x & n_y & n_z \end{bmatrix}$ ,
3. les offsets de mesure  $b_a = \begin{bmatrix} b_{ax} & b_{ay} & b_{az} \end{bmatrix}^T$ .

Ces différents types d'erreurs sont représentés par la figure 2.10.

Ainsi, un modèle d'accéléromètre pourrait s'écrire

$$a = S N (\hat{a}_l - a_g) + b_a + n_a \quad (2.11)$$

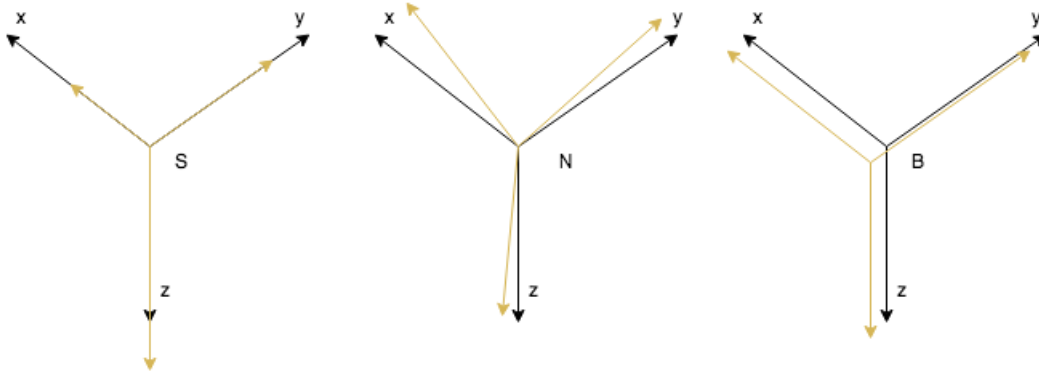


FIGURE 2.10 – Erreurs d'instrumentation sur les capteurs inertiels.

ou  $a$  est l'accélération mesurée par le capteur,  $\hat{a}_l$  est l'accélération idéale que l'on cherche à déduire,  $a_g$  est l'accélération de la gravité et  $n_a$  est un bruit gaussien centré en 0, et  $S$  et  $N$  les matrices données aux points 1 et 2 précédents respectivement.

Le biais  $b_a$  est souvent estimé en positionnant le capteur dans des positions connues et en moyennant le résultat. Il est également possible d'aligner chaque axe de l'accéléromètre avec la gravité et de faire la moyenne des lectures pour fournir à la fois les biais de décalage et la correction d'échelle des axes.

Cependant cette mesure de calibration est imparfaite, car elle peut malheureusement évoluer dans le temps. En particulier la température impacte fortement le biais de mesure [69] et chaque axe du capteur peut être affecté différemment.

**Le gyroscope** mesure le vecteur de vitesse angulaire.

Comme l'accéléromètre, ce capteur permet généralement des mesures rapides. Il est sujet à une dérive lente et variable  $b_g$  (bruit basse fréquence) de l'ordre de 1-2 °/minutes pour les capteurs à bas coût. On peut écrire un modèle de gyroscope sous la forme :

$$\omega = \hat{\omega} + \mathbf{b}_g + \mathbf{n}_g \quad (2.12)$$

où  $\omega$  est la vitesse angulaire mesurée,  $\hat{\omega}$  est la vitesse angulaire réelle que l'on cherche à estimer et  $\mathbf{n}_g$  est un bruit blanc gaussien.

Le biais  $\mathbf{b}_g$  peut être estimé en plaçant le capteur dans un état immobile et en mesurant la dérive pendant quelques instants. Sur les capteurs MEMS à bas coût, ce biais n'est pas constant dans le temps et cette estimation ne peut être considérée comme fiable.

**Le magnétomètre** mesure le champ magnétique terrestre (et donc l'orientation par rapport au nord magnétique).

Sur les capteurs MEMS à bas coût, cette mesure est souvent relativement lente (typiquement 10 fois plus lente que pour les deux précédents), et liée à l'environnement extérieur.

Les mesures du magnétomètre sont, comme pour l'accéléromètre, sujets à des erreurs liées aux facteurs d'échelle entre les axes  $S$ , aux défauts d'orthogonalité des axes du capteur  $N$ , et aux offsets de mesure. À cela s'ajoute les

interférences d'origine magnétiques liées à l'environnement :

1. les champs magnétiques permanents liés au capteur ou les phénomènes d'hystérésis magnétiques, généralement appelé "hard iron". Ces champs magnétiques ne sont pas dépendants de l'orientation du capteur et peuvent être assimilés à un offset que l'on notera  $b_{hi}$  ;
2. les interactions complexes avec les matériaux ferromagnétiques locaux, qui peuvent être dépendants de l'orientation par rapport au champ magnétique local et généralement appelé "soft iron". Ils peuvent être modélisés sous la forme d'une matrice

$$A_{si} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} .$$

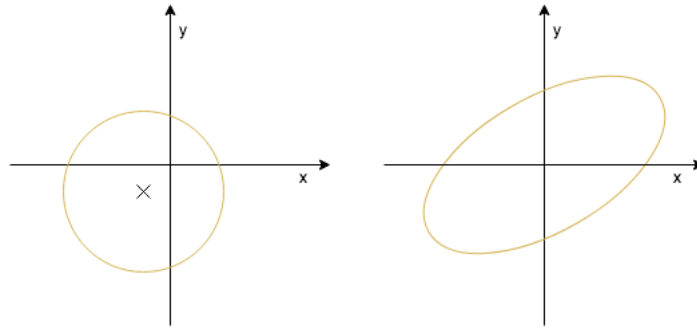


FIGURE 2.11 – Représentation symbolique des effets de fer dur (gauche) et de fer doux (droite) en 2D. Le champ magnétique est soit décentré (hard iron) soit déformé (soft iron).

Le champ magnétique  $h_m$  mesuré peut alors s'écrire sous la forme

$$h_m = S N (A_{si} h + b_{hi}) + b_{so} + n_m , \quad (2.13)$$

expression dans laquelle  $h$  est le champ magnétique local dans le repère du capteur et  $n_m$  le bruit de mesure supposé gaussien. Si on note  $\hat{A}$  une estimation de  $S N A_{si}$  et  $\hat{b}$  une estimation de  $S N b_{hi} + b_{so}$ , alors on peut calculer une mesure post-calibration  $h_n$

$$h_n = \hat{A}^{-1}(h_m - \hat{b}) . \quad (2.14)$$

Dans les conditions idéales, lors des rotations du capteur, les mesures devraient se trouver sur une sphère dont le rayon correspond à la magnitude du champ magnétique terrestre local. On peut montrer (voir par exemple [192]) que les mesures  $h_m$  sont positionnées sur un ellipsoïde, ainsi le processus de calibration s'apparente à un problème de fitting d'ellipsoïde. Plusieurs algorithmes ont ainsi été proposés pour résoudre ce problème ([131, 192, 219]).

Malheureusement cette approche peut être complexe à implémenter sur une électronique embarquée. Une méthode plus simple consiste à centrer l'ellipsoïde sur l'origine et à mettre à l'échelle les réponses axiales [234].

À titre d'illustration, nous donnons sur la figure 2.12 les mesures que nous avons effectuées sur le magnétomètre d'un IMU Inversense MPU9250 collectées en effectuant de larges mouvements sur tous les axes. Sur les données brutes, on visualise aisément le biais de *fer dur* caractérisé par un décentrage de l'ellipsoïde. La première figure montre le

résultat d'une calibration simple de type centrage/scaling alors que la seconde figure montre une calibration par fitting d'ellipsoïde tel que décrit en [131].

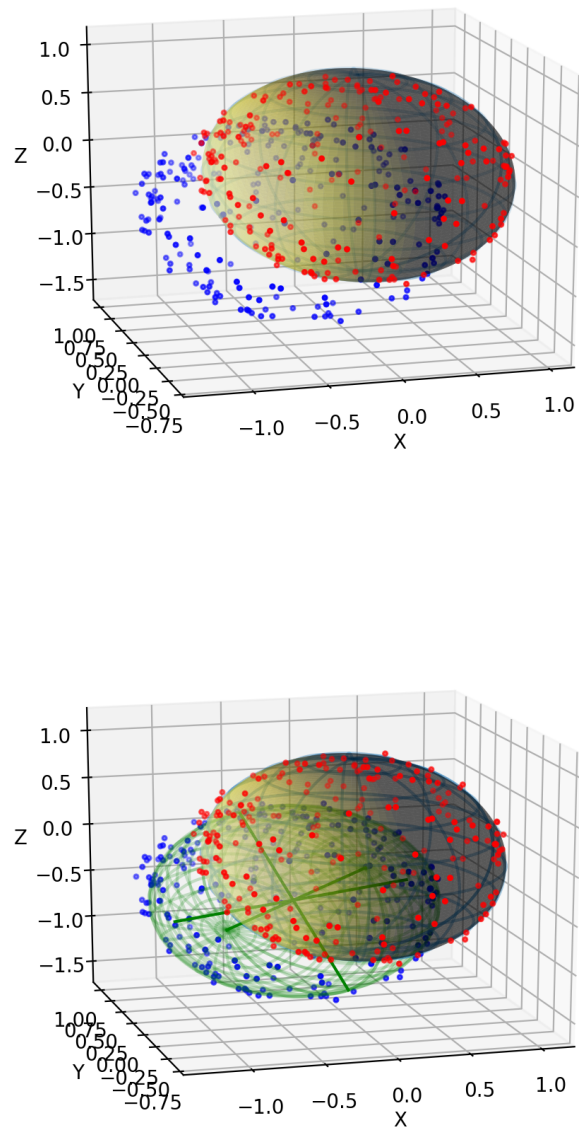


FIGURE 2.12 – Échantillons de données bruts (bleu) collectées par un capteur MPU9250. En haut : Mise à l'échelle et centrage pas la méthode des min-max, En bas : Fitting sur un ellipsoïde (verte), et ajustement sur la sphère unité



### 2.3.2 Estimation de l'orientation

Afin d'obtenir une estimation fiable de l'orientation, les données de ces trois capteurs doivent ainsi être fusionnées. Après fusion des données on parle généralement de système AHRS (Attitude and Heading Reference System). Les méthodes typiques de fusion de données sont :

- La combinaison d'un filtre passe-haut sur les données du gyroscope et passe-bas sur les données de l'accéléromètre [194] souvent appelé *filtres complémentaires*. Si on considère les mesures (calibrées) du gyroscope  $\omega = [\omega_x, \omega_y, \omega_z]^T$  et de l'accéléromètre  $\mathbf{a} = [a_x, a_y, a_z]^T$ . On peut écrire les deux filtres respectivement sous les formes

$$\hat{a}_{t+dt} = (1 - \alpha)a_{t+dt} + \alpha\hat{a}_t$$

et

$$\hat{\omega}_{t+dt} = (1 - \alpha)\hat{\omega}_t + (1 - \alpha)(\omega_{t+dt} - \omega_t)$$

où  $t$  est le temps d'échantillonnage et  $\alpha$  est un paramètre du filtre (une constante généralement supérieure à 0.5 liée à la confiance dans la mesure de l'accéléromètre).

Finalement la mesure de l'angle peut s'écrire :

$$Ang_{t+dt} = (1 - \alpha)(Ang_t + \omega_{t+dt}dt) + \alpha a_{t+dt}.$$

- L'utilisation d'un filtre de Kalman [160]. Une approche du problème consiste à considérer le biais du gyroscope (dans chaque orientation) dans la variable d'état en plus des angles à estimer. Le filtrage de Kalman permet d'estimer dynamiquement le biais du gyroscope et, de ce fait, fournit une estimation supérieure au filtre complémentaire [110]. Cependant, il nécessite une puissance de calcul souvent trop importante pour les petits systèmes embarqués et, de ce fait, reste peu utilisé.
- L'utilisation de méthode de filtrage spécialisée. Le filtre de Mahony [71] est un filtre complémentaire utilisant un contrôleur proportionnel integral (PI) pour estimer le biais du gyroscope en temps réel et fournir une bonne estimation de l'attitude. Le filtre de Magdwick [153] est lui aussi dérivé d'un filtre complémentaire, il utilise un algorithme de descente de gradient et une représentation en quaternions. Ce type de filtre permet une précision de l'estimation comparable au filtre de Kalman mais avec un coût machine bien plus faible. Le filtre de Magdwick est aujourd'hui très largement utilisé pour les petites centrale inertielle disposant de peu de ressources (typiquement pour la stabilisation des drones).

## 2.4 Autres techniques, empreintes

Par souci d'exhaustivité, on peut mentionner d'autres systèmes utilisant les empreintes de signaux. Par exemple, la technologie d'empreintes magnétiques utilise les variations locales du champ magnétique. Ces variations peuvent être liées à la structure des bâtiments et des installations électriques. Cette méthode cherche à tirer partie de ces anomalies dans le champ magnétique pour déduire une position ([93, 101]). Elle présente l'avantage d'être totalement passive puisqu'aucune installation n'est nécessaire. Il semble cependant qu'elle comporte certaines limites (relevées par [144]) en particulier en raison des interférences liées aux activités humaines (déplacement d'objets, ascenseurs, etc.).

## Chapitre 3

# Vers un système de localisation pour le guidage des sportifs non voyants

La localisation des personnes en intérieur comme en extérieur est un sujet de recherche très actif qui a permis l'essor de dispositifs commerciaux. Cependant peu d'entre eux sont capables de répondre à notre cahier des charges pour une application au guidage des personnes DV. Nous pourrions résumer nos besoins de la façon suivante :

- Localisation en intérieur et en extérieur (pas nécessairement avec un unique dispositif) ;
- Précision décimétrique ;
- Estimation rapide (plusieurs fois par seconde) ;
- Rayon d'action minimum d'environ 30-40 m pour un fonctionnement dans un gymnase ;
- Faible coût ( $\ll$  1000 euros).

Les GNSS différentiels ou multibandes sont des technologies qui satisfont pratiquement tous ces critères, à l'exception du prix et de la capacité à fonctionner en intérieur. On peut donc estimer qu'ils peuvent répondre aux besoins pour une utilisation en extérieur. Nous nous focalisons donc maintenant sur les technologies destinées à la localisation en intérieur.

### 3.1 Localisation par réseau de capteurs UWB

Comme évoqué plus haut, les réseaux de capteurs associés à des méthodes de latération sont très répandus pour la localisation en intérieur. Dès 1992, le système Active Badge [225] utilisait un réseau des capteurs infra-rouge pour localiser un nœud mobile en intérieur. Par la suite de nombreux autres systèmes reprenant ce concept ont été proposés comme le système Bats (1999) [102] qui utilise un réseau de capteur Ultra-son ou la localisation à l'aide de la puissance des signaux radiofréquence.

En revanche, les caractéristiques prometteuses des systèmes Ultra WideBand nous ont conduit à développer un module de localisation en intérieur s'appuyant sur cette technologie. L'Ultra WideBand est en effet une technologie de choix pour la localisation temps réel en intérieur. Associé à un algorithme d'estimation de distance (TWR, TDOA) aujourd'hui directement géré par le matériel, elle permet de mesurer des distances de façon fiable et rapide. Par exemple les tranciever DM1000 de la société Decawave permettent d'accéder à des précisions de l'ordre de 10cm (données constructeur). Lors du démarrage de ce projet de recherche, en 2015, en l'absence de solution industrielle abordable

et interfacable, nous avons été amenés à développer notre propre solution de positionnement par balise UWB, à la fois pour certains aspects matériels et surtout concernant le logiciel de localisation temps-réel. L'aspect matériel ne sera pas développé dans ce document, nous nous sommes principalement appuyés sur le projet *arduino-dwm1000* qui permet d'interfacer les transceivers Decawave sur microcontrôleur via l'API arduino. La bibliothèque [217] fournit toute l'interface de communication bas-niveau avec les modules Decawave permettant le développement d'application de *ranging par TWR*.

### 3.1.1 Positionnement temps réel par filtre de Kalman et calibration par méthode des moindres carrés (LSE - *Least square Estimation*)

Dans une installation typique de localisation par multilatération, les balises sont placées à des coordonnées connues dans un repère cartésien (voir fig. 3.1). Ce type d'installation nécessite un positionnement précis qui peut être délicat (il s'agit typiquement d'effectuer des mesures sur plusieurs dizaines de mètres) et fastidieux en particulier pour une installation mobile ou temporaire.

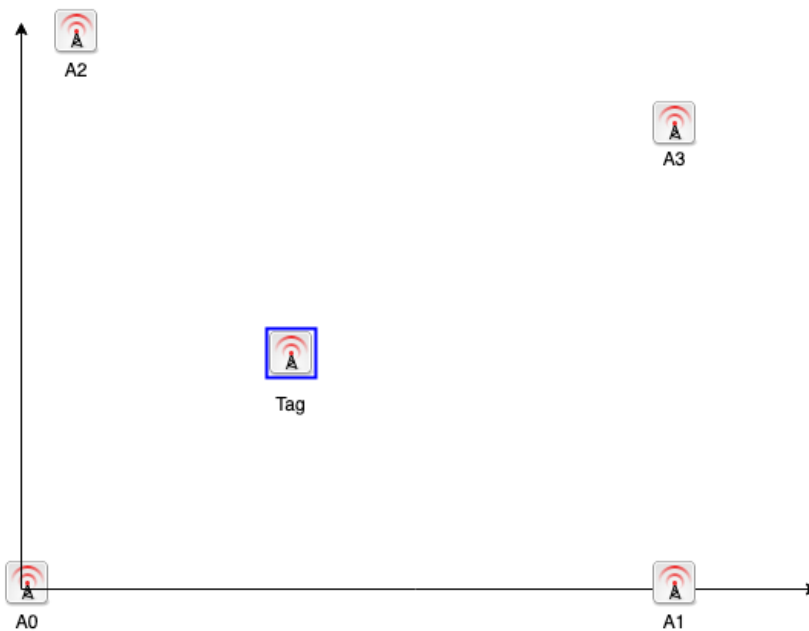


FIGURE 3.1 – Installation typique d'un système UWB en 2D. Les quatre balises extérieures doivent être positionnées précisément afin de pouvoir estimer la position du Rover (Tag).

Dans le domaine de la localisation par réseau de balises il est d'usage d'appeler *ancres* les stations de base (BS) fixes et *tag* le transceiver placé sur l'objet mobile (OM) à localiser. Nous utiliserons cette terminologie par la suite.

Par ailleurs, comme présenté au chapitre précédent, les mesures de chaque (paire de) transceiver sont affectées d'un biais que l'on peut supposer fixe dans nos conditions d'usage (voir l'équation (2.9)). On peut donc envisager d'estimer ce biais de mesure associé à chaque balise si on souhaite obtenir des estimations de distance précises. Résoudre ce double problème (positionnement des balises et estimation du biais) constitue donc un enjeu important pour l'utilisation de ces technologies. Une procédure de calibration utilisant les méthodes LSE [105] est alors envisageable.

Nous proposons donc une méthode de localisation qui fonctionnera en deux phases :

1. une phase préliminaire de localisation des balises par méthode LSE ;
2. une phase de localisation temps-réel de l'objet mobile par filtrage de Kalman.

### 3.1.2 Calibration

Pour évaluer la position des balises de référence nous utilisons une méthode de minimisation par moindres carrés décrite ci-dessous.

Considérons  $m$  balises fixes numérotés de 1 à  $m$ . Nous considérons la première balise comme l'origine du repère et la seconde définira l'axe des abscisses. Dans ce repère, les balises sont situées aux coordonnées  $a_i = (a_{x,i}, a_{y,i})$ . Effectuons  $n$  mesures de distance par rapport aux balises, d'un objet mobile situé (à l'instant  $j$ ) aux coordonnées  $(t_x^j, t_y^j)$ , pour  $j$  allant de 1 à  $n$ . En appelant  $d_i^j$  la distance entre l'objet mobile et la  $i$ -ème balise à l'instant  $j$ , nous pouvons écrire le système d'équations quadratiques surdéterminé suivant

$$(d_i^j)^2 = (t_x^j - a_{x,i})^2 + (t_y^j - a_{y,i})^2, \text{ pour } 1 \leq j \leq n \text{ et } 1 \leq i \leq m. \quad (3.1)$$

Comme  $a_{x,1} = a_{y,1} = a_{x,2} = 0$ , il s'agit d'un système de  $m * n$  équations et  $2(m + n) - 3$  inconnues.

En pratique, cette phase consiste à se déplacer dans l'environnement proche des balises et à effectuer des mesures pendant quelques dizaines de secondes. À l'issue de ces mesures, une méthode LSE permet d'estimer la position des balises. La question du parcours à réaliser pour une convergence optimale de ce type de méthodes est aujourd'hui un sujet mal connu [59]. On considérera qu'il faudra se déplacer sur une surface importante, de façon non rectiligne et s'approchant successivement de chacune des balises.

Chercher une approximation du système (3.1) permet d'obtenir une estimation de la position de l'objet mobile mais aussi des balises fixes à chaque mesure. Il s'agit d'un problème d'optimisation non linéaire pour lequel il existe des algorithmes standards comme les méthodes de type Newton ou les méthodes de régions de confiance. Notre implémentation en Python repose sur la méthode *Trust Region Reflective* décrite dans [34] qui est implémentée dans Scipy.

Les méthodes LSE étant très sensibles aux aberrations de mesures (*outliers*), on procède au préalable à un filtrage qui permettra de les éliminer. Pour cela, pour chaque mesure on va comparer la mesure à une médiane glissante. Chaque mesure dont la différence à la médiane mobile dépassera un seuil fixé empiriquement (et qui dépend de la vitesse de déplacement du sujet) sera préalablement éliminée.

Enfin, si l'on intègre les biais de mesure dans le système (3.1) on est aussi capable d'évaluer simultanément les biais de mesure et la position des balises.

D'un point de vue pratique, la mise en place des balises (par exemple dans un gymnase) doit pouvoir se faire facilement et rapidement. Les mesures de distances sont donc proscrites. En revanche, il est souvent aisé de placer les balises aux quatre coins d'un rectangle grâce au marquage au sol. Nous avons donc prévu la possibilité de tenir compte de ce type de géométries dans notre algorithme.

À titre d'exemple, un parcours semi-circulaire du tag est réalisé Figure 3.2. Pour tester la robustesse de la méthode, la simulation est effectuée avec un bruit additif gaussien d'écart type  $\sigma = 1m$  qui est sensiblement plus important que dans les cas d'usage typiques de cette technologie.

Nous avons simulé la méthode de calibration dans les différents cas d'usage. Pour chaque cas (géométrie libre ou rectangulaire, biais connu ou inconnu) et différentes quantités de données collectées la figure 3.3 montre l'évolution de la précision moyenne (moyennes sur 50 simulations par cas). La métrique d'erreur telle que définie en (3.2) est le cumul

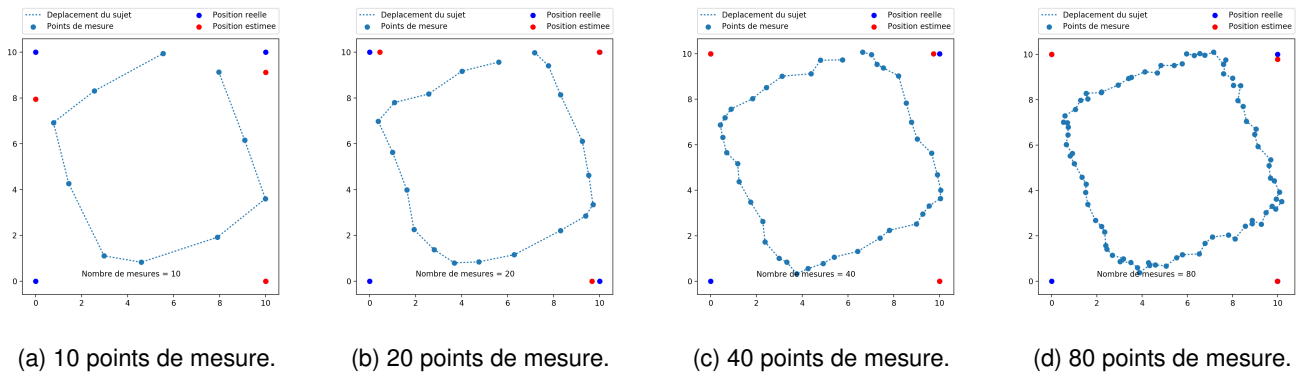


FIGURE 3.2 – Exemple d'évolution de l'estimation de la position des ancrs suivant le nombre de mesures effectuées. Cas de mesures fortement bruitées (bruit gaussien,  $\sigma = 1m$ ).

des distances euclidiennes entre les positions trouvées et les positions attendues pour les trois ancrs dont la position est inconnue.

$$\text{Erreur\_moyenne\_cumulee} = \frac{\sum \sqrt{(X_{\text{reel}} - X_{\text{estime}})^2 + (Y_{\text{reel}} - Y_{\text{estime}})^2}}{\text{Nombre d'estimations}}. \quad (3.2)$$

Dans le cas d'un biais connu, l'estimation converge rapidement avec un avantage sensible dans le cas d'une géométrie fixée (rectangulaire). Pour 160 mesures, l'erreur cumulée est de 12cm (max 28cm).

Prendre plus de mesures implique de se déplacer plus longuement dans la zone, mais nécessite également un temps de calcul qui devient rapidement grand par rapport au temps d'acquisition des données. Sur une machine performante de type PC moderne, les temps de calcul sont de l'ordre de 10s pour 80 mesures et 40s pour 160.

L'intérêt de la méthode pour l'estimation simultanée des biais et des positions est discutable en particulier pour une géométrie libre. Dans le cas d'une géométrie fixée, l'erreur moyenne cumulée tombe à environ 30cm (max 81cm) pour 640 mesures ce qui peut présenter un réel intérêt pratique, mais le temps de collecte de données comme le temps de calcul (temps moyen de 420 secondes sur PC) est pénalisant.

Ces résultats conduisent à l'élaboration d'une stratégie qui consiste à évaluer la position des ancrs avec un nombre de points limité avant de commencer le filtrage temps réel. Au fil de la collecte des données, de nouvelles mesures peuvent être prise en compte pour affiner l'estimation. Cette stratégie présente un intérêt pratique réel, mais nécessite de s'assurer que la tag n'est pas fixe pendant la seconde phase de la calibration.

### 3.1.3 Filtrage temps réel

Une fois que les positions des ancrs sont déterminées, on peut entamer la phase de localisation temps-réel (2). Pour cela un filtre de Kalman permettra tout à la fois d'effectuer la trilatération et le filtrage des données en utilisant directement les distances dans la fonction de mesure du filtre.

Dans les méthode LSE utilisées ci-dessus, aucune information préalable concernant la dynamique du système n'est utilisée pour l'estimation. Les techniques de filtrage bayésien, tel que le filtre de Kalman, consistent à estimer l'état des variables d'un système dynamique sujet à des perturbations et des observations partielles et bruitées en utilisant un modèle lié à la connaissance des déplacements. Dans le cas de modèles linéaires de bruit normalement distribués, le filtre de Kalman fournit l'estimateur optimal.

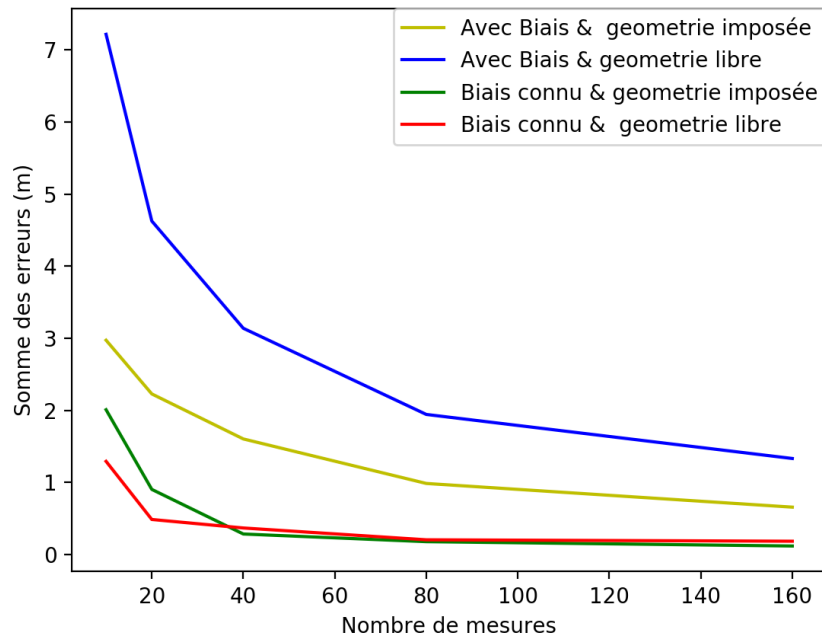


FIGURE 3.3 – Précision de l'estimation de la position des ancrs dans différents cas d'usage ( $\sigma = 0.2m$ ).

Les filtres de Kalman permettent d'estimer l'état d'un système dynamique linéaire bruité :

$$\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{B}_k \mathbf{u}_k + \mathbf{w}_k \quad (3.3)$$

où  $\mathbf{x}_k$  est l'état à l'instant  $k$ ,  $\mathbf{F}_k$  est la matrice de transition d'état du modèle,  $\mathbf{u}_k$  et  $\mathbf{B}_k$  respectivement le vecteur de contrôle et la matrice du modèle de contrôle et  $\mathbf{w}_k$  le bruit supposé gaussien et de moyenne nulle (de covariance  $Q$  tel que  $\mathbf{w}_k \sim \mathcal{N}(0, Q)$ ).

Les mesures sont liées linéairement à l'état du système, et bruitées :

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k \quad (3.4)$$

où  $\mathbf{z}_k$  est le vecteur contenant les mesures effectuées à l'instant  $k$ ,  $\mathbf{H}_k$  est la matrice de mesure, et  $\mathbf{v}_k$  le bruit de mesure de covariance  $R$  ( $\mathbf{v}_k \sim \mathcal{N}(0, R)$ )

Le filtre de Kalman est un estimateur qui fonctionne en deux phases, une étape de **prédiction** qui utilise l'état prédit à l'instant précédent pour estimer l'état courant à l'instant  $k$  et une étape de **mise à jour** qui prend en compte l'observation. Le fonctionnement de ce type de filtre est rappelé en annexe B.

Concevoir un filtre de Kalman consiste donc typiquement à déterminer les matrices  $\mathbf{F}$ ,  $\mathbf{B}$  et  $\mathbf{H}$  quand  $Q$  et  $R$  sont utilisés comme des paramètres souvent déterminés de façon empiriques suivant les caractéristiques des capteurs et les aspects non pris en compte par le modèle.

Les modèles *de position seule* (P), *position-vitesse* (PV) et *position-vitesse-accélération* (PVA) utilisant différentes

variables d'état sont généralement discutés et choisis en fonction de la dynamique des objet à suivre [202, 228]. Il est délicat de modéliser les mouvement d'une personne se déplaçant, mais on peut supposer une certaine régularité dans les mouvements, tout particulièrement dans les cas d'usage qui nous intéressent (roller, course à pied). Par la suite nous considérerons les modèles PV et PVA qui sont le plus souvent utilisés dans ce cas d'usage.

Le tableau ci-dessous définit les vecteurs d'états et les matrices de transition d'état en 2D pour les deux modèles étudiés :

	Modèle PV	Modèle PVA
Vecteur d'état	$\mathbf{x} = \begin{pmatrix} x \\ y \\ \dot{x} \\ \dot{y} \end{pmatrix}$	$\mathbf{x} = \begin{pmatrix} x \\ y \\ \dot{x} \\ \dot{y} \\ \ddot{x} \\ \ddot{y} \end{pmatrix}$
Matrice de transition	$\mathbf{F} = \begin{pmatrix} 1 & 0 & dT & 0 \\ 0 & 1 & 0 & dT \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$	$\mathbf{F} = \begin{pmatrix} 1 & 0 & dT & 0 & 1/2dT^2 & 0 \\ 0 & 1 & 0 & dT & 0 & 1/2dT^2 \\ 0 & 0 & 1 & 0 & dT & 0 \\ 0 & 0 & 0 & 1 & 0 & dT \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

TABLE 3.1

On pourra alors écrire la fonction de mesure sous la forme :

$$\mathbf{H}(X) = \begin{pmatrix} \sqrt{(x - a_{x,1})^2 + (y - a_{y,1})^2} \\ \sqrt{(x - a_{x,2})^2 + (y - a_{y,2})^2} \\ \vdots \\ \sqrt{(x - a_{x,n})^2 + (y - a_{y,n})^2} \end{pmatrix}. \tag{3.5}$$

La matrice de covariance Q du bruit de processus tient compte des facteurs non modélisés du système, ainsi, dans le modèle PV l'accélération est traitée comme du bruit aléatoire.

Pour ce filtre, la fonction de mesure est non linéaire, ce qui interdit l'utilisation d'un filtre de Kalman standard. Nous avons donc implémenté ce filtre avec un *Unscented Kalman Filter* (UKF). Grâce à l'Unscented Transform (UT), les UKF peuvent déterminer des points d'échantillonnage (points Sigma) pour propager la caractéristique statistique de la variable d'état au fil du temps [116]. L'implémentation, réalisée en python, s'appuie sur le module FilterPy de R. Labbe [134] qui implémente les fonctions essentielles pour réaliser un UKF.

**Mise en œuvre** : Le dispositif que nous mettons en oeuvre pour la localisation et le guidage de personnes aveugles est typiquement constitué d'un réseau de quatre balises placées au sommet d'un rectangle. La personne à guider pourra évoluer soit à l'intérieur de ce rectangle, soit à l'extérieur. Avec les transceivers que nous utilisons la surface de guidage maximale est de l'ordre de 1000m<sup>2</sup> ce qui correspond à un petit gymnase. Dans la pratique, on effectue la latération en 2D, les ancrs et le tag sont donc placés dans un même plan et pour se placer autant que possible en condition de vue direct ("LOS") le tag est placé au dessus de la tête de l'athlète, sur un casque de protection.

La figure 3.4 montre une mise en œuvre typique avec un athlète non-voyant.



FIGURE 3.4 – Personne à localiser à proximité de balises UWB. Si on résout en 2D, les ancrs doivent être dans le plan du tag.

**Algorithme :** L'algorithme 1 décrit la méthode utilisée. On distingue une phase préalable de calibration et une phase temps-réel. La phase de calibration peut être prolongée parallèlement à la phase temps-réel pour affiner les informations de calibration. Cependant l'intérêt pratique de la méthode reste limité dans la mesure où une précision suffisante pour notre application de guidage est rapidement atteinte.

---

**Algorithm 1:** Algorithme de localisation par UWB

---

```
# Mettre en place les ancrs dans la zone. On choisira préférentiellement une géométrie
  rectangulaire
# Phase de calibration
for  $k = 0$  to  $Max$  do
  | Collecter les mesures
end
Éliminer les mesures aberrantes (médiane mobile)
Résoudre le système pour obtenir la position (biaisée) des ancrs.
# Phase de mesure temps-réel
while  $Mesures$  do
  | Kalman : Predict
  | Mise à jour de la matrice de covariance de bruit de mesure
  | Mise à jour de l'indicateur d'état du système (complet/dégradé)
  | Kalman : Update
end
```

---

Nous avons mis en place le système dans deux configurations :

1. Le tag est placé sur un train électrique qui circule sur un circuit circulaire d'environ 1.5m de rayon. Le circuit est placé au centre d'un carré de 3m où on a placé les ancrs à chaque sommet. L'ensemble est installé en intérieur et la position des ancrs est mesurée avec une précision centimétrique (mesuré avec un mètre ruban). Le train évolue de façon parfaitement plane.



2. Le tag est fixé sur un casque cycliste porté par un sujet se déplaçant librement en extérieur sur un terrain de 15m x 30m. Les ancrs sont placées sur des supports à hauteur du tag aux quatre coins d'un quadrilatère d'environ 10m x 10m (centré sur le terrain) et positionnées avec une précision décimétrique (mesuré avec un mètre ruban).

Dans les deux cas la résolution se fait en 2D, les mouvements verticaux du sujet sont donc susceptibles d'influencer la précision de la solution.

La solution de localisation peut être évaluée suivant les critères d'exactitude et de robustesse.

**L'exactitude** : Dans le cas (1) du tag placé sur un circuit en condition idéale on mesure une erreur maximum de 13 cm pour le cas à 4 ancrs et 20 cm sur le cas à 3 ancrs, sans différence significative entre les modèles PV et PVA. Les résultats pour un modèle PV sont visibles figure 3.5. Pour le cas (2) d'un tag placé sur la tête d'une personne dans

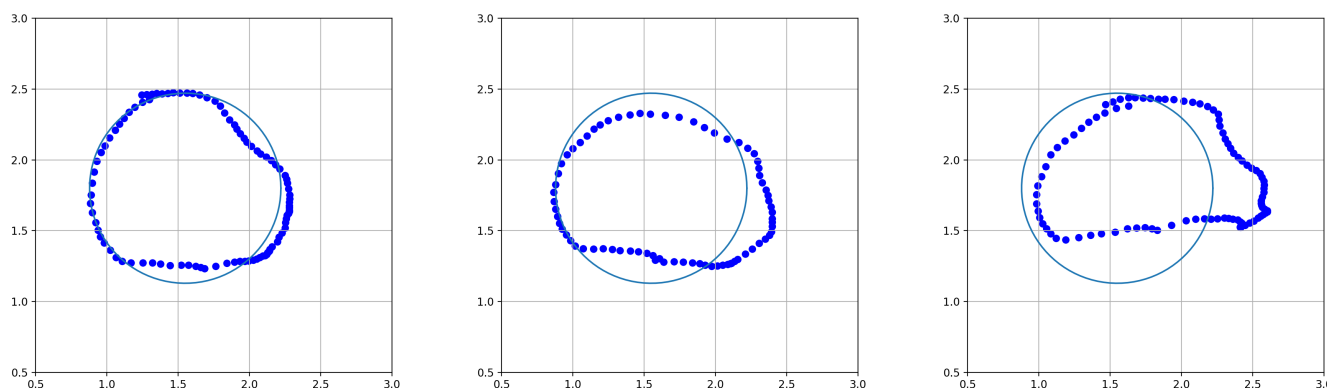


FIGURE 3.5 – Localisation sur un circuit circulaire de petite taille (en 2D). De gauche à droite : Avec 4 ancrs, 3 ancrs et 2 ancrs

des conditions d'utilisation réelle, avec un sujet marchant ou courant et un parcours complexe comportant des lignes droites et des courbes plus ou moins serrées, l'exactitude des mesures est reportée dans le tableau ci-dessus :

	PV		PVA	
	3 ancrs	4 ancrs	3 ancrs	4 ancrs
Err. Moyenne (m)	0.23	0.19	0.23	0.19
Err. RMS (m)	0.28	0.25	0.27	0.25
Err. Max (m)	0.81	0.59	0.79	0.61

Là encore les deux modèles considérés fournissent des résultats très comparables. Pour une comparaison plus précise un histogramme de répartition des erreurs est donné en Figure 3.6 et le parcours réalisé ainsi que les résultats de localisation sont visibles sur la Figure 3.7.

**Robustesse** : Pour un dispositif de guidage, la question de la robustesse est primordiale, ce qui implique de gérer le défaut de réception d'une ou plusieurs ancrs.

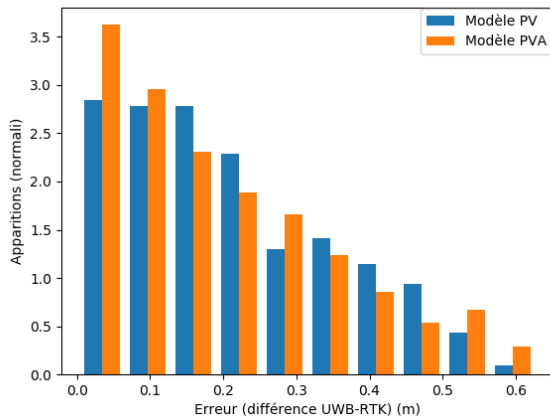


FIGURE 3.6 – Histogramme de répartition de l'erreur (distance par rapport à la référence RTK) pour les modèles PV et PVA sur le parcours de droite

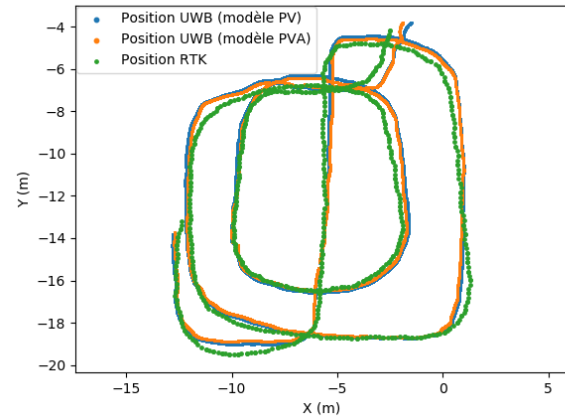


FIGURE 3.7 – Localisations UWB et référence RTK pour un parcours en extérieur à la vitesse de la marche et de la course lente (<10km/h). Les deux solutions UWB (vert et bleu) sont pratiquement confondues

De façon simplifiée, le protocole TWR consiste en l'envoi d'une trame *broadcast* à chaque ancre, alors suivi d'une réponse. Il n'y a donc pas de difficulté à identifier une borne qui ne fournit pas d'information. Une méthode simple et efficace en pratique pour gérer l'absence de mesure consiste à modifier la matrice de covariance de bruit d'observation  $R$  en fixant temporairement une valeur de variance très grande pour le capteur à l'information manquante. Suivant le nombre de balises accessibles, le guidage peut alors se poursuivre de façon dégradée ou être interrompu. D'une façon générale l'état du système peut être remonté vers l'utilisateur pour l'informer d'une défaillance partielle ou totale.

Nous avons donc évalué la possibilité de poursuivre la localisation en cas de défaillance d'une ou plusieurs ancres. Dans le cas du circuit circulaire un exemple montrant la poursuite dégradé de la localisation avec deux ancres est donnée en Figure 3.5. Dans les conditions d'utilisation réelles, on peut constater des erreurs conduisant à des résultats de localisation faux. La Figure 3.8 est un exemple significatif où la localisation avec deux ancres fonctionne correctement pendant plusieurs dizaines de mètres puis conduit l'algorithme à converger vers un parcours totalement en erreur.

**Nombre minimal de balises :** La localisation en 2D nécessite trois balises pour un fonctionnement fiable et la solution obtenue avec ces trois balises est d'une précision suffisante pour notre application. Cependant, l'ajout de balises supplémentaires permet de rendre le dispositif plus robuste en cas de défaillance (masquage, limite de réception) ce qui est un avantage significatif. À noter que le coût en temps de calcul lié aux balises surnuméraires est tout à fait négligeable sur les plateformes matérielles que nous utilisons. Les seules contraintes sont donc le coût et les éventuelles difficultés de mises en place des balises.

### 3.1.4 Conclusion

Le matériel et les algorithmes mis en œuvre permettent une localisation en temps-réel efficace et robuste dans un petit gymnase, avec une précision décimétrique et à une fréquence d'environ 10Hz. Le système peut potentiellement passer à l'échelle en augmentant le nombre de balises et ainsi augmenter la zone de couverture. En outre, la technique

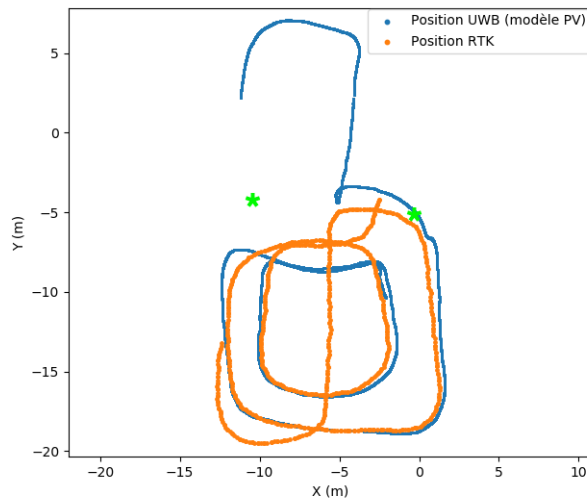


FIGURE 3.8 – Localisation UWB et référence RTK avec seulement 2 ancres fonctionnelles. Après quelques difficultés de convergence initiales, le système fournit une localisation correcte pendant quelques dizaines de mètres avant de converger vers une position en erreur, symétrique à la position réelle par rapport aux ancres. Les ancres fonctionnelles sont matérialisées par des étoiles vertes.

permet d’avertir l’utilisateur en cas de défaillance. La mise en œuvre pratique est relativement aisée : il est possible d’équiper un lieu en une quinzaine de minutes (calibration incluse). D’un point de vue matériel, l’algorithme a pu être implémenté sur une plateforme embarquée type Raspberry portée par l’utilisateur (cet aspect sera détaillé en partie 3). Nous avons ainsi pu utiliser cette technologie pour une partie des expériences de guidage qui seront décrites plus tard dans ce document. Aujourd’hui, des dispositifs de localisation complets existent sur le marché<sup>1</sup>, avec des caractéristiques comparables aux systèmes décrits ci-dessus.

### 3.2 Localisation par réseau de capteurs radar Doppler hyperfréquence

Les petits radars Doppler hyperfréquences à onde continue sont largement utilisés dans les détecteurs de présence (e.g. ouverture de portes) ainsi que pour le contrôle de vitesse de véhicules. Mais contrairement aux radars modulés en fréquence (Frequency-modulated continuous-wave radar - FMCW) ou à impulsions, ils ne sont pas destinés à mesurer des distances. Si le radar et l’objet à localiser sont en mouvement l’un par rapport à l’autre, l’effet Doppler produira une modulation du signal réfléchi. Le radar Doppler fonctionnant comme un démodulateur, la composante basse fréquence, disponible en sortie du capteur permet de déterminer les composantes de vitesses radiales de l’ensemble des objets en mouvements dans le champ du radar. D’un point de vue pratique, nous proposons d’utiliser des radars conçus pour la détection de présence. De conception très simple, ils sont à la fois très compacts (typiquement 3-4cm de côté, c.f. Fig. 3.9) et très économiques

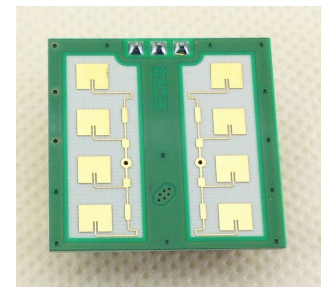


FIGURE 3.9 – Module radar doppler CDM324/IPM165

1. par exemple *Pozix*

( $\approx 10$  euros). Ces radars n'étant pas a priori destinés à des applications de mesures, leur précision et les caractéristiques de leur bruit de mesure ne sont pas garanties par le fabricant.

Une revue de littérature montre que ces modules ont été peu exploités dans la localisation des personnes, l'absence d'information de distance étant probablement le principal défaut pour ces applications. Ils sont par contre plus facilement utilisés à des fins d'évitement d'obstacles [212]. Dans le domaine de la localisation, une étude [224] a montré la faisabilité d'un dispositif de suivi en associant ces radars à la technologie UWB qui permet alors de pallier l'absence d'information de distance. Afin de ne pas avoir à utiliser d'autres capteurs, nous proposons ici une approche différente et originale dans laquelle nous utilisons l'intensité du signal reçu (RSS) du radar afin d'obtenir une évaluation de la distance. Nous montrerons que cette estimation de la distance, associée aux mesures de vitesse des capteurs, présente une précision suffisante pour une application pratique de guidage.

### 3.2.1 Estimation de la vitesse et de la distance

Selon les équations classiques de l'effet Doppler, la vitesse  $v$  mesurée par le radar peut s'écrire :

$$v = \frac{c f_d}{2 f_{tx} \cos(\alpha)} \quad (3.6)$$

avec  $c$  la vitesse de la lumière,  $f_d$  la fréquence doppler,  $f_{tx}$  la fréquence du signal radar (typiquement 24GHz) et  $\alpha$  l'angle formé entre la direction du mouvement et le faisceau radar.

Il faut noter que dans la pratique, pour la localisation de personnes, le signal doppler résultant contient plusieurs composantes de fréquences liées aux mouvements des membres. Bien que ces capteurs ne soient pas conçus pour fournir des informations de distance, nous montrerons que l'intensité du signal reçu (RSSI pour *Received Signal Strength Indication*) fournit une information qui peut être néanmoins exploitable.

Pour les radars, la puissance reçue  $P_r$  est généralement écrite en termes de puissance d'émission  $P_t$  comme :

$$P_r = P_t \frac{G^2 \lambda^2 \sigma}{(4\pi)^3 R^4} \quad (3.7)$$

où  $G$  est le gain d'antenne,  $\sigma$  la section efficace (réflexion) de la cible, et  $\lambda$  la longueur d'onde.

Pour un signal émis  $s_t(t) = \cos(2\pi f_c t)$ , en négligeant le terme de phase, le signal reçu peut être écrit sous la forme  $s_r(t) = \alpha \cos(2\pi(f_c + f_d)t)$  où  $\alpha$  est un facteur d'atténuation dépendant de la distance et que l'on déduit de (3.7).

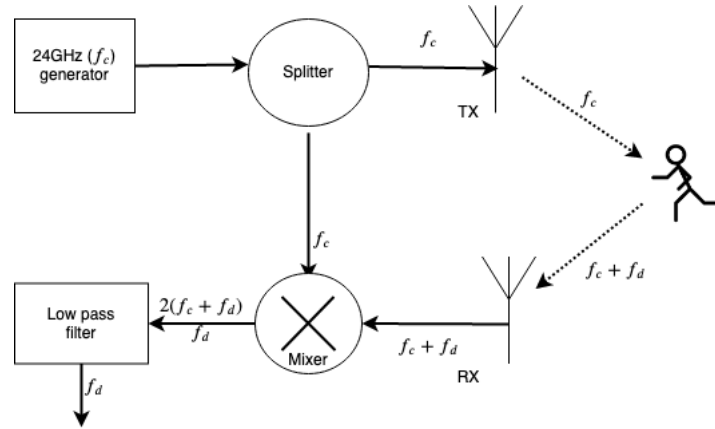


FIGURE 3.10 – Principe d'un radar à onde continue. Le signal du générateur 24GHz est diffusé via l'antenne d'émission. Lors de la réception, le signal reçu est multiplié par le signal source. Le signal résultant est alors filtré pour ne conserver que la composante à basse fréquence. L'ensemble des composants du capteur est intégré dans le module de la figure 3.9.

Dans le capteur, le signal reçu est mixé (i.e. multiplié) avec le signal émis (cf. fig. 3.10). Le signal devient alors :

$$\begin{aligned} s_{\text{mixed}}(t) &= \cos(2\pi f_c t) \cdot \alpha \cos(2\pi(f_c + f_d)t) \\ &= \frac{\alpha}{2} \cos(2\pi(2f_c + f_d)t) + \frac{\alpha}{2} \cos(2\pi f_d t) \end{aligned} \quad (3.8)$$

Il en résulte une somme de deux signaux, un basse fréquence et un haute fréquence. Dans une dernière étape, le signal est filtré par un filtre passe-bas. L'amplitude du signal résultant  $S(t)_{IF}$  est directement proportionnelle à la puissance reçue et peut être utilisée pour estimer une distance.

$$s(t)_{IF} \propto \frac{\alpha}{2} \cos(2\pi f_d t) \quad (3.9)$$

Nous proposons d'utiliser la magnitude de la fréquence prédominante dans le signal ce qui devrait permettre également de mieux sélectionner le signal direct.

Finalement, la distance  $R$  est déduite de cette magnitude, en supposant que la section efficace de la cible est constante pendant le mouvement et que la magnitude suit un modèle en champ libre  $k/R^4$  (cf. eq. (3.7)). Comme aucun émetteur ou récepteur n'est porté, cette méthode est naturellement insensible aux problèmes d'ombrage ou d'alignement de la polarisation de l'antenne dont souffrent habituellement les techniques RSS.

Ces capteurs possèdent une antenne plane gravée sur un circuit imprimé (*patch-plane antennas*) peu directionnelle et ont une atténuation inférieure à 3dB à  $\pm 60^\circ$  sur le plan horizontal (cf. Fig 3.11). Il est donc souhaitable de les éloigner de quelques mètres de l'espace de travail pour le couvrir complètement sans introduire trop d'atténuation liée à la directivité.

Typiquement, la sortie Doppler de chaque capteur est reliée à un circuit d'amplification approprié (environ 60dB) comprenant un filtre passe-bande 5Hz-900Hz (correspondant à une vitesse maximale de 20km/h). Dans notre dispositif, la sortie de l'amplificateur est connectée à l'entrée d'une carte audio multicanal qui permettra l'échantillonnage du signal. Notre programme pourra alors calculer la fréquence Doppler et l'amplitude du signal reçu par chaque capteur.

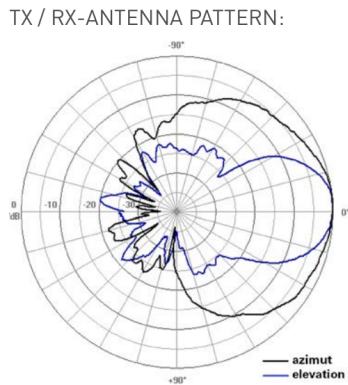


FIGURE 3.11 – Diagramme de directivité du module Doppler Innosent IPM-165. Extrait du datasheet du capteur [109]

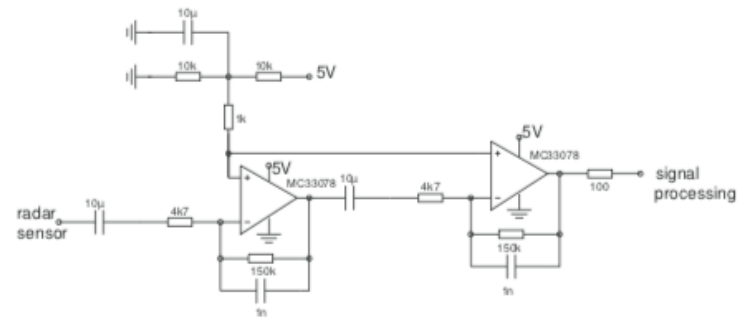


FIGURE 3.12 – Nous avons utilisé le circuit d'amplification préconisé par le constructeur dont nous avons adapté la bande passante (filtres RC). Circuit extrait de la note d'application InnoSent [108].

### 3.2.2 Estimation de la distance utilisant le RSS

Nous avons cherché à évaluer la précision de l'estimation de distance (*ranging*) avec un modèle simple en  $1/r^4$  qui utilise l'amplitude du signal Doppler. Le capteur ne permet pas de faire de mesure dans le cas statique ce qui rend la mise en place de tests expérimentaux moins aisée. Pour se placer dans un contexte proche de la réalité, nous avons effectué des mesures avec un sujet se déplaçant (1) dans l'axe du capteur, puis (2) de façon orthogonale au capteur. Pour évaluer l'erreur, le sujet est équipé d'une balise UWB et une seconde balise est placée près du module radar. Ce couple de balises UWB fournira alors une mesure de référence (à environ 10cm prêt).

1. Dans un premier temps on va mesurer la valeur du RSS lorsqu'un sujet fait des allers et retours dans l'axe du capteur. Le sujet effectue donc plusieurs aller-retours à des vitesses variables (pour évaluer un éventuel impact de la vitesse). On peut alors comparer les mesures utilisant le modèle RSS avec les mesures de distances UWB. L'expérience est réalisée en intérieur et en extérieur. La figure 3.14-a permet d'illustrer des mesures avec deux aller-retours en intérieur (pièce de 7mx4m) puis en extérieur. On peut constater que le modèle en  $k/R^4$  fournit une estimation réaliste à la fois en extérieur et en intérieur. On notera néanmoins que les mesures à faible distances sont faussées par la saturation de l'amplificateur, il est en effet difficile de gérer une dynamique aussi importante avec un amplificateur à gain fixe.
2. Il peut sembler a priori moins évident de prédire la réponse du radar en amplitude lors des déplacements orthogonaux puisque la section efficace de la cible risque d'être sensiblement différente du cas précédent. Pour vérifier expérimentalement notre approche, un sujet se déplace de façon orthogonale au capteur, c'est-à-dire en décrivant des arcs de cercle, de façon à garder une distance constante. La figure 3.14-b présente les mesures de distance pour différentes distances. On constate que la dispersion augmente fortement à mesure que la distance augmente mais fournit tout de même une information de distance a priori exploitable.

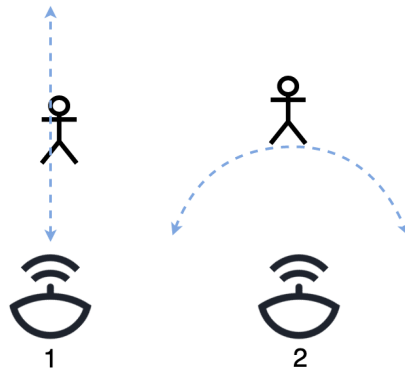


FIGURE 3.13 – Schéma de principe des expériences de mesure de distance par RSS avec les radars doppler. Dans le cas (1) le sujet se déplace face au capteur. Dans le cas (2), le sujet conserve une distance constante

La méthodologie décrite ci-dessus permet d'évaluer assez grossièrement la précision de la mesure de distance mais montre que le modèle proposé est cohérent avec les mesures expérimentales. Le tableau 3.2 donne l'erreur RMS lors des séries de mesures effectuées dans les deux types de déplacement étudiés :

	Intérieur	Extérieur		2m	4m	6m
RMS Error (m)	0.65	0.7	RMS Error (m)	0.6	0.7	0.8

(a) Déplacements face au capteur (cas 1)                      (b) Déplacement à distance constante (cas 2)

TABLE 3.2 – Erreurs de mesures constatés lors des deux types de déplacement. Les expériences montrent des résultats comparables en extérieur et en intérieur. Sur les déplacements latéraux, on peut constater une augmentation de la dispersion avec l'éloignement du sujet.

Le modèle fonctionne de façon comparable en extérieur et en intérieur. On peut noter que les capteurs sont très sensibles aux mouvements des objets environnants (mouvement des arbres, de l'herbes) qui peuvent induire un bruit de mesure visible en extérieur.

### 3.2.3 Filtre particulière pour la localisation en temps réel

Le système expérimental pour la localisation temps réel est constitué d'un ensemble de capteurs statiques comme le montre la figure 3.15 ici avec trois capteurs.

La méthode mise en œuvre pour le *tracking* devra tenir compte de plusieurs caractéristiques du système :

- Les capteurs des radars fournissent des mesures très bruitées ou manquantes (pas de pics de fréquence ou signal trop faible).
- Il est impossible de distinguer les situations où le sujet est immobile de celles où il n'y a pas de mesures parce que le sujet est hors de portée.
- Le bruit de mesure n'est pas gaussien et difficile à modéliser.

Pour résoudre ces problèmes, nous utilisons un filtre particulière qui est bien adapté aux modèles non linéaires et au bruit de mesure non gaussien. Comme

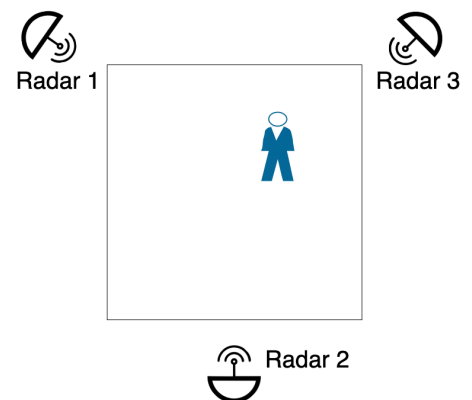
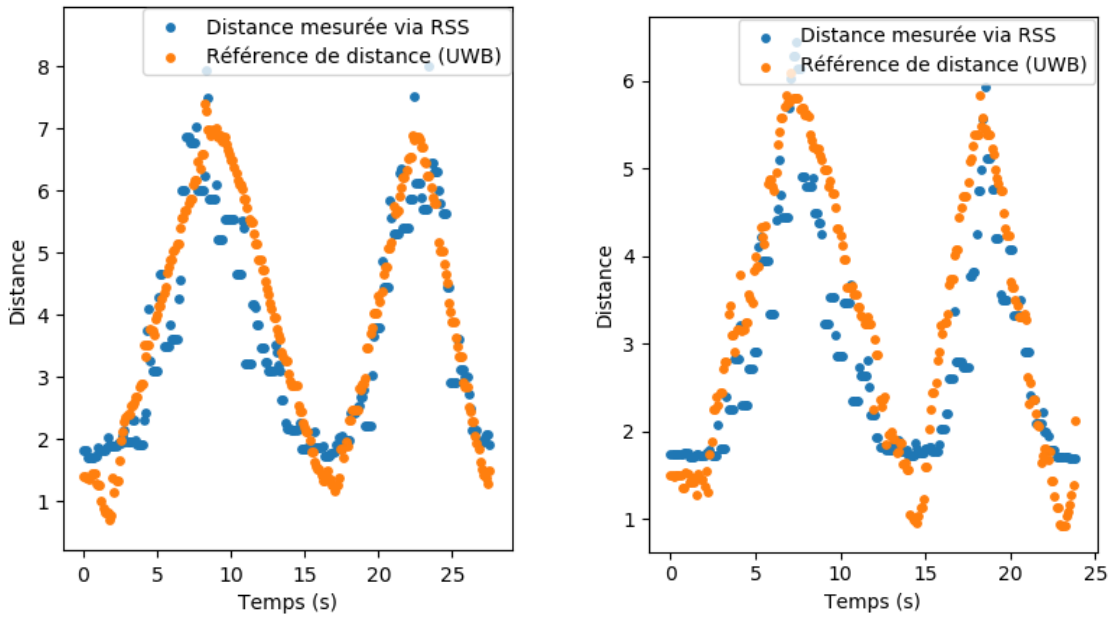
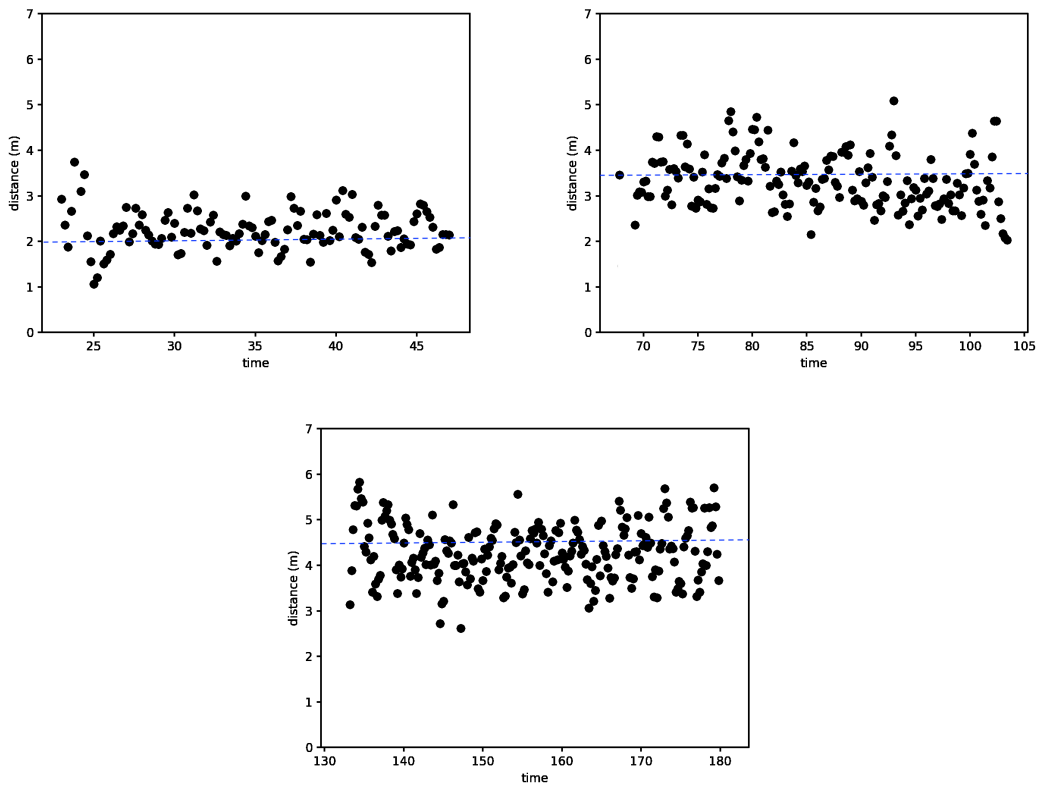


FIGURE 3.15 – Le dispositif expérimental avec trois capteurs



(a) Mesures de distance pendant des aller-retours dans l'axe du capteur. À gauche l'expérience est réalisée en extérieur, à droite en intérieur. La mesure de distance de référence est donnée par une balise UWB portée par le sujet. Dans chaque cas, le sujet réalise deux aller-retours, d'abord en marche lente puis en marche rapide



(b) Mesures de distance pendant trois déplacements orthogonaux au capteur (distance réelle constante, en pointillés bleu). De gauche à droite, les mesures sont effectuées à des distances de 2m, 3.5m et 4.5m.

FIGURE 3.14 – Mesures de distances lors de déplacements dans l'axe ou orthogonaux au capteur.



précédemment, on utilise un modèle à vitesse constante (PV), et le filtre utilise cette dynamique comme fonction d'importance.

Afin d'augmenter la précision du système, l'utilisateur peut en outre être équipé d'une IMU avec fusion AHRS donnant l'orientation du sujet et par extension son cap. Dans les expériences ci-dessous, nous avons utilisé un capteur TDK-Inversense MPU9250 à bas coût connecté à un microcontrôleur qui exécute l'algorithme de fusion de données Magdwick [154]. Un tel dispositif peut fournir des informations d'orientation avec une précision d'environ 3-5°. Il s'agit d'un petit dispositif sans fil (de la taille d'une boîte d'allumettes), qui est porté à la ceinture.

Dans la méthode décrite ci-dessous, on fait l'hypothèse que l'utilisateur se déplace dans la direction du plan sagittal (i.e. s'oriente dans sa direction de déplacement). L'IMU permet en outre de pallier la difficulté d'évaluer le sens de déplacement avec les radars.

Le filtre particulaire est un estimateur qui fonctionne en plusieurs étapes : une étape de *prédiction* utilisant le modèle dynamique, suivi d'une étape de *mise à jour* et de *sélection* utilisant les mesures si elles sont disponibles. On utilise des *particules* qui représentent des états possibles du système. Elles seront propagées et sélectionnées à chaque itération. L'étape de prédiction consiste donc à propager les particules suivant le modèle dynamique. Dans le modèle PV, les particules représentent le vecteur d'état  $\mathbf{X} = (x, y, \dot{x}, \dot{y})$  dont l'état évoluera en considérant une vitesse constante (dynamique décrite en Table 3.1 au chapitre précédent). En outre du bruit est ajouté dans l'état des particules pour représenter les mouvements non modélisés. À chaque particule on associe un poids qui, lors de l'étape de mise à jour, évoluera en fonction de son adéquation avec la mesure. Dans l'algorithme SIR utilisé ici (voir annexe B) le poids (pondéré) correspond à la probabilité relative que la particule représente l'état à estimer. Dans le cas présent, le poids de particules sera mis à jour suivant l'équation (B.20) (décrite en annexe B) qui permettra de tenir compte de :

- l'adéquation de la position de la particule par rapport aux mesures de distance (modèle RSS) ;
- l'adéquation de la vitesse de la particule avec la mesure de vitesse des capteurs ;
- l'adéquation de l'orientation du vecteur vitesse de la particule  $\theta = \arctan \dot{y}/\dot{x}$  avec l'angle Yaw de l'IMU.

Les détails mathématiques concernant le fonctionnement d'un filtre particulaire sont données en annexe B.

---

**Algorithm 2:** Algorithme du filtre à particules du Radar Doppler.

---

**Result:** Filtre particulaire

(Initialisation) Création aléatoire d'un ensemble de particules représentant les états possibles (vitesse et position)

**for**  $k = 0$  **to**  $Max$  **do**

- Pour toutes les particules : prédiction de l'état de la particule suivante en supposant un modèle à vitesse constante
- Mesure de la vitesse radiale des capteurs, distance déduite de l'orientation RSS et IMU.
  - Identifier les situations où la cible est statique (vitesse et RSS nulle pour tous les capteurs)
  - Ecarter les mesures inconsistantes;
- Pour toutes les particules : Mise à jour du poids des particules en tenant compte de la mesure;
- Suppression des particules de poids trop faible et rééchantillonnage;
- Calcul de l'estimation (en utilisant une moyenne pondérée);

**end**

---

### 3.2.4 Mise en oeuvre et tests

Afin d'évaluer la méthode, nous avons mis en œuvre le dispositif décrit dans la Figure 3.15 sur une surface de 8m x 8m. Tous les tests qui vont suivre sont effectués avec une seule personne marchant lentement (vitesse de 0,5 à 1,5 m/s) dans cette zone.

Nous avons mis en œuvre la méthode en utilisant tout ou partie des capteurs sur des mesures effectuées lors d'un parcours circulaire simple afin de montrer l'intérêt de la fusion de données. Ces résultats sont visibles sur la Figure 3.16. On peut constater que si on utilise uniquement la mesure de vitesse Doppler, la position est rapidement affectée par une dérive, alors que le RSSI seul présente une trace plutôt erratique en raison de l'imprécision des mesures. La fusion de ces données permet alors d'obtenir une localisation qui apparaît plus précise.

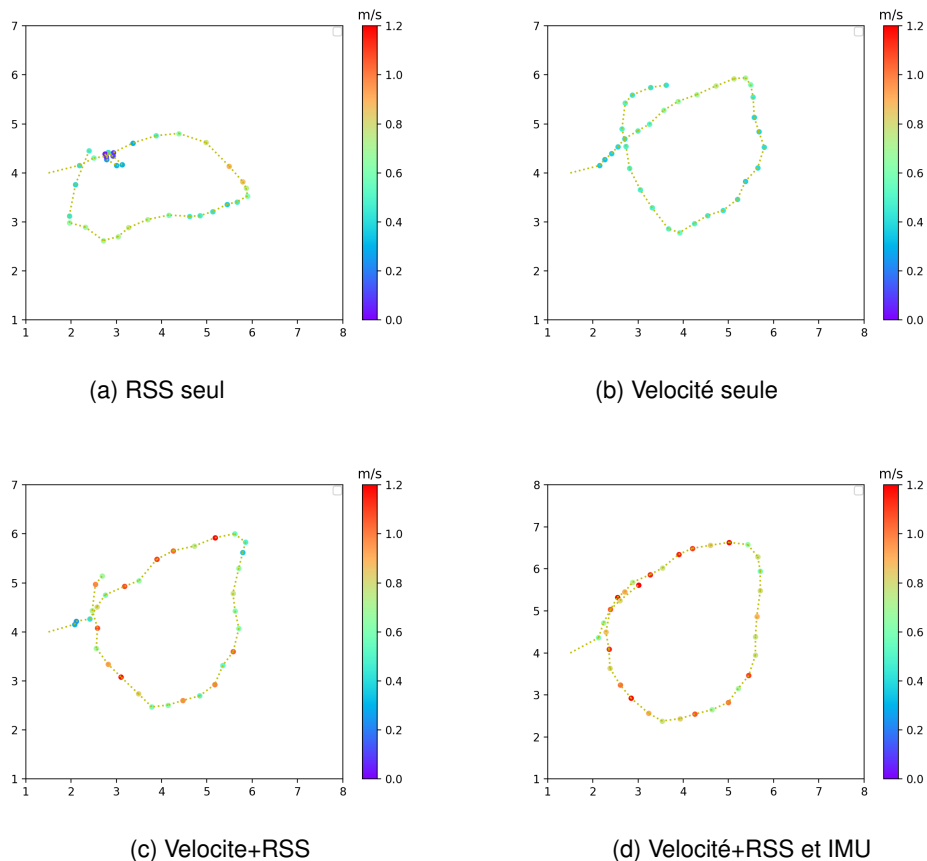


FIGURE 3.16 – Résultats de fusion de données typiques pour un circuit circulaire.

Afin d'identifier un éventuel problème de dérive, nous avons testé la méthode avec un parcours de type circulaire d'une dizaine de tours (c.f. Fig. 3.17) pendant lequel le sujet a varié sa vitesse (marche rapide/lente).

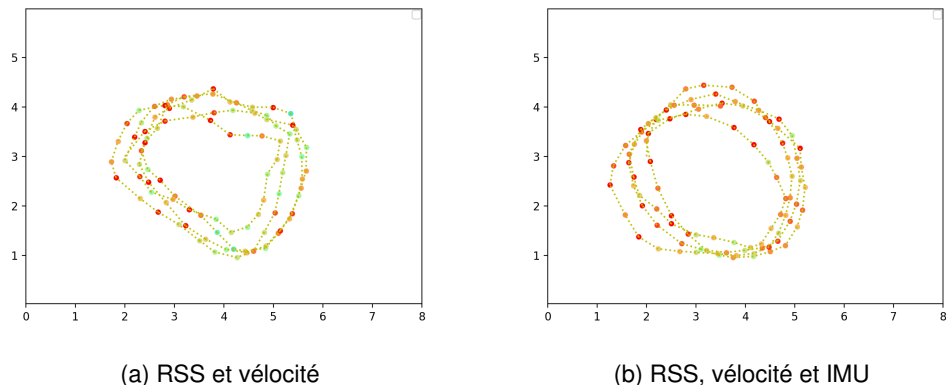


FIGURE 3.17 – Suivi sur un parcours circulaire avec plusieurs tours.

En l'absence de position de référence sur ce parcours il est difficile d'évaluer la précision de la localisation dans cette expérience, mais on peut noter qu'il n'y a pas de dérive significative, même dans les expériences comportant de nombreux tours successifs.

Afin de montrer les capacités de suivi du système, nous avons effectué différents types de déplacements avec des parcours comportant des déplacements doux ou abrupts. Ces tests ont été effectués en utilisant les données RSS et les vitesses Doppler mais aussi en intégrant l'orientation de la personne donnée par l'IMU. La figure 3.18 montre en particulier une comparaison de résultats obtenus avec ou sans la fusion des données de l'IMU. Pour évaluer et comparer les différents scénarios nous avons comparé les points d'arrivées attendu avec les points d'arrivées mesurés. Nous avons également mesuré l'écart maximal de la trajectoire mesurée par rapport à la trajectoire attendu (ce qui ne représente pas nécessairement l'erreur de positionnement maximal sur le trajet). Avec le tracking par capteur Doppler+IMU, sur différents trajets réalisés (multiples aller-retours en ligne droite, croix, cercles multiples, zig-zag) correspondant à des distances parcourues de 6m à 30m, l'écart par rapport à la trajectoire de référence est resté inférieur à 1,5 m et les erreurs de positionnement à l'arrivée étaient comprise entre 0.3m et 1m.

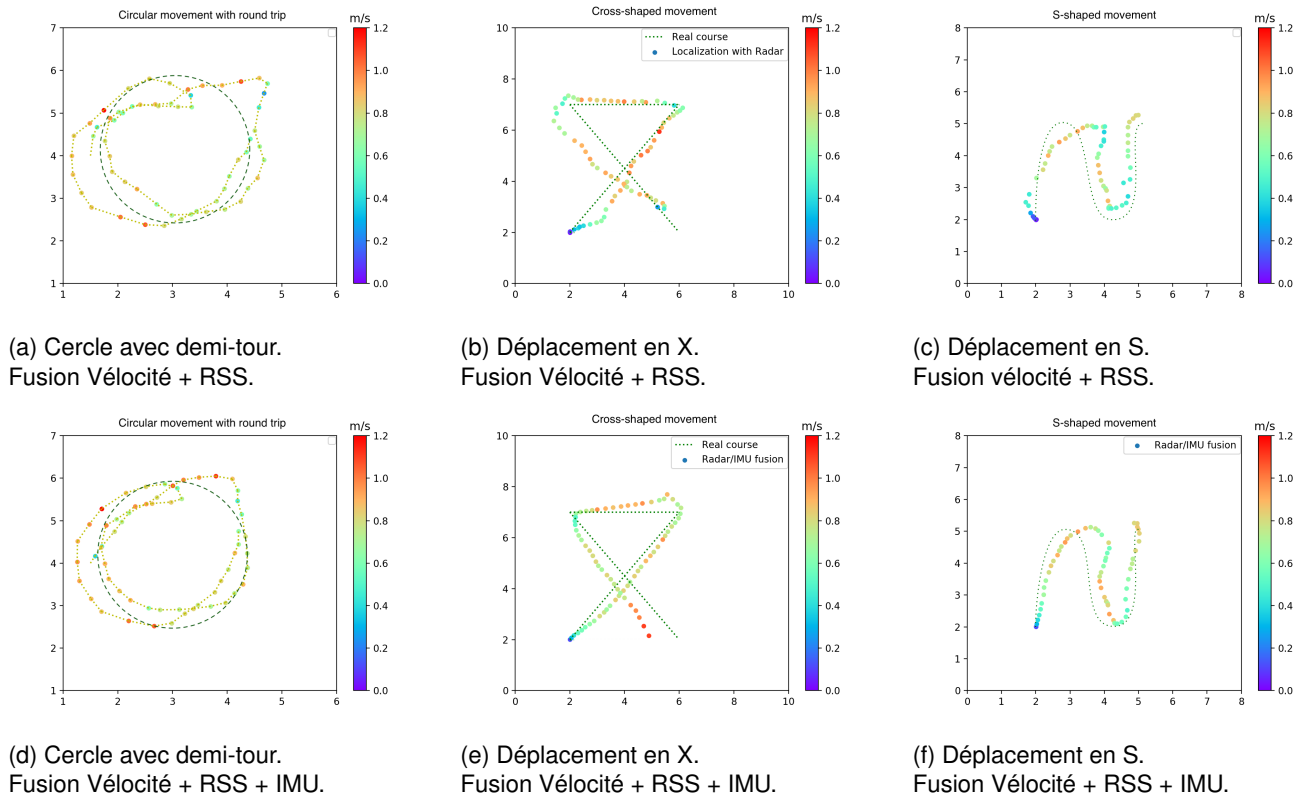


FIGURE 3.18 – Exemples de suivi de personne par notre méthode. Les lignes pointillées représentent le parcours de la personne.

### 3.2.5 Conclusion

Cette méthode est encore à l'état de preuve de concept car nous n'avons pu la mettre en œuvre que sur des surfaces de quelques dizaines de  $m^2$  ce qui limite son intérêt pratique pour une application de guidage. Cependant il ne semble pas y avoir de réels obstacles, soit à la création d'un réseau de capteur plus important, soit à l'utilisation de capteurs à plus longue portée. La précision, métrique, est inférieure à d'autres techniques de localisation en intérieur, mais son faible coût pourrait constituer un atout important pour certains cas d'usage.

## 3.3 Flot optique et localisation piétonne

Les techniques de navigation à l'estime sont encore rarement utilisées pour le suivi et la localisation des personnes en raison de la difficulté de construire un odomètre piéton robuste. Il est en effet délicat d'estimer la vitesse d'une personne qui marche ou court avec un dispositif qu'elle embarquerait. Les approches dans le domaine utilisent typiquement des centrales inertielles pour estimer le rythme et l'amplitude des pas (voir par exemple [84]) cependant ces méthodes, basées sur une modélisation de la démarche sont assez peu robustes et nécessitent souvent de mettre en œuvre des IMU de haute qualité dont le coût reste élevé. Il s'agit cependant de techniques présentant un grand intérêt, par exemple pour pallier à l'indisponibilité temporaire d'un signal GPS de qualité. Dans ce paragraphe, nous proposons

une approche qui nous paraît originale et montrons qu'il est possible d'utiliser un capteur de flot optique pointé vers le sol pour fournir un indication d'odométrie relativement robuste si certains pré-requis sont réalisés. Finalement nous réaliserons des expériences de suivi utilisant un capteur de flot optique, une centrale inertielle ainsi qu'une fusion de données avec l'UWB.

### 3.3.1 Odométrie piétonne par flot optique

Le flot optique permet d'estimer le mouvement apparent des objets filmés par une caméra en comparant les séquences d'images successives (c.f. paragraphe 1.3). L'odométrie par flot optique conventionnelle utilise un capteur orienté vers l'avant, mais une méthode alternative consiste à utiliser un capteur filmant le sol. Cette seconde approche, proposée pour l'odométrie de robot spatiaux [64] est aujourd'hui régulièrement utilisée pour les drones en phase d'atterrissage [87, 147]. Nous avons opté pour cette technique car elle présente plusieurs avantages. Elle simplifie la question de l'illumination de la scène dans la mesure où il est facile d'éclairer de façon uniforme et constante la zone sous le sujet plutôt que l'ensemble de l'environnement proche. La technique permet également d'éviter que des objets en mouvement se trouvent dans le champ de la caméra, ce qui simplifie grandement le traitement de l'image. Cette caractéristique va permettre d'utiliser des capteurs optiques à très bas coût. Malgré les quelques avantages ci-dessous, une revue de la littérature montre que ces méthodes ne sont pas habituellement utilisées pour la localisation de personnes. Une difficulté étant d'évaluer le bruit induit par les mouvements liés à la démarche.

**Capteur de flot optique :** Il est difficilement possible d'utiliser les caméras d'imagerie traditionnelles en raison de leurs trop faible vitesse (les ordres de grandeur typiques sont de 25 à 50 images/s). Ainsi nous utiliserons des capteurs de flots conçus pour les souris optiques. Il s'agit en réalité de caméras très rapides (jusqu'à 3000 images par seconde), à très faible résolution (typiquement moins de 1000 pixels) et associés à un processeur de traitement du signal mesurant le flux optique lié à l'aspérité des surfaces. Elles ont également l'avantage d'être extrêmement économiques (quelques euros). Associés à une optique adaptée, ces capteurs peuvent être utilisés pour mesurer le flux optique en pointant le sol depuis des hauteurs de quelques centimètres à plusieurs mètres (voir image 3.19).

En utilisant un modèle de caméra *pinhole*, et en considérant que le capteur se déplace sur un plan parallèle au sol, on peut montrer facilement que les valeurs de distances mesurées par ce type de capteurs sont proportionnelles à la focale de l'optique et inversement proportionnelles à la hauteur de mesure [87, 147] suivant la relation ( $f$  est la focale) :

$$\frac{d_{\text{mesure}}}{d_{\text{réel}}} = \frac{f}{\text{hauteur}} . \quad (3.10)$$

Ces capteurs fournissent un vecteur de flot optique correspondant aux déplacements sur les deux axes, ainsi qu'une valeur de *Surface Quality* (SQUAL) correspondant au nombre de points caractéristiques suivis sur l'image. La valeur de SQUAL varie suivant la texture du sol, la focalisation et les conditions d'éclairage. Notre algorithme devra savoir traiter les mesures à faible SQUAL car elles sont susceptibles de conduire à une sous estimation des distances [49].

La valeur du flot optique correspond à une grandeur de déplacement à chaque pas de temps. Cette valeur est donc homogène à une vitesse. Le flot optique est généré à la fois par les mouvement de translation mais aussi par les mouvement de rotation (ou oscillation) sur les axes  $x$ ,  $y$  et  $z$  qui peuvent éventuellement être corrigés utilisant les mesures d'une IMU [147].

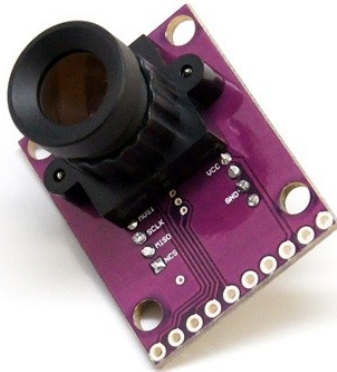


FIGURE 3.19 – Capteur de flot optique composé d'une caméra de souris optique et d'une optique. Ce type de capteurs est commercialisé pour l'équipement des drones en phase d'atterrissage.

**Déplacement d'un capteur dans le plan parfaitement parallèle au sol :** La Figure 3.20 montre la mesure de flot optique pour des déplacements successifs de 10 m avec un capteur situé à un mètre d'un sol plat (route asphaltée), maintenu parfaitement normal au sol et pour des vitesses croissantes (environ 4 à 7 km/h). La distance est déduite de la somme des flots mesurés pendant les déplacements. Pour dix expériences, on mesure un écart type de 35cm pour un déplacement de 10m.

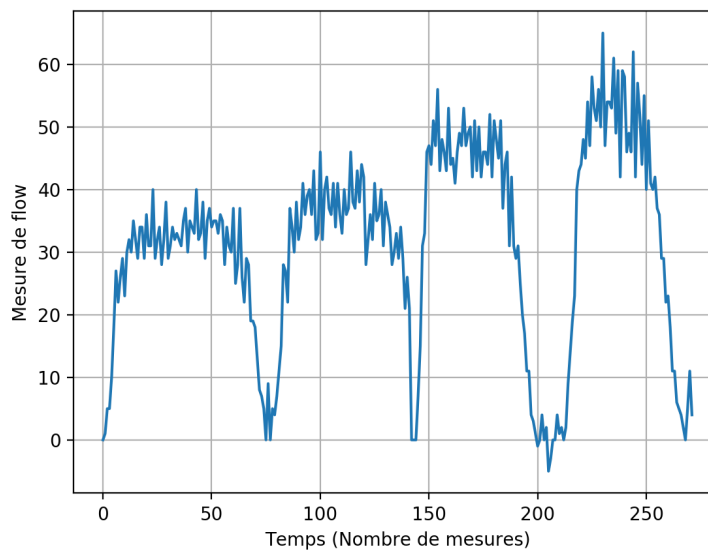


FIGURE 3.20 – Mesure de flot optique (vitesse) pour un capteur normal au sol à une hauteur de 1 m. Quatre déplacements d'environ 10m sont effectués à vitesse croissante.

Les tests ci-dessus sont effectués sur une route goudronnée correctement éclairée. La texture du sol est un élément clé dans le fonctionnement du capteur qui ne fonctionne pas correctement sur un sol trop lisse. Nous avons obtenu des résultats satisfaisants sur route, sur un gazon entretenu ou sur le revêtement de tartan d'une piste d'athlétisme.

**Capteur en conditions réelles :** Le dispositif expérimental utilisé pour les tests en conditions réelles est présenté Fig. 3.21. Il est constitué d'un capteur de flot optique qui est fixé à la ceinture et pointé vers le sol. En outre le boîtier est équipé d'une IMU qui sera utilisée par la suite. L'ensemble est parfaitement solidaire du sujet.

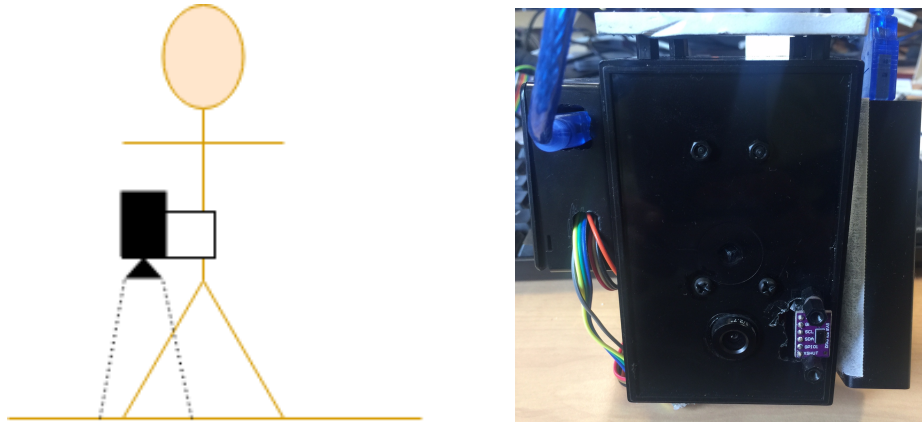


FIGURE 3.21 – À gauche : Schéma de principe d'un dispositif d'odométrie par flot optique avec un capteur fixé à la ceinture et pointé vers le sol. À droite : le dispositif vu de dessous. On peut distinguer l'optique, un lidar permettant de mesurer la distance au sol, une batterie et un émetteur sans fil pour transmettre les données sans avoir à transporter l'ordinateur.

La marche comme la course à pied conduit à des mouvements périodiques qui affecteront à la fois les mesures de l'IMU et les mesures de flot optique. Un capteur de flot optique fixé à la ceinture sera donc affecté d'un bruit quasi-périodique qu'il pourra être utile de filtrer. Dans notre cas d'usage, pour le suivi d'un coureur, il sera possible de faire l'hypothèse que cette période reste relativement stable tant que l'athlète reste en situation de course. Il est en effet généralement admis que la cadence de course d'un coureur reste assez stable lors d'une course même si l'intensité de l'effort varie, la vitesse étant principalement liée à l'amplitude du mouvement [58].

Le signal de flot optique est affecté d'un bruit qui semble corrélé aux oscillations de l'IMU à la démarche. Les pics du signal de flot sont visibles de façon périodique et légèrement déphasés par rapport aux mesures de l'IMU. À noter que les valeurs très basses du flot mesurées sont typiquement associées à une indice SQUAL faible et sont très vraisemblablement attribuables à une défocalisation du capteur. Il est préférable d'éliminer les valeurs de SQUAL faibles car elles conduisent à une sous-estimation des distances [49].

Pour filtrer ces données en temps réel tout en limitant la latence de mesure on pourra évaluer la cadence (en considérant qu'elle varie peu) en mesurant la période du signal IMU puis appliquer un filtrage passe bas sur la base de cette période. On peut par exemple utiliser la méthode suivante :

- *En tâche de fond* : on calcule la période  $\omega_{yaw}$  du signal de l'IMU (FFT sur une fenêtre de quelques secondes) qui correspond à la cadence de la démarche.
- De  $t - \omega_{yaw}$  à  $t$  on élimine les valeurs associées à des indices SQUAL faibles (par exemple celles sous la valeur médiane pendant cette période).

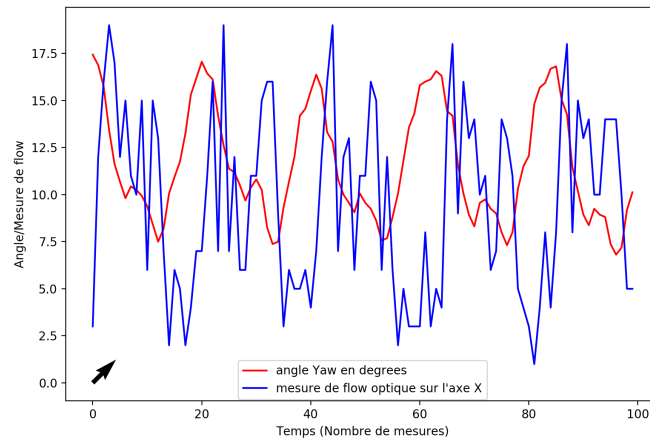


FIGURE 3.22 – Mesure de flot optique et mouvements mesurés par l'IMU (axe Yaw) pour un capteur fixé à la ceinture lors d'une course. Sur le signal de l'IMU on peut observer les oscillations périodiques en Yaw liées à la démarche. Le signal de flot optique comporte un bruit qui semble corrélé avec les oscillations de l'IMU, en particulier les pics apparaissent avec la même période.

- Sur la période, on identifie les valeurs pics du signal où l'on calcule la valeur médiane. Les pics ont l'avantage d'être souvent corrélés avec les valeurs de SQUAL élevées. Dans une application robuste de guidage, on peut imaginer que les valeurs SQUAL puissent être utilisées pour signaler un problème potentiel.

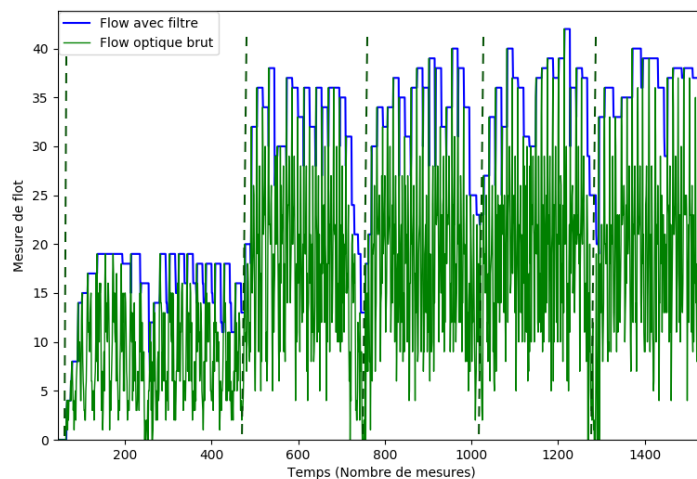
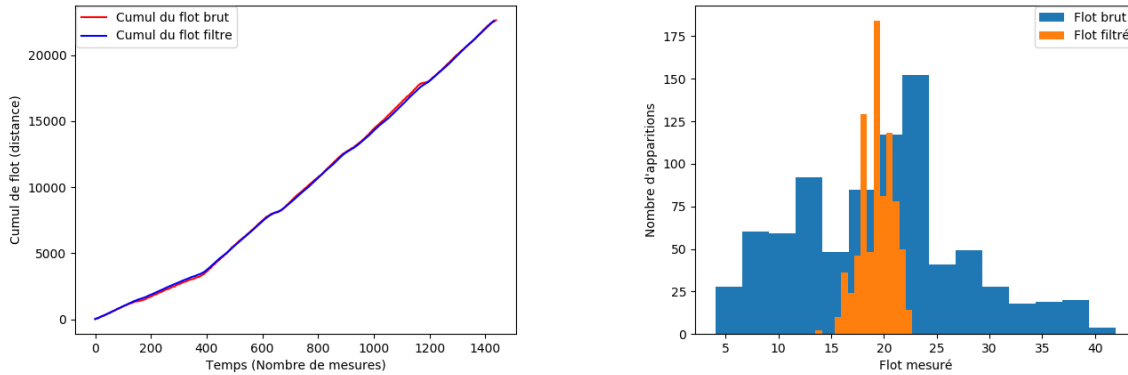


FIGURE 3.23 – Mesure de flot optique avec un capteur portée à la ceinture. Le sujet effectue cinq déplacements d'environ 15m entrecoupés d'arrêts, un premier à basse vitesse et quatre plus rapides. Le filtrage proposé est affiché en bleu.

Le figure 3.23 montre l'effet du filtrage sur le signal obtenu avec un capteur fixé à la ceinture sur un sujet se déplaçant en extérieur (sol goudronné convenablement éclairé). Le sujet effectue plusieurs déplacements d'environ 15m



à différentes allures, en marchant ou en course rapide et avec des enjambées de taille variable. Le filtrage n'entraîne pas de latence supérieure à la durée d'une enjambée ce qui est a priori suffisant pour nos applications. On peut évaluer la précision de l'odométrie en comparant les mesures cumulées de flot (filtré ou non) et les mesures réelles comme on peut le voir sur la Figure 3.24a. Le flot filtré est mis à l'échelle avec le flot via un coefficient linéaire simple à déterminer par calibration. La technique n'a pas de conséquence mesurable sur la précision qui reste de l'ordre de 3% d'erreur. Cependant, une fois le signal filtré le bruit présente les caractéristiques d'un bruit additif gaussien, dont la dispersion est très sensiblement réduite.



(a) Flot cumulé (homogène à une distance) correspondant à l'expérience précédente. Flot filtré (mis à l'échelle) et flot brut fournissent des mesures de distance cumulée comparable.

(b) Histogramme des mesures avant et après filtrage.

### 3.3.2 Exemples de localisation par flot optique

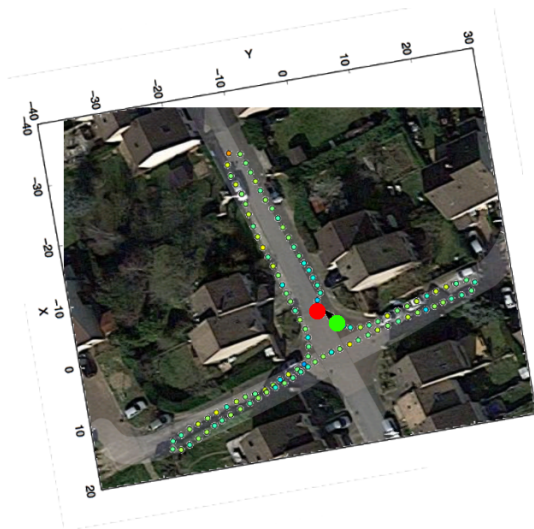
Pour montrer l'intérêt des capteurs de flot optique pointés vers le sol pour la localisation piétonne, et tirer parti du caractère gaussien du bruit, nous allons implémenter deux dispositifs de localisation utilisant un filtrage de Kalman. Nous utiliserons une IMU (couplée avec une fusion AHRS) avec un modèle simple utilisant la vitesse et le cap du sujet. Dans cette approche on suppose que le mouvement de la personne suit la direction donnée par le capteur d'orientation (cap).

**Fusion AHRS et flot optique** Nous considérons le vecteur de mesure  $z = (v, \theta)^T$  où  $v$  est la vitesse estimée à partir du capteur de flot et  $\theta$  est l'angle yaw de l'IMU. Nous cherchons alors à estimer la position et la vitesse dans un repère cartésien  $x = (x, y, \dot{x}, \dot{y})^T$ . Nous pouvons alors définir une fonction de mesure

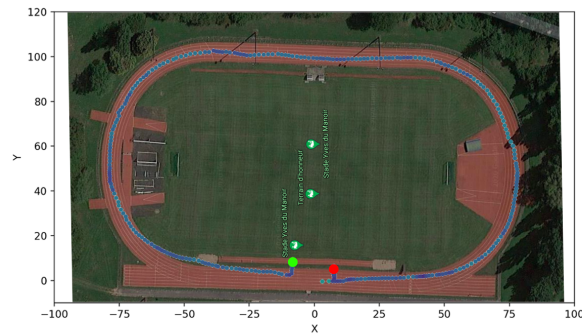
$$\mathbf{H}(x) = \begin{pmatrix} \sqrt{\dot{x}^2 + \dot{y}^2} \\ \arctan \frac{\dot{y}}{\dot{x}} \end{pmatrix} \quad (3.11)$$

Dans ce modèle, où l'on exclut les mouvements latéraux, la composante en ordonnée du vecteur vitesse mesuré doit toujours être nulle, le capteur étant solidaire du marcheur et en principe toujours orienté dans le sens de la marche.

Sur la figure 3.25a, on peut voir un parcours d'une centaine de mètres à pied sur une route goudronnée. À la fin de la boucle on arrive presque au point de départ (erreur d'environ 2 mètres). Un second test est effectué sur une piste d'athlétisme au revêtement de tartan. La figure 3.25b montre un parcours d'environ 400m et l'erreur finale (différence entre le point de départ et le point d'arrivée) est d'environ 20m, l'erreur étant également imputable à l'utilisation d'une IMU à bas coût. Les différents tests réalisés ont fait apparaître des erreurs de 2 à 5% quand ils sont effectués avec des sols adaptés.



(a) Parcours sur route.



(b) Parcours sur une piste d'athlétisme.

FIGURE 3.25 – Deux parcours effectués en marchant (environ 4 km/h) avec retour au point de départ. Les points de départ et d'arrivée estimés sont matérialisés par des disques respectivement rouges et vert. Le parcours réel est effectué sur le couloir central avec départ au point rouge.

### 3.3.3 Conclusion

Nous avons montré qu'un capteur de flot optique pointé vers le sol peut constituer une solution utilisable pour l'odométrie piétonne et permet des applications à la localisation temps-réel des personnes. Cette technique peut constituer une solution de localisation sur une distance limitée, par exemple pour pallier à une défaillance GPS ou en zone de mauvaise réception. Il est également possible de coupler ce système avec un système de localisation par proximité (réseau de balises) ou des références de distance pour pallier certains de ses défauts en évitant les problèmes de dérive, un exemple de fusion flot optique-UWB est donné en annexe C. La technique nécessite cependant des développements supplémentaires pour être plus largement utilisable en pratique. Certains aspects techniques doivent cependant être améliorés pour une utilisation en condition réelles. Un travail semble nécessaire sur l'éclairage du sol sous le capteur. Il serait également nécessaire de trouver une solution optique assurant une bonne focalisation du capteur. Les distances mesurées étant fonction de la hauteur du capteur, il pourrait également être utile de mesurer en temps réel la hauteur du capteur (avec un lidar par exemple) et de développer les méthodes de filtrage et de calibration utilisant cette information.



# Conclusion de la première partie :

Les méthodes de guidage décrites plus loin dans ce document doivent s'appuyer sur une localisation temps-réel suffisamment précise (typiquement métrique) et pouvant fonctionner, selon le cas, en extérieur ou en intérieur.

Les techniques les plus avancées de navigation globale par satellite (augmenté et multibande) sont des solutions satisfaisantes mais elles ne sont pas toujours disponibles, par exemple en intérieur mais également en environnement urbain. C'est pourquoi nous avons travaillé sur plusieurs méthodes alternatives de localisation utilisables là où le signal GNSS s'avère insuffisant.

La localisation en intérieur est un sujet de recherche très actif et de très nombreuses méthodes ont été proposées, c'est pourquoi nous nous sommes concentrés sur quelques technologies émergentes ou nouvelles ayant un réel potentiel d'applications pratiques.

La localisation par réseau de balise UWB, telle que nous l'avons implémentée, constitue une solution totalement fonctionnelle, robuste et adaptée à notre cas d'usage. Certains résultats de guidage présentés en troisième partie de ce document utilisent cette technologie dans des cas de guidage réel en intérieur. Cette solution, encore émergente au commencement de ces travaux est maintenant assez largement développée et des applications industrielles proposant des systèmes de localisation sont disponibles sur le marché (e.g. Pozix).

Nous avons également travaillé sur des méthodes originales utilisant d'autres types de capteurs pour lesquels il n'existait pas, à notre connaissance, d'applications en localisation piétonne. Nous avons pu montrer que ces capteurs (radar doppler et capteurs de flot optique), peuvent constituer des outils pour la localisation dans des cas d'usage spécifiques. Nous avons pu mettre en œuvre ces techniques sous forme d'applications temps-réel embarquables mais nous ne les avons pas encore utilisées dans des applications de guidage. Ces méthodes sont donc encore à l'état de preuves de concepts, des développements techniques supplémentaires sortant du champ de cette thèse étant encore nécessaires pour étendre leur rayon d'action, leur robustesse et leur précision. Différentes voies de développement sont possibles pour rendre ces technologies totalement utilisables en dehors de l'environnement maîtrisé du laboratoire, par exemple la multiplication des balises pour les techniques Doppler ou l'amélioration du capteur pour les techniques de flot optique.

Une synthèse des technologies que nous avons étudiées est donnée Tableau 3.3.

	GNSS			UWB	Doppler	Flot-optique
	Mono-bande	RTK	bi-bande			
Précision <sup>(1)</sup>	<10m	<10cm	<3m	10 à 60cm	<2m	<1m pour 20m parcouru
Utilisable en intérieur	Non	Non	Non	Oui	Oui	Oui
Rayon d'action	Couverture mondiale	500m à 10km	Couverture mondiale	30-40m <sup>(2)</sup>	8m <sup>(2)</sup>	potentiellement sans limite
Coût	généralement << 100 euros	<500 euros	<200 euros	<200 euros	<100 euros	<50 euros
Maturité	Industriel / grand public	Industriel/ Professionnel	Industriel / grand public à court terme	Industriel	Expérimental	Expérimental

(1) Ordre de grandeur de l'erreur maximale dans des cas normaux d'utilisation (ciel visible pour les technologies GNSS).

(2) Peut être étendu en multipliant les balises.

TABLE 3.3 – Comparaison des technologies de localisation étudiées dans ce document.

**Deuxième partie**

**Synthèse Binaurale**



# Introduction

Dans la continuité d'autres travaux de la communauté du son binaural, nous pensons que le son spatialisé peut constituer un outil efficace pour guider des personnes aveugles. D'une part, parce que les personnes aveugles sont dans leurs vies quotidiennes souvent amenées à s'orienter en utilisant des sources sonores, et d'autre part, car nous pensons que le son spatialisé permet de présenter une information spatiale de façon immédiatement perceptible, sans introduire de délais liés à un encodage particulier.

Dans un premier temps, nous allons faire un état de l'art synthétique sur les techniques de synthèse binaurale et sur les questions de localisation des sources sonores chez les personnes valides et les personnes déficientes visuelles.

Dans un deuxième temps, nous nous intéresserons à la façon de produire efficacement des stimuli binauraux sur des plateformes embarquées afin de les utiliser dans un dispositif de guidage. Pour cela, nous pourrions nous appuyer sur le savoir faire du groupe audio du CMAP dans le domaine de la spatialisation par filtrage binaural, sans lequel ces recherches n'auraient pu aboutir.

Enfin nous décrivons les expériences que nous avons réalisées pour caractériser les capacités des sujets à localiser des sources sonores. Ces expériences permettent de comparer les capacités de localisation des sources sonores naturelles et des sources sonore spatialisées. En outre nous comparerons plusieurs filtres pour la spatialisation sonore et évaluerons la capacité des sujets à suivre des sources sonores en mouvement.





## Chapitre 4

# Écoute spatialisée et localisation sonore

Pour bien comprendre la question de la localisation sonore, nous allons dans ce chapitre présenter le fonctionnement du système auditif, les indices de localisation et présenter les connaissances actuelles sur les capacités humaines à localiser des sources sonores.

### 4.1 Système auditif

Le système auditif périphérique humain se compose de trois parties : l'oreille externe, l'oreille moyenne et l'oreille interne [205] (cf. Fig. 4.2).

L'oreille externe est constituée du pavillon, du canal auditif et de la membrane tympanique. Cette partie de l'oreille joue un rôle actif dans la perception sonore puisqu'elle modifie sensiblement les sons entrants. En particulier, le couplage avec le canal auditif provoque une série de modes de résonance à moyenne et haute fréquence.

L'oreille moyenne se situe entre l'oreille externe et l'oreille interne. Elle se compose d'une cavité remplie d'air appelée cavité tympanique et comprend les trois osselets et leurs ligaments de fixation, le tube auditif et les fenêtres ronde et ovale. L'oreille moyenne joue principalement un rôle d'adaptation d'impédance entre l'air et l'oreille interne remplie de liquide.

L'oreille interne se trouve dans l'os temporal, dans une cavité complexe appelée *labyrinthe osseux*. Les mouvements de l'étrier sont transmis à la cochlée via la fenêtre ovale et le vestibule. La cochlée est un organe spiralé en forme de coquille. Les cellules ciliaires présentes dans la cochlée au sein de l'organe de Corti, sont responsables de la transformation des signaux mécaniques en stimuli électriques. Elles sont sensibles à la déformation de la membrane basilaire. De part leur position le long de la cochlée, elles agissent comme un analyseur de fréquence : les cellules les plus proches de la base sont plus sensibles aux sons les plus aigus, tandis que les cellules les plus proches de son extrémité le sont aux sons plus graves [57].

Enfin, il faut noter que l'oreille interne contient non seulement l'organe de l'ouïe, mais aussi le vestibule et les canaux semi-circulaires qui sont les organes de l'équilibre.

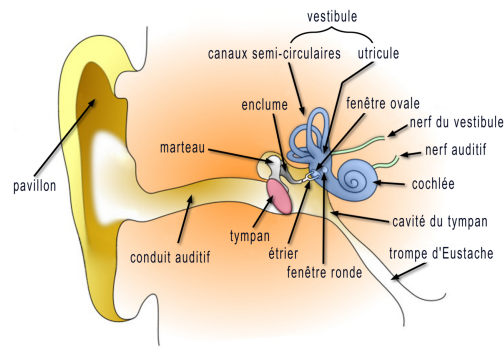


FIGURE 4.1 – Oreille humaine, d'après Dan Pickard et B. Guillot, Wikimedia Common

## 4.2 Référentiel

Dans le domaine de l'audition spatiale, la position d'une source sonore est généralement spécifiée en termes de direction et de distance par rapport au centre de la tête de l'auditeur. Le centre de la tête est généralement défini par le point équidistant des deux oreilles. Il est situé sur l'axe *interaural* qui est la droite passant par le centre des deux oreilles. Le *plan horizontal* (également appelé transversal ou azimutal) est le plan horizontal passant par l'axe interaural séparant le bas et le haut de la tête. De la même façon, le *plan frontal* est un plan vertical passant par l'axe interaural et séparant l'avant et l'arrière de la tête. Le *plan médian* (ou sagittal), quant à lui, est un plan vertical qui sépare les deux demi-espaces gauche et droite et passant par le centre de la tête. Il est perpendiculaire à l'axe interaural.

Plusieurs conventions d'orientation existent mais le système *polaire vertical* (dit *convention rayon-longitude-latitude* ou *repère du géographe*) est le plus communément utilisé. La position de la source sonore est décrite par  $(r, \theta, \phi)$  où  $r$  est la distance par rapport au centre de la tête,  $\phi$  l'angle entre le vecteur direction de la source sonore et le plan horizontal, et  $\theta$  correspond à l'angle entre la projection horizontale du vecteur direction et l'axe frontal.

## 4.3 Indices de localisation

Les théories psychoacoustiques modernes sur la capacité humaine à identifier la direction d'une source sonore sont apparues au milieu du 19ème siècle (voir par exemple [240]). En 1907, Lord Rayleigh [191] présenta la théorie du duplex qui tente d'expliquer la localisation des sons par les *indices interauraux*. En 1939, Hens Wallash identifia des *indices dynamiques* en expliquant comment le mouvement de la tête limite les confusions avant-arrière [221]. Enfin, dans les années 60, on a commencé à comprendre le rôle joué par le pavillon de l'oreille [20] et on a alors identifié les *indices spectraux*.

### 4.3.1 Différence Interaurale de temps (*Interaural Time Difference* ou *ITD*)

Il s'agit d'un indice binaural (ayant trait aux deux oreilles) qui correspond à la différence de temps d'arrivée d'une onde sonore à chacune des deux oreilles. On peut le définir formellement [185] par la valeur

$$\Delta T = \Delta d / v \quad (4.1)$$

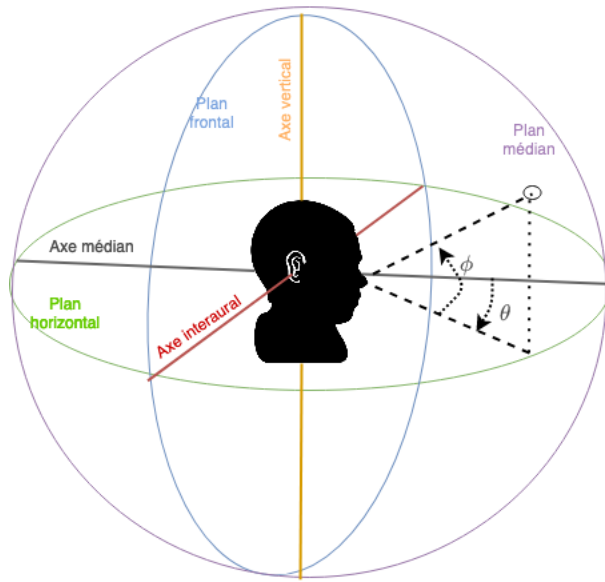


FIGURE 4.2 – Référentiel, plans et axes remarquables

où  $\Delta d$  est la distance supplémentaire (en m) que le son doit parcourir pour atteindre la deuxième oreille après avoir atteint la première, et  $v$  est la vitesse du son (en m/s). Considérant une vitesse du son de 340m/s,  $\Delta T$  vaut environ  $29\mu s$  pour chaque cm supplémentaire que le son doit parcourir. On peut considérer que pour une source sonore située dans le plan médian, cet indice est nul dans la mesure où les deux oreilles sont équidistantes de la source. De même, si l'on considère le son comme une onde plane (i.e. dont la source à l'infini) dans le plan horizontal et que l'on modélise la tête humaine par une sphère de rayon  $r$  sans pavillon auditif, l'ITD pour un angle azimuthal  $\theta$  peut être approchée par la formule de Woodworth [235] :

$$ITD(\theta) = \begin{cases} \frac{r}{c}(\sin \theta + \theta) & \text{pour } 0 \leq \theta \leq \frac{\pi}{2} \\ \frac{r}{c}(\sin \theta + \pi - \theta) & \text{pour } \frac{\pi}{2} \leq \theta \leq 3\frac{\pi}{2} \\ \frac{r}{c}(\sin(3\pi - \theta) - 2\pi + \theta) & \text{pour } 3\frac{\pi}{2} \leq \theta \leq 2\pi \end{cases} \quad (4.2)$$

où  $c$  représente la célérité du son dans l'air. La Figure 4.3 représente l'ITD obtenue par la formule de Woodworth en fonction de l'azimuth  $\theta$ .

Une formule similaire, tenant compte de l'élévation  $\phi$  a été proposée Larchet et Jot [140] :

$$ITD = \frac{r}{c}[\arcsin(\cos \phi \sin \theta) + \cos \phi \sin \theta]. \quad (4.3)$$

L'ITD est un indice fondamental pour la localisation des sources sonore dont la longueur d'onde est supérieure au diamètre de la tête [199], c'est-à-dire pour un son pur de fréquence n'excédant pas environ 1500Hz. Dans la pratique, pour les sons de la vie quotidienne, l'ITD peut rester un indice utile au delà de cette fréquence grâce à la perception de la différence de temps d'arrivée de l'enveloppe sonore [241].

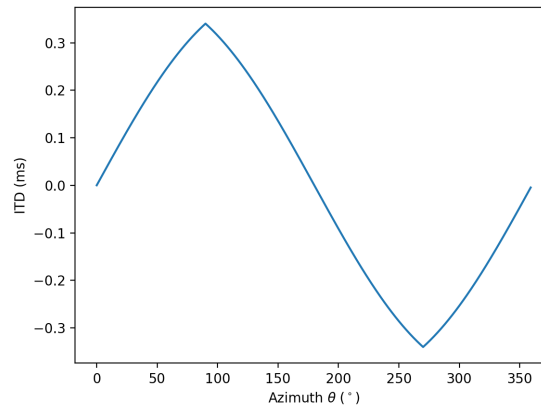


FIGURE 4.3 – ITD obtenues à partir de la formule de Woodworth.

### 4.3.2 Différence Interaurale de niveau (*Interaural Level Difference* ou *ILD*)

Il s'agit de la différence de niveau sonore perçue entre les deux oreilles. Cette différence est due à l'effet d'ombrage de la tête. L'oreille ipsilatérale qui fait face à la source entend le son plus fort que l'oreille controlatérale située de l'autre côté. Pour les longueurs d'onde plus courtes que le diamètre de la tête, cet effet d'ombrage diminue partiellement l'énergie acoustique par réflexion et absorption. Par contre, l'ILD est pratiquement nulle en dessous de 1500 Hz. L'ILD varie fortement avec l'azimut, et naturellement, plus l'incidence est latérale, plus l'ILD est importante. La plus petite ILD perceptible est de 0,5 dB [169].

### 4.3.3 Cônes de confusion

Pour toutes les positions de la source sonore situées dans le plan médian, les ITD et les ILD sont nulles, et en conséquence il est a priori impossible de dissocier des sources sonores en utilisant ces indices. Il existe également des positions où ces indices sont identiques mais non nuls (dans le cas d'une tête sphérique). Ces positions décrivent des cônes appelés *cônes de confusion*. Pour chaque azimut il existe un cône dont la génératrice est l'axe interaural sur lequel les ITD et ILD sont constants.

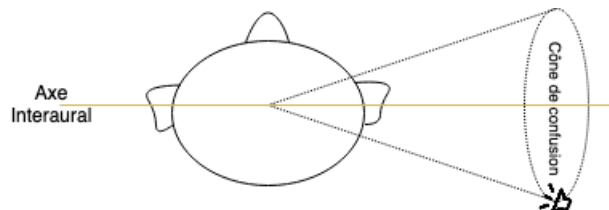


FIGURE 4.4 – Illustration d'un cône de confusion. Le long du cercle de droite, les ILD et ITD sont identiques par symétrie axiale de la figure autour de l'axe interaural.

### 4.3.4 Indices dynamiques et spectraux

Pour tenir compte des limites de la théorie du duplex, en 1939 Hans Wallach [221] émet l'hypothèse que les changements d'ITD et d'ILD induits par les mouvements de tête pourraient constituer un indice de localisation dynamique. Depuis, des expériences ont montré que la rotation de la tête autour d'un axe vertical était effectivement essentielle pour résoudre l'ambiguïté avant-arrière dans la localisation horizontale [231].

Les ITD et ILD ne sont utiles que pour la localisation sonore latérale. Or les être humains ont également des capacités de localisation sur l'axe d'élévation [31]. Cette capacité peut être expliquée par les déformations spectrales liées aux réflexions et à la diffraction dues au pavillon de l'oreille [21], à la tête et au torse. Ces caractéristiques, dépendantes de la direction (cf. Fig. 4.5), joueraient également un rôle dans la localisation avant-arrière en plus des indices dynamiques.

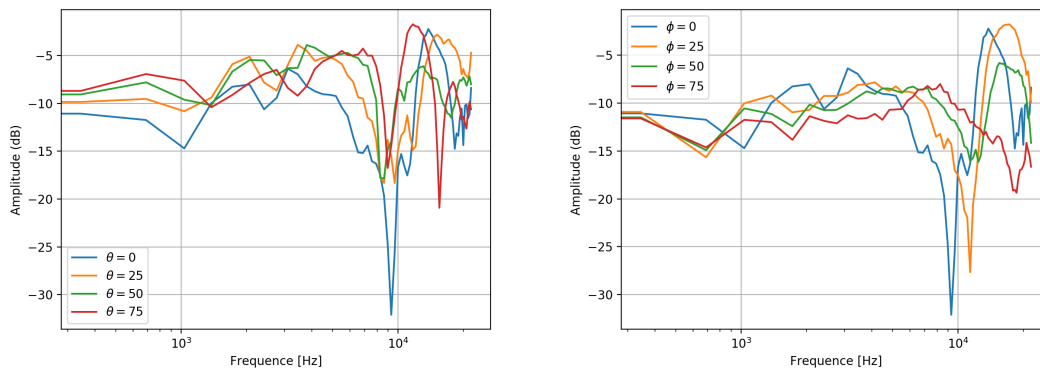


FIGURE 4.5 – Filtrage en fréquence, Gauche : variation en azimuth avec  $\phi = 0$ , Droite : variation en élévation avec  $\theta = 0$ . Données, base IRCAM Listen (sujet 1040).

## 4.4 Capacités humaines de localisation sonore

### 4.4.1 Perception de la direction de la source sonore

#### Localisation sur le plan azimuthal

La précision de localisation horizontale est fortement dépendante de la position. Avec des *bursts* de bruit blanc<sup>1</sup>, selon Blauert [31], l'incertitude de localisation est de 3-4° pour les sources frontales (au voisinage de l'azimut 0°), 5-6° pour les sources situées derrière la tête sur l'axe frontal (azimut 180°), et environ 10° pour les sources latérales (azimut 90° ou 270°). D'une manière générale, la localisation est meilleure pour les sons à large bande. Pour les sons à bande étroite, la précision est supérieure pour les sons centrés sur des fréquences inférieures à 1000 Hz ou supérieures à 3000 Hz [242] ce qui peut s'expliquer par l'insuffisance des repères binauraux aux fréquences intermédiaires. Le niveau sonore par contre ne semble pas affecter la précision de la localisation de façon notable [239].

1. Salves successives de stimuli de quelques dizaines à centaines de millisecondes entrecoupés de silence également courts.

### Localisation dans le plan médian

Selon Blauert, pour la localisation de la parole, la précision est d'environ 9° pour une élévation de 0°; 13° pour  $\phi=75^\circ$  et de 15° à 22° sur l'arrière de la tête.

La localisation sur le plan vertical est essentiellement due aux indices spectraux. Elle est plus facile pour les hautes fréquences et les sons large bande, typiquement au dessus de 7 kHz [88].

### 4.4.2 Perception de la distance

En champ libre, l'intensité sonore reçue d'une source sonore ponctuelle lointaine dépend de l'intensité de la source  $I_0$  et de la distance  $r$  selon

$$I \sim \frac{I_0}{r^2}. \quad (4.4)$$

Sans référence sur l'intensité absolue  $I_0$  de la source sonore, la distance  $r$  reste ambiguë. Par ailleurs, l'impédance acoustique de l'air variant avec la fréquence, le spectre sonore perçu pourrait fournir un indice de distance supplémentaire puisque les hautes fréquences sont plus atténuées que les basses fréquences au cours de la propagation du son. Une fois encore, cette indication reste ambiguë sans référence au spectre initial du son. Les capacités de perception de la distance sont donc limitées.

En environnement réverbérant, quand la source sonore s'éloigne de l'auditeur le rapport d'énergie entre le signal direct et le champ réverbéré diminue et il a été montré que ce rapport pouvait constituer un indice de perception de la distance [163].

D'un point de vue psychoacoustique, il apparaît que les auditeurs sous-estiment souvent de manière significative la distance aux sources éloignées et surestiment généralement les distances aux sources plus proches que 1 m [243]. Par ailleurs, il a été montré que la précision de la perception était supérieure pour les sons auxquels nous sommes habitués [55].

## 4.5 Perception sonore chez les personnes aveugles

### 4.5.1 Perception de l'intensité et de la fréquence

Pour les tâches simples de discrimination de fréquence ou de niveau sonore, les études comparant les capacités des personnes aveugles congénitales avec les personnes voyantes n'ont généralement pas montré de différences notables [24, 174].

Pour les tâches de perception plus complexes cherchant à déterminer la résilience au bruit [174] ou la discrimination fréquentielle utilisant la mémoire auditive [196] les personnes aveugles congénitalement semblent disposer d'aptitudes sensiblement supérieures.

Notons que les musiciens obtiennent généralement des performances supérieures à ces tests, et les personnes aveugles ont souvent une expérience musicale plus importante que la moyenne.

### 4.5.2 Localisation sonore chez la personne Déficiante Visuelle (DV)

Les aptitudes des personnes aveugles pour la localisation sonore sont sujettes à discussion. A priori, en l'absence de retour visuel, qui joue un rôle important dans le développement de l'audition spatiale, les capacités de localisation des personnes aveugles pourraient être altérées. Cependant, certaines études ont conclu à des aptitudes supérieures par rapport au normo-voyant pour la localisation horizontale [142]. Ces capacités supra-normales semblent plus nettement marquées dans l'espace périphérique [220].

Paradoxalement, sur le plan vertical, certains auteurs [143, 248] ont montré des capacités inférieures chez les personnes aveugles .

### 4.5.3 Echolocalisation

Certaines personnes aveugles ont développé la capacité de percevoir l'environnement en détectant les obstacles d'une certaine taille sans les toucher. Cette capacité est parfois qualifiée de *Sens des masses*. Il est aujourd'hui largement reconnu que cette capacité est liée au sens auditif. Les personnes pratiquant cette technique sont capables de percevoir et d'interpréter la réverbération et les échos et d'en tirer des informations sur l'espace environnant. Cette technique peut être pratiquée de façon passive (utilisant les bruits ambiants) ou active avec des clics de la langue ou des claquements de mains. Elle est parfois enseignée dans les apprentissages de locomotion. Bien que cette pratique soit également accessible aux personnes normo-voyantes, elle nécessite un apprentissage poussé, et il existe des preuves d'un meilleur traitement de l'écho et d'une meilleure utilisation des signaux d'écho à des fins de navigation chez les personnes aveugles [210].

### 4.5.4 Des performances supérieures aux personnes valides ?

Il est souvent admis que par un processus de compensation sensorielle, les personnes aveugles développent leurs sens auditif et tactile. Au vu des études en psychoacoustique, il apparaît que les supposées capacités *supra-normales* des individus aveugles par rapport aux individus normo-voyants ne deviennent vraiment significatives que dans des situations complexes. D'une manière générale, on peut imaginer que par l'apprentissage lié à une formation et une concentration prolongées, les personnes aveugles acquièrent une meilleure expérience de la collecte d'informations sur le monde environnant par le biais de l'ouïe en interprétant plus efficacement les caractéristiques des sons.

Dans le cadre de la conception d'un dispositif de guidage, en plus des capacités listées ci-dessus, on peut noter l'aptitude à comprendre la parole synthétique ultra-rapide (jusqu'à 25 syllabes par seconde à comparer aux 8 à 10 syllabes par seconde chez les normo-voyants) [104] souvent utilisée dans les interfaces homme-machine (IHM) pour aveugles.





## Chapitre 5

# HRTF et son spatialisé

### 5.1 HRTF

Les ITD, ILD et les indices spectraux peuvent être inclus dans une unique fonction de transfert appelée *fonction de transfert relative à la tête* ou *Head-Related Transfer Function - HRTF*. Par conversion dans le domaine temporel, on obtient la *Head-Related Impulse Response - HRIR* correspondante. Les paires de HRTFs,  $H_l(r, \theta, \phi, a)$  et  $H_r(r, \theta, \phi, a)$  associées à chaque oreille sont fonctions de la fréquence  $f$ , de la position de la source sonore en termes de distance  $r$ , d'azimuth  $\theta$ , de l'élévation  $\phi$  et de la morphologie du sujet  $a$ .

Les HRTFs peuvent être obtenues par mesures sur des personnes ou sur des mannequins. En général, l'erreur de mesure sur une tête artificielle est moindre, mais les mesures sur les sujets humains permettent d'obtenir des HRTFs individualisées. Les HRTFs sont généralement mesurées dans une chambre anéchoïque afin de minimiser l'influence de la réverbération sur la réponse mesurée. Pour une source sonore suffisamment éloignée, les HRTFs sont des fonctions indépendantes de la distance. Les HRTFs sont donc souvent réalisées avec des sources à distance constante souvent supérieure à 1m. Ainsi les sujets sont généralement placés au centre d'un dispositif qui déplacera une source audio sur une sphère centrée sur leur tête. Les HRTFs sont mesurées avec des incréments de quelques degrés. La mesure se fait généralement en plaçant des microphones peu après l'entrée du canal auditif (canal bloqué) ou au niveau des tympans sur une tête artificielle.

Depuis les années 1970, de nombreuses équipes de recherche ont entrepris des mesures de HRTF en champ lointain sur des groupes de sujets ou des mannequins constituant des bases de données souvent mises à disposition pour la communauté de recherche.

Parmi les bases de HRTFs ayant les meilleures résolutions spatiales, les HRTFs mesurées sur des sujets humains ont généralement une résolution jusqu'à 5° (CIPIC, RIEC, ITA) alors que les bases de données mesurées sur des têtes artificielles ont des résolution jusqu'à 2° voire 1° (TU-Berlin, Kemar). Le tableau 5.1 donne un panorama des bases de données publiques de HRTF, et plus particulièrement des bases à haute résolution spatiale. La Figure 5.2 permet de visualiser des exemples de données en temps / fréquence / phase de sujets extraits de ces bases.

Base /origine	Contenu	Méthodologie	Distance	Résolution spatiale
CIPIC U.C.Davis 2004 [6]	45 sujets	Salle acoustiquement traitée 44.1-KHz, 200 échantillons Format Matlab	1 m	5° en $\theta$ et $\phi$
Listen [226] IRCAM 2003	51 sujets	Anechoic chamber, 44.1 KHz, 512 échantillons (Compensé) Format Matlab	1.95 m	15° en $\theta$ et $\phi$
ITA [32], RWTH Aachen 2016	48 sujets	Semi-anechoïque 44.1 KHz, 256 échantillons Format Sofa	1.2 m	5° en $\theta$ et $\phi$
ARI [10] Austrian Ac. of Sciences 2016	170 sujets	Semi-anechoic 44.1 KHz, 256 échantillons bande passante 200 à 16000 Hz Format Sofa	1.2 m	5° en $\theta$ et $\phi$ 2.5° en $\theta$ sur la zone frontale
Fabian et Kemar Tu-Berlin 2017 [35]	2 dummy	Semi-anechoïque Format Sofa	50cm 1m 1.5m	2° en $\theta$ et $\phi$
Bili [47] IRCAM	54 suj. 3 dummy	Anéchoïque, 96KHz, 2048 échantillons Format Matlab	2m	2° et 1° en $\theta$ et $\phi$
RIEC [227] Tohoku University 2014	105 sujets	Anéchoïque, 48KHz, avec 512 échantillons 44KHz, 256 échantillons Format Sofa	1.5m	5° en $\theta$ et 10° en $\phi$

TABLE 5.1 – Exemples de bases de données de HRTF et leurs caractéristiques.



FIGURE 5.1 – Installation de mesure des HRTF à l'IRBA (extrait de [8])

## 5.2 Décomposition ITD + filtre à minimum de phase

Les HRTFs peuvent être décomposées dans une composante de phase minimale et une composante de phase excessive [176]. La composante d'excès de phase pouvant elle-même être décomposée en une composante à phase linéaire et un filtre passe-tout.

$$H(j\omega) = |H(\omega)| \cdot e^{j\phi_{\min}(\omega)} \cdot e^{j\phi_{\text{lin}}(\omega)} \cdot e^{j\phi_{\text{all pass}}(\omega)}$$

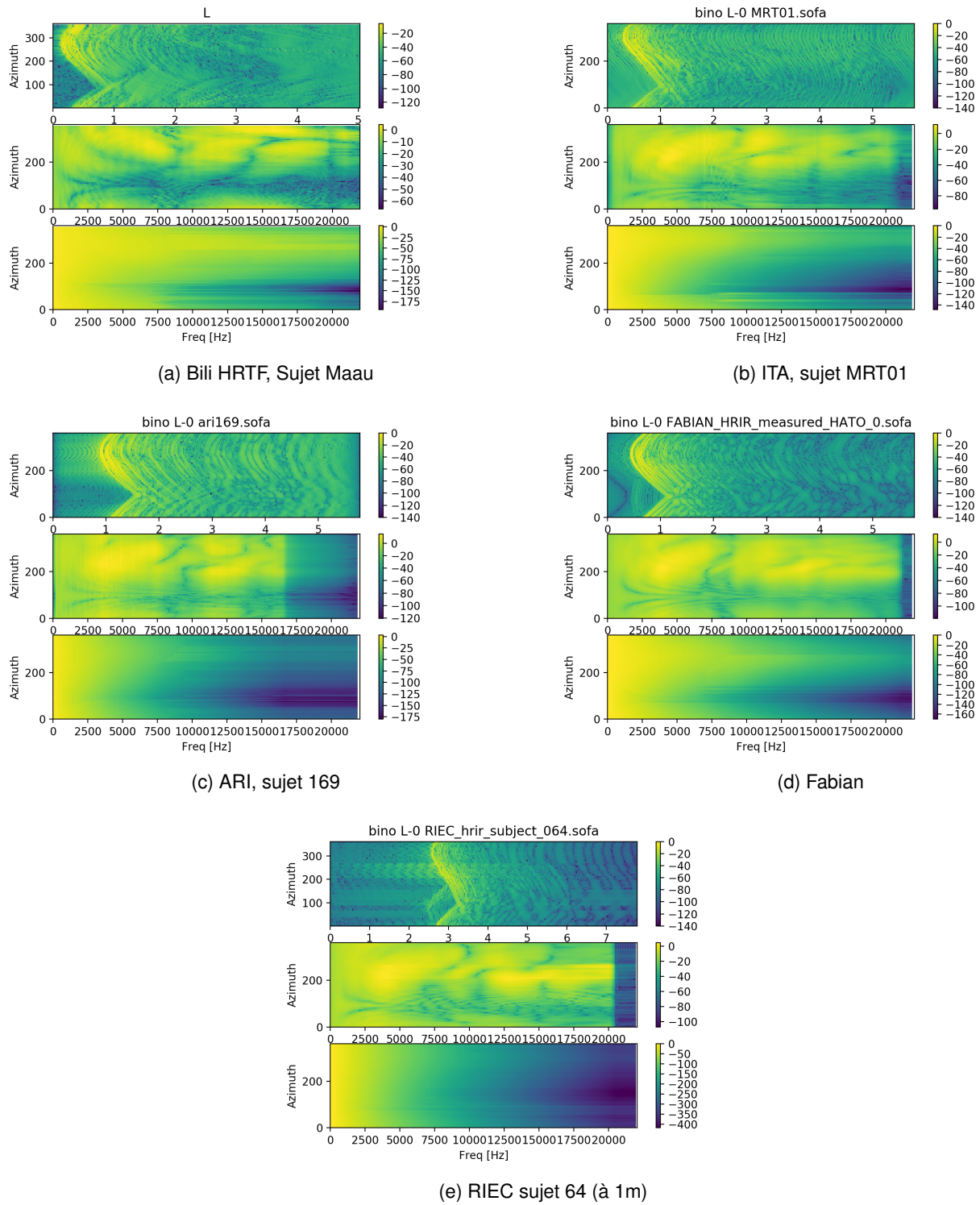
. La composante passe-tout pouvant a priori être négligée [173] et la composante à phase linéaire pouvant être assimilée à un délai, on peut considérer que les HRTFs peuvent être décomposées en un délai (i.e. l'ITD) suivi d'un filtre à phase minimale.

### Extraction des ITD

Plusieurs méthodes ont été proposées pour extraire les ITDs : Cross-corrélation inter-aurale [129], détection de seuil ou évaluation du délai de groupe [125, 140, 170]. Notons que la notion d'ITD est ambiguë : quand l'onde suit un trajet multichemin – typiquement dans le cas contralatéral – on peut distinguer plusieurs pics, de niveaux d'énergie comparables, et correspondant à différents temps d'arrivée.

D'un point de vue pratique, les méthodes de détection de seuil sont simples et globalement efficaces. Il s'agit soit de repérer un seuil dans la HRIR pour distinguer le signal arrivant du bruit (typiquement 10 à 20% du signal peak), soit d'utiliser l'enveloppe du signal, obtenue par lissage). Cette dernière méthode utilisée plus bas est simple mais moyennement robuste quand on doit traiter les données de différentes bases de données. Un sur-échantillonnage de la HRIR permet de gagner éventuellement en précision.

La Figure 5.3 montre l'amplitude moyenne des HRTFs de Fabian pour une élévation  $\phi = 0$ .

FIGURE 5.2 – Comparaison de plusieurs bases de données de HRTF sur le plan  $\phi = 0$ .

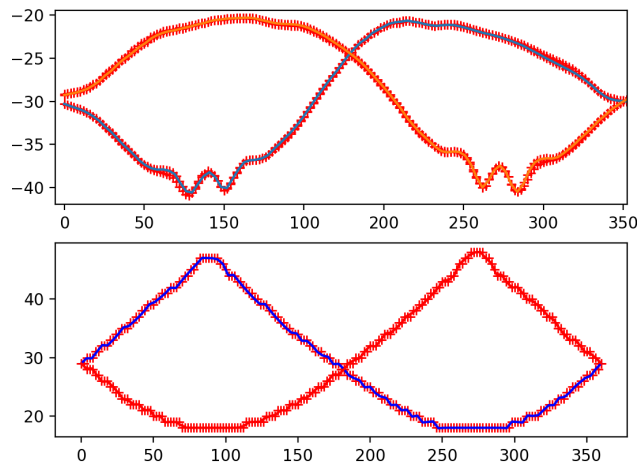


FIGURE 5.3 – Fabian : ILDs (moyennes RMS) en haut et ITDs (pics) en bas, sur le plan azimutal ( $\theta$  en abscisse),  $\phi = 0$ .

### Extraction des ILDs

Par nature, les ILDs sont fortement liées à la fréquence de la source. Une façon d'évaluer les ILDs peut consister à quantifier le niveau moyen de la HRIR, par exemple en considérant la valeur RMS de la RIR. Cependant cette approche peut conduire à surestimer l'ILD. Une autre approche consiste à utiliser la valeur de crête de la RIR (préalablement lissée). Dans les expériences décrites plus bas dans ce document, nous avons utilisé la méthode de la valeur RMS.

## 5.3 Son spatialisé : Enregistrement et synthèse binaurale

Les sons captés à partir d'une tête artificielle dont les micros sont situés à l'emplacement des tympans, ou enregistrés par des micros placés dans le canal auditif (canal bloqué), intègrent toutes les caractéristiques d'une écoute naturelle dans un environnement donné. En conséquence, restitués à l'aide d'un casque ils permettent à l'auditeur de percevoir une sensation comparable à celle de l'écoute dans les conditions réelles et en particulier de percevoir les sources sonores dans l'espace. Des techniques de *Cross-talk Cancellation* peuvent également être utilisées pour une restitution par haut parleur, mais leur mise en œuvre est délicate en pratique.

La convolution d'un signal monophonique par chaque filtre HRTF (gauche et droite) permet de générer un signal stéréophonique pour restituer artificiellement la perception spatialisée. La source sera alors perçue comme provenant d'une direction donnée. Ce principe s'appelle la *Synthèse binaurale*.

**Head tracking** Les études psychoacoustiques ont montré l'importance de la dynamique dans la perception de l'espace sonore [221, 231], ainsi les techniques de synthèse binaurale peuvent inclure un dispositif de suivi de la tête pour plus de réalisme, appelé *head tracking*. Il a pu être montré que ce type de dispositif augmentait sensiblement le sentiment d'externalisation du son et réduisait les confusions avant-arrière [23].

Les dispositifs de head-tracking utilisés sont typiquement des IMU à bas coût tels que ceux présentés en première partie de ce document. Cependant, contrairement aux dispositifs utilisés pour la navigation à l'estime, les critères de

choix seront la latence de mesure et la stabilité. La précision absolue devenant un critère plus secondaire. La latence de tracking est importante car il a été montré qu'elle devenait perceptible autour de 60-80ms [38]. Or cette latence s'ajoute à l'ensemble de la chaîne de traitement. Il est donc important de la limiter à quelque milli-secondes. Ces aspects seront étudiés plus loin dans ce document.

**La question des HRTFs individualisées :** Les HRTF, fonctions de la morphologies de l'auditeur, sont par nature individuelles. Cependant, les difficultés liées à la mise en œuvre des techniques de mesures de HRTF individuelles rend l'utilisation de HRTF individuelles difficiles. L'utilisation de HRTF non individuelles (c'est-à-dire qui n'utilisent pas les filtres correspondant à la morphologie de l'auditeur) a pour conséquence de limiter sensiblement la sensation d'externalisation et de réduire la précision de localisation des sources. En particulier il a été montré que les HRTF non individualisées conduisaient à une augmentation notable des confusions avant arrière et à une baisse des performances de localisation en élévation, en revanche certaines études tendent à montrer que les effets sur la localisation en azimuth restent limités [229, 230]. Pour pallier ces effets il est envisageable d'individualiser les HRTFs. Ces méthodes consistent alors soit à sélectionner les filtres adaptés à l'auditeur (en utilisant des méthodes perceptives, ou par l'analyse de la morphologie), soit à modifier des HRTFs existantes ou bien à utiliser la synthèse directe de HRTFs depuis des données morphologiques. L'individualisation des HRTFs constitue un sujet de recherche en soi qui sortirait du champ de ces recherches. Ainsi, pour ces travaux de guidage des personnes dans le plan azimuthal, nous avons choisi d'utiliser uniquement des HRTFs génériques. L'étude des performances en orientation dynamique et en guidage avec des HRTFs individualisées pourra faire l'objet de travaux supplémentaires pour évaluer leur intérêt.

# Chapitre 6

## Filtrage binaural

La génération de son 3D en environnement anéchoïque consiste généralement à appliquer une paire de filtres (HRTF) sur un signal monophonique. De nombreuses implémentations ont été proposées depuis l'apparition d'ordinateurs puissants ou de processeurs de traitement du signal depuis les années 2000 (SLAB [167], 3D Tune-In [56], SoundScape [90], Ircam Spat [4]). Un moteur audio 3D dénommé *MyBino* [13, 164, 198, 208] a également été développé au CMAP par François Alouges, Matthieu Aussal et François Salmon, en collaboration avec le CNSMDP, avec un focus sur la qualité audio pour le mixage et le monitoring audio professionnel. Les travaux décrits ci-dessous sont directement issus de ce moteur audio.

### 6.1 Implémentation

Pour pouvoir équiper un dispositif autonome, notre moteur audio devra pouvoir fonctionner à la fois sur une plateforme embarquée et sur un ordinateur portable. D'un point de vue fonctionnel, l'accent doit être mis sur la précision spatiale plutôt que sur la fidélité de restitution sonore. Par ailleurs, en contexte dynamique rapide il conviendra de réduire la latence autant que possible.

#### 6.1.1 Algorithme

Pour des raisons de performance, le filtrage est effectué dans le domaine fréquentiel (par convolution rapide en utilisant la FFT). La convolution est effectuée en pseudo-temps-réel avec l'algorithme Overlap-Add [11] qui est classique en traitement du signal. Si l'on considère un filtre de taille  $M$ , et le signal d'entrée divisée en blocs de taille  $L$ . Les blocs sont complétés avec  $M - 1$  zéros (*zero-padding*), puis convertis dans le domaine fréquentiel et multipliés par la transformée de Fourier du filtre. On revient dans le domaine réel en appliquant une transformée de Fourier inverse. On obtient alors plus d'échantillons que dans le bloc d'entrée. Ainsi pour obtenir le signal de sortie, les blocs surnuméraires du calcul effectué sur le bloc précédent sont ajoutés à ce résultat. On conserve les nouveaux blocs surnuméraires pour l'étape suivante et ainsi de suite.

Cette méthode de filtrage dans le domaine fréquentiel réduit la complexité des calculs mais introduit un retard lié au calcul par blocs de la Transformée de Fourier Discrète (TFD). Un bloc entier de TFD doit être mis en mémoire tampon dans le système de traitement du signal avant que le calcul proprement dit ne puisse commencer. Ainsi Overlap-add



induit une latence de  $L - M + 1$  échantillons. Par ailleurs la taille de ce buffer ne peut pas être plus petite que la taille du filtre et, étant donnée la complexité de l'opération en  $\mathcal{O}(n \log n)$  il peut être nécessaire de trouver un compromis entre latence et temps de calcul pour rendre possible le traitement temps réel sur les plateformes peu performantes (en notant qu'augmenter la taille du buffer peut être contre productif considérant qu'il faudra traiter des zéros ajoutés qui ne contribuent pas au résultat). Dans notre implémentation, la taille du buffer est typiquement de 1024 échantillons soit une latence d'un peu plus de 20 ms à 44KHz. À noter que la transformation des filtres dans le domaine fréquentiel n'est réalisée qu'une seule fois, ils peuvent donc être précalculés et stockés.

### 6.1.2 Choix technologiques

Le moteur MyBino initialement écrit en Matlab a été réécrit en Python pour pouvoir fonctionner sur les plateformes embarquées. Le choix de Python se justifie par la volonté d'utiliser un langage de haut niveau permettant le prototypage rapide. Il est par ailleurs adapté à l'audio temps réel (grâce au wrapper pyaudio) et disponible sur plateforme embarquée, avec en outre une faible empreinte mémoire. Par ailleurs Python est incontournable dans le monde Raspberry et ses dérivés car il est préconisé pour la gestion des GPIO (entrées/sorties matérielles). Le principal inconvénient de python étant ses performances limitées, certains éléments ont été développés en Cython ce qui permet d'augmenter très sensiblement les performances des codes mais nécessite d'abandonner certaines facilités de python (en particulier le typage automatique des variables).

Tout comme MyBino, ce moteur permet, de lire des fichiers audio multicanaux, d'associer une position à chacune des sources (canaux), et de modifier en temps réel la localisation des sources et de la tête.

Dans ce moteur :

- Les traitements peu coûteux en temps sont écrits en Python (comme par exemple la lecture des fichiers audio multi-canaux)
- Les traitements efficacement vectorisables ou utilisant des bibliothèques système via Numpy (par exemple les FFT) sont écrits en python et vectorisés avec Numpy.
- Les autres traitements (e.g. les rotations 3D, les conversions de systèmes de coordonnées, la sélection/copie des HRTF) sont écrits sous la forme d'un module Cython.

Ce moteur a été testé sur PC/Mac, processeur ARM 8 (Raspberry Pi 3) et Intel Edison. Le tableau 6.1 montre les performance du moteur anéchoïque sur différents systèmes embarqués.

	1 source	6 sources	24 sources	#max de canaux (buffer 20ms)
Mac	0.9 ms	2.5ms	6.5 ms (fftpack) - 5.9ms (fftw)	(non testé)
Arm 8 (R. Pi 3)	4 ms	11 ms	non	10-12 canaux
Intel Edison (Atom)	9 ms	13 ms	non	1 - 2 canaux

TABLE 6.1 – Comparaison des performance du moteur anéchoïque sur différents systèmes embarqués.

## 6.2 Latence d'un moteur audio

La latence du moteur binaural est un élément crucial pour le guidage. Les mouvements potentiellement rapides de l'utilisateur nécessitent un très bon couplage head-tracking / génération audio. La question de la latence est d'autant plus sensible que les ressources matérielles d'un système embarqué sont limitées. Il est donc nécessaire de bien comprendre les sources de latence et de pouvoir les mesurer afin d'adapter les paramètres logiciels et d'orienter le choix du matériel. Nous pouvons classer les sources de latence de la façon suivante 6.1 :

- La latence liée au head-tracking (et à la localisation).
- La latence liée au traitement audio
- La latence liée à la gestion audio du système d'exploitation
- La latence matérielle de la carte son

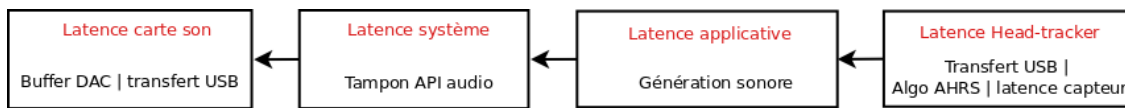


FIGURE 6.1 – Sources de latence d'un moteur binaural head-tracké.

### 6.2.1 Latence audio

De la génération des données audio à l'envoi effectif du signal électrique aux haut-parleurs, le flux audio va transiter par plusieurs tampons qui vont introduire des latences successives. Afin de gérer un flux continu d'échantillons, les systèmes audio vont généralement mettre en œuvre un double tampon : un premier tampon pourra être rempli pendant que l'autre sera lu par le matériel audio. Quand le tampon actif passe sous un seuil de remplissage, la fonction de génération du son *callback* est appelé (c.f. 6.2). Les systèmes audio de MacOS (Core Audio) et linux (Jack, ALSA) implémentent ce type de mécanismes de façon implicite pour le programmeur. À plus bas niveau, les cartes sons possèdent également une mémoire tampon matérielle qui stocke les échantillons enregistrés. Lorsque cette mémoire n'est pas suffisamment pleine, la carte génère une interruption. Le pilote son du noyau va alors transférer des échantillons depuis un tampon d'application. En outre la carte son peut introduire une latence liée au suréchantillonnage (dont l'ordre de grandeur est de 1 à 2 ms).

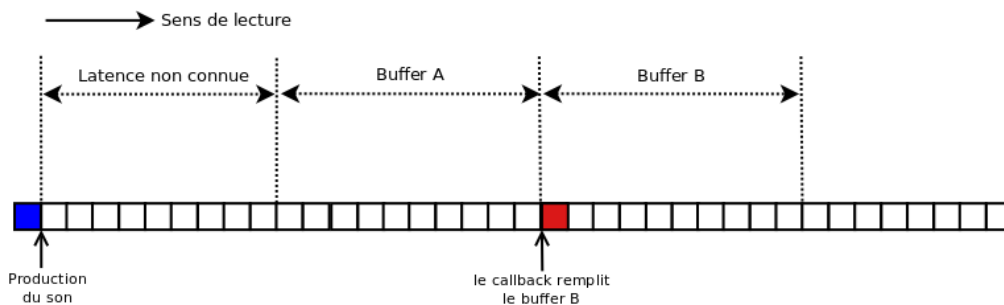


FIGURE 6.2 – Illustration du fonctionnement des buffers audio sous linux (Portaudio/Jack)

## 6.2.2 Temps de traitement des buffers audio

Un système informatique temps-réel doit offrir des garanties sur le temps d'exécution d'une tâche. Ce n'est pas le cas sur les systèmes habituels sur lesquels nous travaillons (linux et MacOS) car l'ordonnanceur ne permet pas de préempter les tâches en toutes circonstances. Il en résulte une importante variabilité dans les temps de traitement des buffers audio qui vont alors devoir être surdimensionnés. Des patches temps réels (RT-linux) ou semi-temps réels (low-latency, preempt-linux) existent pour linux. Ils permettent de réduire les temps de latence maximum, en général au prix d'une baisse globale des performances. Nous avons mesuré les temps de traitement d'un buffer audio sur deux systèmes linux dans deux conditions de charge système (sans charge, et avec un cœur (sur 2) utilisé à 100%)<sup>1</sup>. Dans tous les cas de figure, les temps d'exécution moyens sont de l'ordre de 1ms, mais avec le noyau générique des temps d'exécution supérieurs à 5ms ne sont pas rares (environ 1 cas sur 1000 appels) alors que le maximum est de 2.4ms (machine chargée) dans le cas du noyau à faible latence.

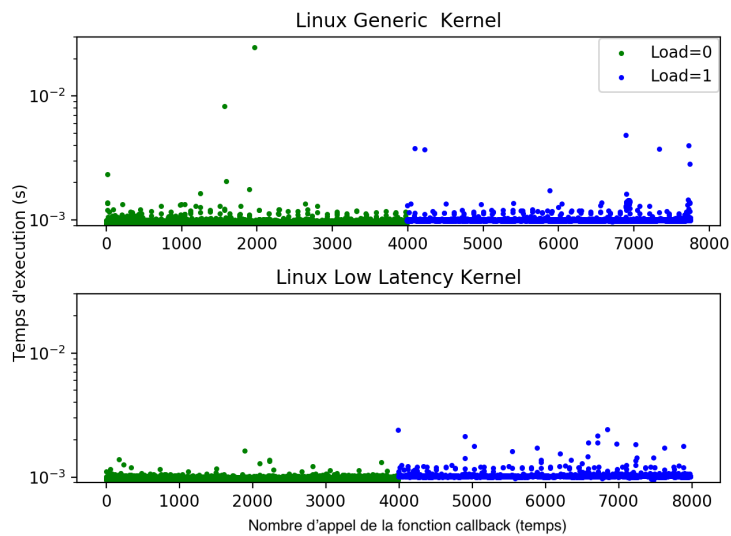


FIGURE 6.3 – Suite de temps d'exécution d'une fonction callback de traitement audio traitant 1024 échantillons avec deux types de noyaux (haut et bas). La partie gauche des graphiques (vert) indique les temps d'exécution pour un système sans charge externe (load 0), la partie droite (bleu) pour un système pré-chargé (load à 1).

## 6.2.3 Mesure des latences système

Déterminer la latence liée aux buffers des systèmes d'exploitation et des drivers audio ainsi que les temps de conversion des DAC<sup>2</sup> est une opération complexe. Les API audio des systèmes peuvent renseigner sur la taille des buffers et la latence qu'ils introduisent, mais il est délicat de connaître précisément la latence de toute la chaîne de conversion DAC. S'il n'est pas possible d'instrumenter les codes à cette fin, on peut utiliser un dispositif matériel pour mesurer le temps écoulé entre un événement externe (une entrée au clavier ou sur un port de communication) et l'émission d'un son. Une telle méthode a déjà été mise en œuvre [216] pour les systèmes embarqués (beagleBone).

1. L'expérience est réalisée avec deux noyaux 4.4, le noyau générique et un autre à faible latence. Pour stabiliser les temps d'exécution, dans tous les cas la vitesse CPU et l'affinité par cœur du processus ont été fixés.

2. Convertisseur Numérique Analogique.

Dans cette recherche, un signal est envoyé sur le GPIO, il déclenche la production d'un son et on mesure le temps à l'aide d'un oscilloscope bicanal.

Nous avons développé et mis en œuvre un dispositif comparable mais d'utilisation plus générique puisqu'il fonctionne sur toutes les plateformes matérielles. Le dispositif est composé d'un microcontrôleur pouvant envoyer un signal sur le bus USB simultanément à un signal logique visualisable sur un analyseur logique. Une fonction est insérée dans le code audio pour générer un son à la réception de la trame USB. En sortie de la carte son, le signal audio produit est ensuite amplifié et seuillé à un niveau logique pour être visualisé sur l'analyseur logique. On peut alors visualiser le signal de déclenchement et le signal audio comme sur la figure 6.4. Cette méthode permet de mesurer la latence sur tous les systèmes possédant un bus USB mais cela introduit une imprécision liée à la latence du bus USB lui-même. Nous avons mesuré ce délai (temps d'aller-retour USB), qui est de l'ordre de 1ms plus ou moins 0.5ms sur les différents systèmes. Les valeurs de latence décrites ci-dessous sont donc corrigées de 1ms.

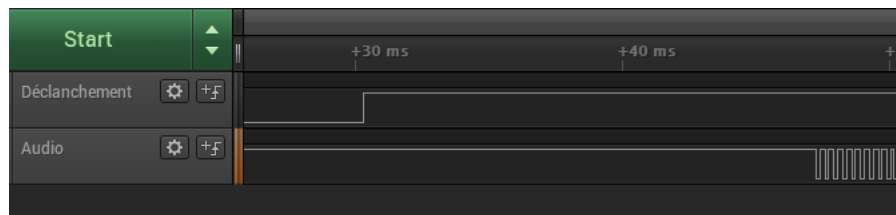


FIGURE 6.4 – Mesure de latence depuis un analyseur logique (Saleae Logic). On peut y visualiser le temps écoulé entre le signal de déclenchement (haut) et le signal audio produit (bas).

Le critère décisif pour optimiser la latence audio sera la taille des buffers applicatifs qui devra être choisie la plus petite possible. Toutefois, un buffer trop petit conduit à des situations de déficit de tampon (appelés *buffer underruns*) et à une surcharge CPU, car cela augmente le nombre d'interruptions.

Le tableau 6.2 montre les latences mesurées sur un Mac et un PC linux. On peut noter que sur chaque système, même pour les tailles de buffers très faibles il existe une latence seuil qu'il ne sera pas possible de réduire.

Taille du buffer applicatif	64	128	256	512	1024
MacOS	10ms	12ms	16ms	30ms	50ms
Linux (noyau standard) x86	20ms	22ms	32ms	40ms	48ms

TABLE 6.2 – Latence totale moyenne mesurée par notre dispositif d'instrumentation pour un traitement audio minimal avec différentes tailles de buffer audio. Comparaison entre MacOS (Catalina, macbook pro 2019, core i5) et linux sur PC (core 2, 2010)

Nous avons également comparé les charges système<sup>3</sup> en fonction de la taille des buffers audio. Les tests sont effectués sur un système linux non chargé<sup>4</sup> qui traite un buffer audio de taille variable. La génération du buffer audio ne nécessite ici pas de traitement spécifique – les échantillons étant stockés en mémoire –, la charge mesurée est donc uniquement liée à la gestion des buffers. On constate (c.f. Table 6.3) que la charge induite par la seule gestion du buffer est importante pour les petites tailles de buffer (typiquement jusqu'à 256 échantillons) ce qui engendre même des situations d'underrun.

3. la charge CPU moyenne représente le nombre moyen de processus en train d'utiliser ou en train d'attendre le processeur. Pour une mesure fiable, nous utilisons la valeur moyennée sur 15 minutes.

4. la charge moyenne sans traitement audio (i.e. au repos) était de 0.02

Taille du buffer applicatif	64	128	256	512	1024
Charge CPU moyenne	0.5	0.4	0.17	0.12	0.1

TABLE 6.3 – Charge système sous linux (moyenne à 15 minutes) pour différentes tailles de buffer audio.

Compte tenu de ces résultats, pour les traitements binauraux nécessitant des latences faibles, nous serons amenés à utiliser des tailles de buffer typiques de 512 échantillons, les tailles inférieures étant de nature à provoquer des charges CPU trop importantes sur les petits systèmes embarqués.

### 6.2.4 Latence du head-tracker

La latence des IMU peut avoir plusieurs sources : la latence des capteurs physiques, les temps liés aux conversions analogiques numériques internes au capteur, le temps de convergence des algorithmes de fusion de données et les délais de transfert vers l'ordinateur. Malheureusement la latence totale des IMUs est rarement présente dans les spécifications constructeur et semble être un sujet peu discuté dans la littérature technique et scientifique : la fréquence de rafraîchissement est souvent la seule métrique considérée, mais cette valeur peut être assez éloignée de la latence réelle.

Pour estimer la latence réelle nous avons mis en place un dispositif simple constitué d'un plateau rotatif qui par un système de tendeur pouvant effectuer un quart de tour en moins de 50ms (nous avons estimé approximativement cette valeur à l'aide d'une caméra 50 img/s). Les IMU sont successivement installées sur le plateau et on mesure leurs réponses. Toutes les IMUs sont configurées pour envoyer à l'ordinateur les données Yaw/Pitch/Roll (le test étant effectué uniquement pour le Yaw) sous forme de messages de 25 octets transmis à 115200 bauds, ce qui représente un temps de transfert d'environ 2ms. Dans la pratique les données sont reçues à des intervalles de 8 à 20ms suivant l'IMU utilisée. En l'absence d'indication sur le temps de déclenchement réel de la rotation, le dispositif ne permet pas de mesurer la latence totale, mais il permet de comparer les IMUs entre elles et d'évaluer les vitesses de convergence des algorithmes de fusion de données.

Les figures 6.6 montrent une grande disparité suivant les IMUs et les algorithmes de fusion associés. On distingue en général deux étapes : dans un premier temps un ajustement très rapide de la valeur mesurée (dont la pente n'est pas mesurable avec le dispositif rudimentaire utilisé) suivi d'une convergence asymptotique vers la position finale. Cette deuxième phase pouvant durer jusqu'à plusieurs secondes sur certains modèles.

Le tableau 6.4 permet de comparer la caractéristique de latence de plusieurs head-trackers, leurs taux de rafraîchissement ( $\Delta T$ ) et les temps pour converger à la valeur finale en considérant les seuils à  $\pm 2^\circ$  et  $\pm 5^\circ$  :

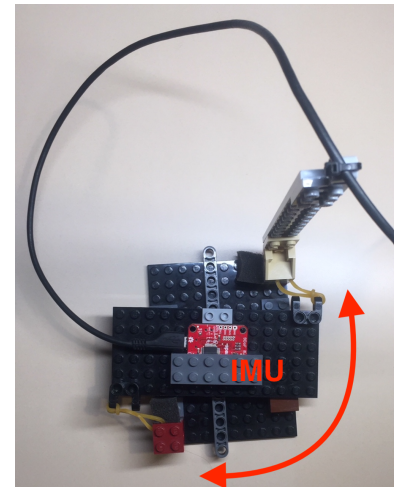


FIGURE 6.5 – Le dispositif utilisé est constitué d'un plateau tournant et deux deux élastiques de rappel. Les élastiques sont tendus puis le système de blocage est retiré pour permettre la rotation rapide du plateau accueillant l'IMU

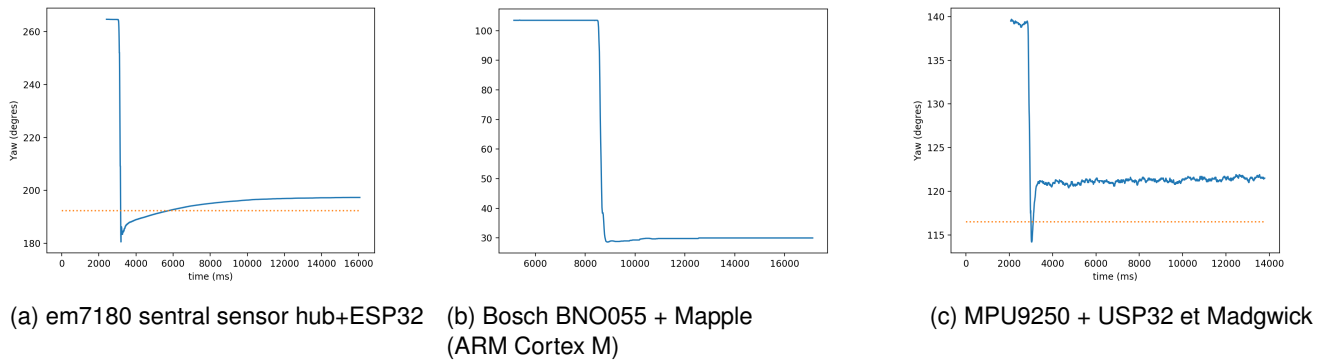


FIGURE 6.6 – Exemples de réponses de différentes IMUs lors d'un mouvement rapide en Yaw.

Capteur	Microcontrôleur	Algorithme	Précision Typique en Yaw	$\Delta T$ mesures	Tps. de stabilisation	
					5°	2°
Bosch BNO055	ARM M3@72MHz	Propriétaire	$\pm 4 - 5^\circ$	20ms	0.2s	0.5s
Em7180 sentral sensor hub	ESP32@240MHz	Propriétaire	$\pm 2 - 3^\circ$	8ms	2.5s	4s
Razor M0	SAMD21@48MHz	DMP (propriétaire)	$\pm 4 - 6^\circ$	10ms	<0.1s	<0.2s
MPU9250	ESP32@240MHz	Madgwick	$\pm 3 - 4^\circ$	8ms	<0.1s	0.2s

TABLE 6.4 – Comparaison de la latence des IMUs.

Sur un temps plus long (quelques dizaines de secondes) certains modèles sont sujets à une déviation, qui est attribuable à un défaut de calibration du magnétomètre. C'est en particulier le cas du Razor M0 qui ne propose aucune procédure de calibration efficace. Ce type de défaut, réhhibitoire, est suffisamment important pour être perçu comme un déplacement de la source sonore. Pour les application de head-tracking, les modèles utilisant l'algorithme de S. Madgwick présentent par exemple un bon compromis stabilité/latence/précision.

### 6.3 Auralisation temps réel pour le guidage audio

La présence de réverbération est un élément important du sentiment d'externalisation à l'écoute d'un son binaural. On sait également que certaines personnes aveugles sont capables d'identifier des éléments de l'espace (taille de la pièce, présence de murs ou d'obstacles, etc.) grâce à la réverbération. C'est pourquoi, sur la base des travaux antérieurs de l'équipe [14, 95] nous avons également développé un moteur permettant l'auralisation temps réel pour un environnement possédant une géométrie simple via la méthode analytique des *sources-images*.

Cette méthode développée par J. B. Allen et A. D. Berkley [7] permet d'estimer la réponse impulsionnelle d'une pièce en forme de pavé droit grâce à une construction géométrique des réflexions sur les parois (représentation des sources, figure 6.7). Cette méthode a par la suite été améliorée [33] pour traiter des géométrie arbitraires, mais nos

travaux utilisent la méthode originale limitée aux pièces rectangulaires.

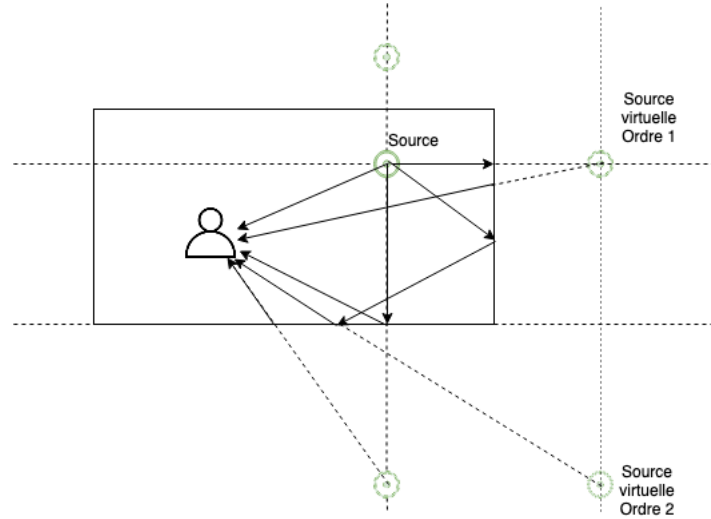


FIGURE 6.7 – Formation des sources image (Adapté de *Lectures on room acoustics*, Ramani Duraiswami, University of Maryland).

La méthode fait l'hypothèse de réflexions spéculaires sur les murs avec absorption partielle. En première approximation, on peut considérer cette absorption comme indépendante de la fréquence, ce qui est connu pour être inexact pour certains matériaux.

L'intensité  $Q_i$  de la source virtuelle  $i$  à la distance  $r_i$  peut être calculée en considérant les coefficients d'absorption  $\beta_s$  du mur  $s$ . Après réflexion sur les différents murs, on a donc :

$$Q_i \propto \frac{\prod_s (1 - \beta_s)}{r_i}. \quad (6.1)$$

On peut alors construire la RIR  $h(t)$  :

$$h(t) = \sum Q_i \delta(t - r_i/c), \quad (6.2)$$

où la fonction delta génère un pic d'intensité  $Q_i$  dans la réponse impulsionnelle, au temps d'arrivée  $r_i/c$ .

Il est alors possible d'utiliser les informations spatiales des sources en intégrant les HRTFs dans la réponse impulsionnelle obtenue afin d'obtenir une BRIR (Binaural Room Impulse Response). D'un point de vue pratique, on peut considérer plusieurs optimisations :

- Au delà de 80ms on considère habituellement que le signal ne transporte plus d'information spatiale perceptible [188, 203], il est donc possible de gérer cette *queue de réverbération* différemment, par exemple en considérant qu'elle est indépendante de la position de l'auditeur ou de la source [247]. On peut alors la précalculer afin de réduire le traitement en temps réel.
- Il n'est pas nécessaire de convoluer chaque source individuellement avec les HRTF (pour les sommer ensuite), en effet, la convolution étant une opération linéaire, on peut construire le filtre en sommant au préalable toutes les HRTFs en respectant les décalages temporel pour effectuer le filtrage une seule fois [247].

Pour un résultat réaliste, la longueur du filtre  $h$  ne devrait pas être inférieure au temps de réverbération de la salle

(c'est-à-dire le temps nécessaire pour que le niveau sonore baisse de 60 dB), soit typiquement de quelques centaines de millisecondes (petite pièce) à quelques secondes (gymnase). Dans ce cas, les techniques de convolution dans le domaine fréquentiel décrites au chapitre précédents introduisent des latences préjudiciables dans une application de guidage. C'est pourquoi nous avons opté pour une approche par convolution partitionnée connue dans la littérature [22] sous l'appellation *Frequency domain delay line*.

La méthode que nous avons implémentée est décrite synthétiquement sur le schéma 6.8. Elle consiste à diviser la réponse impulsionnelle  $h$  en une série de filtres (de tailles identiques) qui seront convertis dans le domaine fréquentiel  $H_0, H_1, \dots, H_{n-1}$ . On utilisera ensuite chaque sous filtre en introduisant le retard correspondant dans la RIR. L'ensemble des produits calculés sont ensuite sommés et convertis à nouveau dans le domaine temporel. L'algorithme est décrit formellement en Alg. 3.

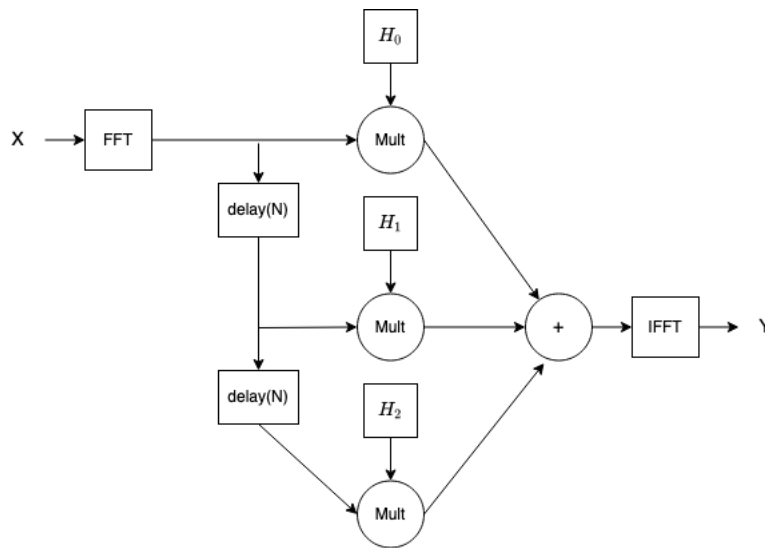


FIGURE 6.8 – Méthode de convolution partitionnée avec sommation dans le domaine fréquentiel (*Frequency Domain Delay Line*), adapté de [22]

---

**Algorithm 3:** Frequency Domain Delay Line avec Overlap-add pour réverbération temps réel

---

```

while Traitement audio en cours do
  — Lit un bloc audio  $X$  de taille  $tailleBloc$ 
  — # Préalcul de la RIR partitionnée uniformément
  — Découpe la RIR en  $n = RirSize/tailleBloc$  blocs  $h_0 \dots h_{n-1}$ 
  — On calcule le spectre de chaque bloc  $H_j$ 
  — Considérant le spectre du bloc audio courant  $X_0$  et les  $n - 1$  spectres des blocs audio précédemment traités
   $X_{M-1}, \dots, X_{M-(n-2)}$ , on calcule la somme :  $Y = \sum_{k=0}^{n-1} H_k \cdot X_{M-k}$ 
  — Calcul de la transformée de Fourier inverse de  $Y$ . Les  $tailleBloc$  premiers échantillons constituent la sortie
  audio, les suivants sont utilisés pour l'overlap à l'itération suivante.
end

```

---

Un découpage *non uniforme* [89] de la RIR est également possible pour traiter des très longs filtres sans impacter la latence, mais nous avons jugé qu'il n'était pas nécessaire de l'implémenter dans notre cas d'usage. En effet, plus que la justesse de la simulation, nous cherchons à obtenir un rendu réaliste pour améliorer le sentiment d'externalisation,



nous nous limiterons donc à des réverbérations de 1 à 2s. En revanche, sur les plateformes embarquées, il s'est avéré que le calcul des sources-images et la construction complète de la RIR était très coûteuse dans les grandes pièces pour les calculs d'ordre élevés (supérieurs à 6-8).

Pour pallier cette difficulté, nous avons développé une méthode qui tire partie de la stabilité de la queue de réverbération dans une pièce donnée. Les techniques qui consistent à construire une RIR en accolant une queue de réverbération pré-calculée à des réflexions précoces directionnelles calculées en temps réel se heurtent à la question de la mise à l'échelle des deux parties et donc à la mise en œuvre d'un modèle de raccordement réaliste. Nous avons pensé qu'il est possible de tirer partie de la méthode de calcul des sources images pour calculer une queue de réverbération suffisamment réaliste pour nos applications. En effet, on peut constater que pour des conditions identiques, si l'on calcule une RIR  $h_{low}$  avec des sources images à un ordre faible (4-6) puis une RIR  $h_{high}$  à un ordre plus élevé (12-14), la contribution des images d'ordres élevés  $h_{diff} = h_{high} - h_{low}$  se situe entièrement dans la queue de la réverbération (i.e. au delà de 80ms) (c.f. Fig.6.9). Ce constat n'est vrai que dans des salles suffisamment grandes, mais c'est précisément dans ces conditions que cette méthode est utile.

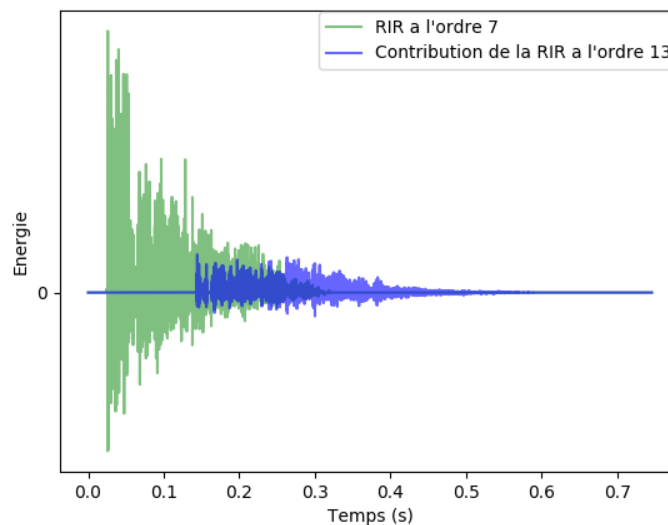


FIGURE 6.9 – Visualisation d'une RIR dans une pièce de taille 7x8x3m avec la méthode décrite ci-dessus. Contribution des sources image calculées à l'ordre 7 et à l'ordre 13 dans le calcul de la RIR.

On peut alors envisager de calculer en temps réel (i.e. dans l'intervalle de temps permis par le traitement du buffer audio) des sources images à un ordre faible et y accoler le filtre  $h_{diff}$  calculé de façon asynchrone en parallèle (typiquement à quelques Hz). La méthode permet en outre de répartir les calculs sur des cœurs différents.

**Implémentation et performances :** Nous avons implémenté en Cython le calcul des sources images et de la génération de la BRIR. Le reste du traitement étant encore une fois codé en Python/numpy. Notre implémentation permet de calculer une réverbération en temps réel sur Mac ou sur une plateforme embarqué Intel haute performance (Udoo Bolt) avec une durée de réverbération de l'ordre de 2s, ce qui semble réaliste pour une grande salle. Sur ARM les performances sont plus limitées : de l'ordre de 150ms sur raspberry pi 3 et 300ms sur odroid XU4 ou Raspberry Pi 4.

## 6.4 Conclusion

Notre travail sur le filtrage binaural a été guidé par les problématiques spécifiques de notre cas d'usage. En effet certaines contraintes usuelles du filtrage binaural comme la préservation du timbre ou de la qualité d'écoute apparaissent secondaire, alors que la question de la latence et des performances CPU sont prépondérantes. Implémenter efficacement un moteur de rendu audio binaural embarqué à faible latence pour le guidage implique une bonne compréhension de la chaîne de traitement, du head-tracking à la restitution audio. Les principales sources de latence étant liés au tracking, aux buffers matériels et logiciels et au traitement binaural lui même. Nous avons pu optimiser l'ensemble de ces paramètres par la sélection du matériel (ex capteurs Bosch ou Inversens/TDK+Magdwick pour le tracking), les paramétrages systèmes et l'implémentation efficace du traitement audio. S'appuyant largement sur les implémentations antérieures de l'équipe, notre le moteur audio anéchoïque est capable de fonctionner sur des plateformes aux performances très réduites et à faible consommation (consommation <5W sur Raspberry pi 3B avec 10 sources spatialisée). Nous avons également pu implémenter un moteur d'auralisation temps réel embarquable permettant une meilleure externalisation. L'intérêt de l'auralisation temps réel pour le guidage (expérience utilisateur, précision, information sur l'environnement) restera à étudier dans de futures recherches.



## Chapitre 7

# Évaluation de filtres pour le guidage

Les études sur les capacités humaines de localisation d'une source sonore sont assez nombreuses et anciennes. Parmi ces travaux, on peut distinguer d'une part des expériences de *localisation sonore relative*, qui consistent typiquement à évaluer l'angle minimum audible entre deux sources identiques situées à la même distance de l'auditeur [99, 169], et d'autre part des tests de *localisation absolue* qui visent à évaluer la capacité du sujet à désigner directement la position d'une source sonore dans l'espace [37, 157]. Ce type d'études a également été mené pour la localisation de sons spatialisés par les méthodes de filtrage binaural [166, 229, 244]. La plupart de ces études portent sur la localisation de sources sonores fixes dans l'espace. En revanche, la capacité des sujets à suivre les changements de position des sources sonores en mouvement a été relativement peu étudiée. Des membres du groupe audio du CMAP ont également menés des travaux sur la perception de filtre audio dégradés [198] ou sur la localisation de stimuli audio binauraux multicanaux [29].

Dans les expériences suivantes, nous chercherons à caractériser les capacités des sujets à localiser des sources sonores statiques mais aussi en mouvement. L'étude porte sur des sources situées sur le plan azimutal puisque nous visons des applications pour le guidage au sol.

Ci-dessous nous allons présenter deux séries d'expériences. Nous souhaitons comparer les résultats obtenus pour la localisation de sons spatialisés au casque avec des sons naturels. Dans chaque cas, nous utiliserons les mêmes stimuli et une méthodologie la plus semblable possible.

## 7.1 Filtres

### 7.1.1 HRTF

Les expériences sur le son binaural sont réalisées avec deux jeux de HRTFs différents. Nous choisissons les HRTFs parmi deux projets aux caractéristiques très distinctes :

- Une jeu de HRTF issu du projet IRCAM Listen [226] :

Nous avons sélectionné le sujet 1040 (sujet humain), reconnu dans la communauté binaurale pour ses qualités d'externalisation [103]. Les HRTFs du projet Listen de l'IRCAM sont mesurées avec une résolution spatiale de 15°. Pour accéder à une résolution plus élevée, nous avons effectué une interpolation à 2° en utilisant les méthodes d'interpolation développées au CMAP [13, 15].

— Les HRTFs du projet Fabian à TU-Berlin [35] :

Ces HRTFs sont mesurées avec une tête et un torse artificiels et ont une résolution spatiale native de  $2^\circ$ , un rapport signal/bruit de 55db (contralatéral) à 80db (ipsilatéral) et une large bande passante (environ 100Hz-20KHz). Nous n'effectuons aucun traitement supplémentaire.

## 7.1.2 Panning

En utilisant les ITD extraites de HRTF mesurées et les ILD (sous forme d'amplitudes moyennes, c.f. chapitre 4) on peut construire un filtre de panning. Un tel filtre peut être visualisé sur la figure 7.1-A.

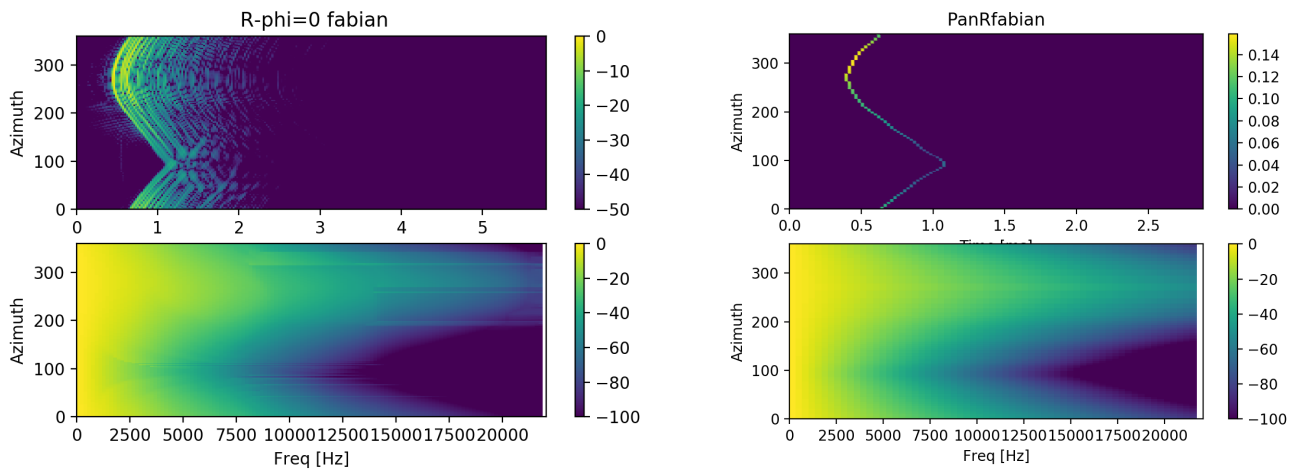


FIGURE 7.1 – Visualisation de la HRIR (R) Fabian et d'un panning dérivé des ITD/ILD selon la méthode décrite au chapitre 4. En haut, diagramme niveau/temps(ms), en bas une représentation phase/fréquence du signal.

La figure 7.2 compare l'estimation de l'ILD ( $\text{Level}_{\text{right}} - \text{Level}_{\text{left}}$ ) pour FABIAN et Ircam-listen-1040 en  $\phi = 0$ . On peut observer une différence d'amplitude des ILD de 4dB qui est susceptible d'avoir une influence sur la localisation des sources avec ce filtre.

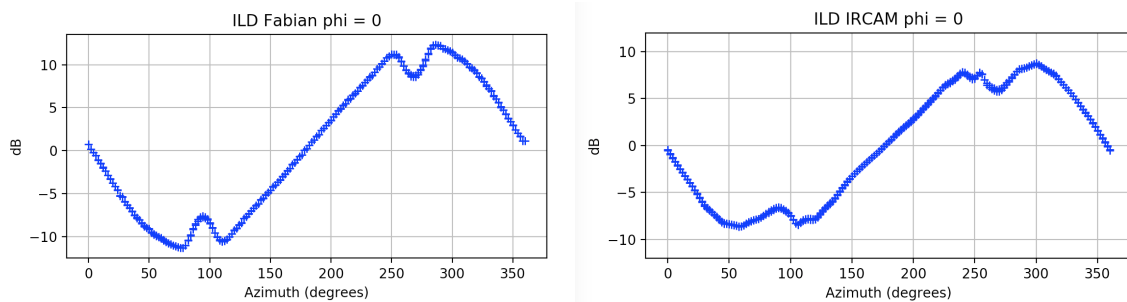


FIGURE 7.2 – Comparaison des ILD Fabian vs. Ircam. Différence de niveau  $\text{Level}_{\text{right}} - \text{Level}_{\text{left}}$  par azimuth à  $\phi = 0$ .

## 7.2 Méthode

Dans nos expériences, nous demandons à chaque sujet de s'orienter vers la source sonore (plutôt que de la pointer du doigt comme c'est souvent l'usage), ainsi c'est l'orientation de la tête qui sert de référence de pointage. Cette technique a été choisie car elle est très naturelle et qu'elle se rapproche des conditions écologiques de suivi de source sonore. Ce *nose pointing* est par ailleurs simple à mettre en œuvre (capteur léger et peu coûteux simplement fixé sur la tête à l'aide d'un bandeau ou d'un casque), fiable et sans risque pour l'utilisateur.

Dans toutes les expériences, les stimuli sonores utilisés seront des bursts de bruit blanc enchaînant 100ms de bruit et 300ms de silence, qui sont couramment utilisés dans les expériences de localisation sonore [31].

Afin de limiter les risques pour les sujets, les niveaux acoustiques sont calibrées pour induire une exposition de moitié inférieure à celle autorisée par la réglementation du code du travail (décret 2006-892 du 19 juillet 2006). Dans l'ensemble des expériences, le niveau sonore des stimuli de localisation est de 85dBa (hors son de masquage). Les participants sont parfaitement informés du dispositif et des objectifs recherchés et sont brièvement entraînés à l'usage du matériel.

La trajectoire de la source sonore a été préalablement générée par une marche aléatoire sur un cercle autour de l'auditeur, elle comporte donc des périodes de mouvements lents et d'autres plus rapides, la vitesse de la source sonore varie de 0 à environ  $20^\circ/s$ . Un exemple de déplacement de la source et de suivi par un participant est donné par la figure 7.3

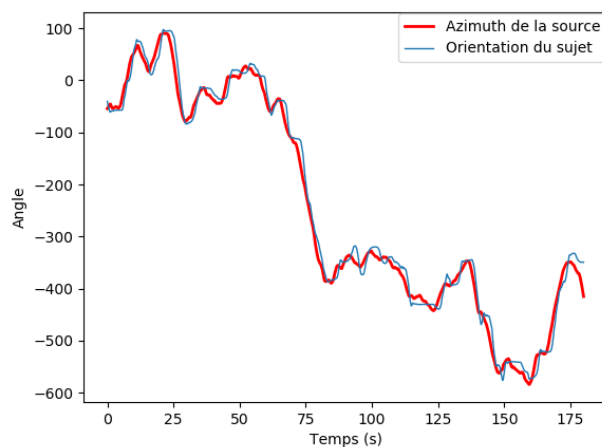


FIGURE 7.3 – Déplacement de la source sonore en azimuth pendant les 3 minutes de l'expérience (courbe rouge). Un exemple d'orientation utilisateur est donné en bleu.

Qu'il s'agisse de filtre de panning ou de filtrage HRTF, le rendu audio des expériences décrites ci-dessous est assuré par le moteur audio anéchoïque présenté au chapitre 4. Les sujets participants aux expériences sont âgés de 20 à 55 ans et ne présentent pas de déficiences auditives connues.

**Identifications des expériences :** Quatre expériences ont été réalisées, avec des sources sonores fixes ou en mouvement, réelles ou rendues par son binaural. Pour les besoins de l'exposition des résultats, nous les avons respec-

tivement nommées suivant la nomenclature suivante. La **Série N** correspondant aux sons naturels et pour lesquels la source peut être soit fixe (**N.Stat.**) soit en mouvement (**N.Dyn.**), et la **Série V** dans laquelle la source est rendue virtuellement en utilisant du filtrage binaural et pour laquelle, de nouveau, la source peut être soit fixe (**V.Stat.**) soit en mouvement (**V.Dyn.**).

Les sujets participants aux expériences de la série **N** sont différents de ceux participants à la série **V**. En revanche un même groupe de sujets participe à toutes les expériences statiques et dynamiques au sein d'une même série.

### 7.3 Expériences N.Stat. et N.Dyn. : Localisation statique et dynamique de sons naturels

Ces expériences ont été réalisées en partenariat avec l'Institut de Recherche Biomédicale des Armées (IRBA Brétigny) qui a mis à notre disposition une salle semi-anéchoïque et nous à apporté un soutien technique (Jean-Christophe Bouy) et scientifique (Guillaume Andéol) en participant à la mise au point du protocole.

Les sujets sont assis sur un fauteuil rotatif, les yeux bandés, dans la salle semi-anéchoïque. Un dispositif composé d'un véhicule roulant sur une circuit circulaire permet d'émettre des sons à l'aide d'un haut-parleur. Ainsi la source sonore se déplace sur un cercle de rayon 1.5m, centré sur l'utilisateur et sur le plan horizontal situé à hauteur des oreilles du sujet. Il est demandé au sujet de s'orienter vers la source sonore de façon à la positionner devant lui.

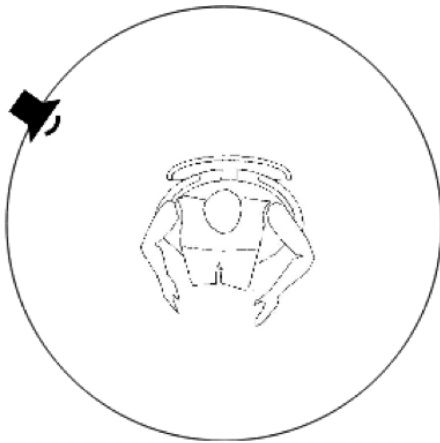


FIGURE 7.4 – À gauche, schéma du dispositif expérimental des expériences **N.Stat.** et **N.Dyn.**. Le cercle sur lequel se déplace le haut parleur est situé sur le plan horizontal au niveau des oreilles du sujet. À droite, une photo du dispositif et d'un sujet.

**L'expérience N.Stat. (source fixe)** se déroule de la façon suivante :

- Le dispositif place un haut parleur à une position aléatoire. Une fois positionné, le dispositif émet un stimulus sonore pendant 5 secondes.
- Le sujet s'oriente vers le son.
- À  $t=5s$ , l'orientation du sujet (i.e. la position de la tête) est mesurée avec le head tracker et l'erreur de positionnement est calculée.
- La séquence reprend.

L'expérience dure 3 minutes, ce qui, compte tenu du délai du positionnement des sources correspond environ à 30 mesures. Pendant les séquences de déplacements du dispositif le long du cercle autour du sujet, les bruits sont couverts par des haut-parleurs disposés dans la salle. Enfin, les séquences de positionnement sont identiques pour tous les utilisateurs.

**L'expérience N.Dyn. (source mobile)** se déroule de la façon suivante :

- Le dispositif fait déplacer un haut parleur sur tout le cercle, de façon aléatoire, continue et régulière. Le haut parleur émet les stimuli en permanence.
- L'utilisateur doit continuellement s'orienter vers la source sonore.
- 10 fois par seconde l'orientation de la tête de l'utilisateur et la position du haut-parleur sont enregistrées et comparées.

L'expérience dure 3 minutes. À la fin de l'expérience, l'erreur cumulée est calculée et l'ensemble des données de mesure est conservé.

### 7.3.1 Notes concernant le dispositif expérimental

Pour permettre un déplacement parfaitement fluide de la source sonore nous avons choisi de construire un dispositif roulant sur la base d'un train électrique et supportant un haut parleur. Les contraintes de poids et d'encombrement liées au transport du haut parleur et l'inertie du dispositif ont conduit à choisir un haut parleur de type JBL flip 3 dont la bande passante mesurée est d'environ 150-12000Hz.

La position du système roulant est évaluée en temps réel par un dispositif composé à la fois d'un réseau de balises RFID positionnées sous les rails et d'une IMU de classe professionnelle Xsens MTI-7 (précision angulaire en yaw :  $1.5^\circ 1\sigma$ ). Le train est contrôlé pour effectuer un parcours parfaitement défini. À cette fin, le dispositif est asservi par un contrôleur PID via une liaison radio. L'ensemble du dispositif est conçu pour minimiser les bruits parasites et en particulier ceux liés au frottement avec les rails.

La position de la tête est captée par une IMU de classe professionnelle xsens MTi-30 (précision yaw typique  $1^\circ 1\sigma$ ). Dans le pire cas, et en négligeant l'apport du système de balises RFID, l'erreur de mesure typique du système de pointage (train + tête) est donc de  $2.5^\circ 1\sigma$ . Ces valeurs ont pu être validées avec un mesureur d'angle Bosch électronique GAM.

### 7.3.2 Résultats

Les expériences ont été réalisées avec 12 participants (11 hommes, 1 femme) âgés de 25 à 55 ans et sans problème auditif connu. En raison des contraintes liées à la crise sanitaire 2020/2021 nous n'avons pas été en mesure d'effectuer ces expériences sur un nombre plus important de sujets.

L'erreur de pointage présentée ci-dessous est définie par la valeur absolue de l'angle formé entre l'orientation du sujet et la source (c.f. éq. 7.1).

$$e_{pointage} = |\theta_{sujet} - \theta_{source}|. \quad (7.1)$$

Dans le cas de la localisation de sources statiques, l'erreur médiane de pointage mesurée est de  $3.6^\circ$  (de  $2^\circ$  à  $5.6^\circ$  suivant les sujets). Dans le cas dynamique l'erreur médiane monte à  $6.5^\circ$ . Les résultats pour chaque sujet sont visibles sur la figure 7.5.



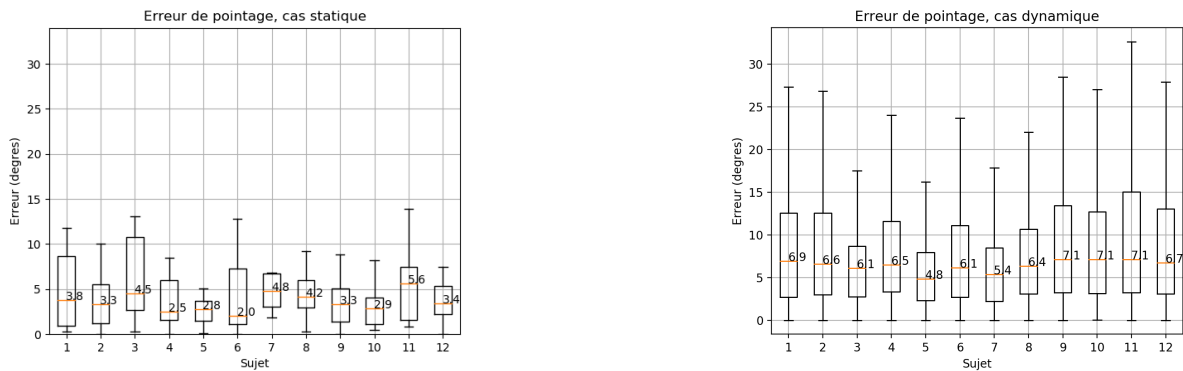


FIGURE 7.5 – Résultats de pointage sonore tel que défini en eq. 7.1 dans le cas des sons naturels statiques (gauche) et en mouvement (droite). Les mêmes sujets participent aux deux expériences et sont identifiés par un même numéro.

On peut également noter que, de manière prévisible, l'erreur médiane augmente avec la vitesse de la source, même si elle reste très contenue (la différence de précision entre les déplacements de sources rapides – i.e. où la vitesse angulaire est supérieure à  $10^\circ/s$  – et lents est d'environ  $2^\circ$ ). L'auditeur a donc des capacités de perception ainsi que les capacités motrices suffisantes pour suivre aisément ce rythme de déplacement.

## 7.4 Expériences V.Stat. et V.Dyn. : Localisation statique et dynamique de sons spatialisés

Ces expériences sont comparables aux expériences de la **série N.** mais cette fois avec des sons spatialisés. Les sujets sont placés dans une salle isolée des bruits extérieurs et aveuglés à l'aide d'un bandeau. Ils sont assis sur un fauteuil rotatif. Les conditions sont donc comparables à celles des expériences 1. Les sujets sont munis d'un casque audio semi-ouvert (Beyer Dynamic DT 990 Pro) équipé d'un head-tracker. Dans toutes les expériences de la **série V.**, le moteur de rendu binaural anéchoïque va permettre de produire des sons spatialisés. Ce moteur est en outre couplé au head-tracker de telle façon que la source sonore garde sa position absolue même lorsque le sujet tourne.

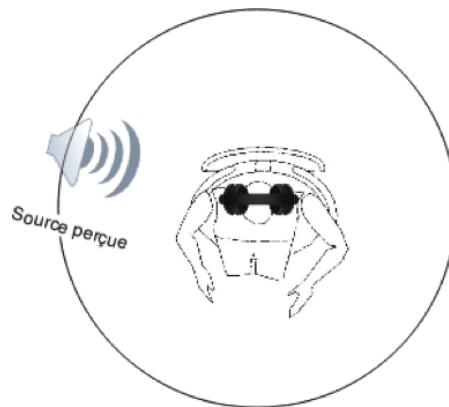


FIGURE 7.6 – Schéma du dispositif expérimental des expériences de la **série V.**

**L'expérience V.Stat. (source fixe)** se déroule de la façon suivante :

- L'utilisateur écoute un son spatialisé simulant une source sonore sur un cercle égocentré de rayon 1.5 m et situé sur le plan horizontal à hauteur des oreilles du sujet. Le stimulus est émis pendant une durée de 5 secondes. La séquence de positionnement pré-définie reproduit le comportement du haut-parleur des expériences de la série 1 avec une réorganisation en sous-séquences différente à chaque expérience.
- Le sujet s'oriente vers le son.
- A  $t=5s$ , l'orientation du sujet (i.e. la position de la tête) est enregistrée et l'erreur de positionnement est calculée. L'expérience dure 3 minutes, ce qui correspond à environ 20 mesures.

**L'expérience V.Dyn. (source mobile)** se déroule de la façon suivante :

- L'utilisateur va écouter un son spatialisé sans discontinuer simulant une source sonore sur le cercle. La séquence de positionnement est identique à celle de l'expérience **N.Dyn.**
- L'utilisateur doit continuellement s'orienter vers la source sonore
- 10 fois par seconde l'orientation de la tête de l'utilisateur et la position de la source sonore virtuelle sont enregistrées et comparées.

L'expérience dure 3 minutes. À la fin de l'expérience, l'erreur cumulée est calculée et l'ensemble des données de mesures est conservée.

Pour tous les sujets, l'expérience **V.Stat.** est réalisée avant l'expérience **V.Dyn.**

Dans les expériences **V.Stat.** et **V.Dyn.**, plusieurs fonctions de transferts (HRTF) seront testées en répétant l'expérience plusieurs fois. De façon à contrebalancer les erreurs statistiques possibles liées à l'ordre de présentation des filtres, on prédéfini 8 ordres de passage. Ainsi, suivant les sujets, on présentera en premier l'un ou l'autre des panning, l'une ou l'autre des HRTFs, et aussi les panning avant ou après les HRTF. Cette méthode permettra d'éviter les artefacts liés à l'ordre de présentation des filtres (possibles en raison de l'apprentissage ou de la fatigue).

### 7.4.1 Résultats

Les tests ont été réalisés par 16 sujets normo-entendants (12 hommes, 4 femmes) âgés de 20 à 55 ans recrutés au sein du laboratoire.

**Tests statiques V.Stat. :** Dans ce test, nous comparons la localisation azimutale de la source sonore pour le filtre HRTF par rapport au panning. La figure 7.7 donne les résultats, sujet par sujet, pour les tests avec les filtres IRCAM-Listen1040 et FABIAN. Dans les deux cas, l'erreur d'orientation mesurée par la moyenne quadratique de l'erreur (RMSE) ou la médiane est inférieure pour les HRTFs par rapport au panning. La comparaison sujet par sujet montre également que la plupart des participants obtiennent de meilleurs résultats avec les HRTFs. Comparés ensemble, les HRTFs FABIAN et celles de IRCAM-Listen1040 interpolées donnent des résultats de précision très similaires. Une comparaison des résultats est visible sur le tableau 7.1. Pour autant la valeur statistique de ces résultats est limitée (Test de Wilcoxon : Listen Vs Panning  $p = 0.06$ , Fabian vs Panning  $p \gg 0.05$ ).

**Tests dynamiques :** Les tests dynamiques montrent également de meilleurs résultats de localisation pour les HRTFs par rapport au panning avec une meilleure précision globale (RMSE et médiane). Sujet par sujet (fig 7.8), on constate qu'une large majorité de sujets obtient de meilleures performances avec les HRTFs (aussi bien Ircam que Fabian). Une

Filtre	Listen1040	Listen1040 Panning	Fabian	Fabian Panning
Erreurs médiane/RMSE <sup>1</sup>	<b>3.9/6.5</b>	5.4/11.4	<b>4.0/7.0</b>	4.9/7.0
Confusions avant-arrière <sup>2</sup>	0(0)	3(2)	1(1)	0(0)
#Sujets localisant mieux <sup>3</sup>	<b>11</b>	5	<b>10</b>	6

<sup>1</sup> Erreur médiane et RMSE en degrés

<sup>2</sup> Nombre de confusions avant-arrière sous la forme *nombre de confusion totale rencontrées (nombre de sujets concernés)*

<sup>3</sup> Nombre de sujets ayant une erreur médiane plus faible avec le filtre

TABLE 7.1 – Comparaison des résultats obtenus dans le cas des sources statiques, avec des filtres HRTFs et des filtres de panning.

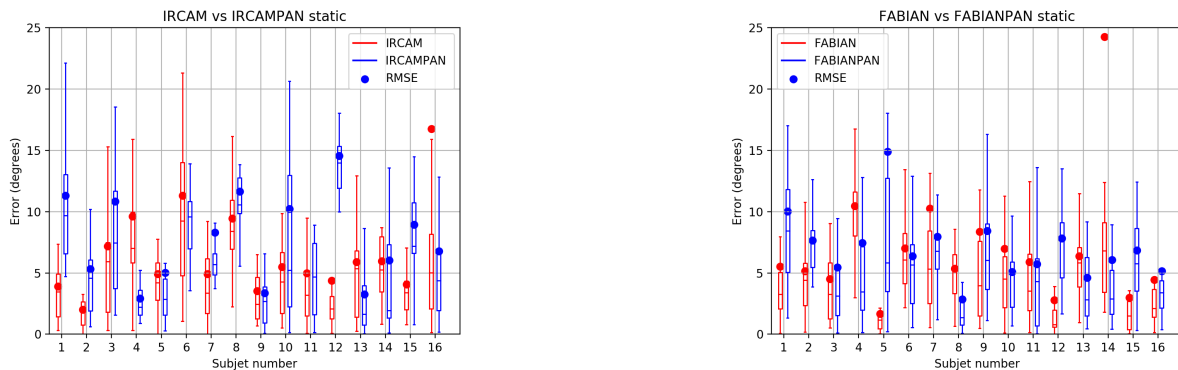


FIGURE 7.7 – Erreurs de pointage vers des sources statiques : HRTF vs Panning pour chaque sujet (IRCAM-Listen1040 à gauche, Fabian à droite).

comparaison des résultats est visible sur le tableau 7.2. Les résultats Listen1040 vs Panning semblent statistiquement significatifs ( $p \ll 0.05$ ), mais ce n'est pas le cas de Fabian vs panning ( $p > 0.05$ ).

Filter	Listen1040	Listen1040 Panning	Fabian	Fabian Panning
Erreur médiane/RMSE <sup>1</sup>	<b>6.9 / 11</b>	9.6/14.6	<b>6.9/10.6</b>	7.6/12.4
#Sujet localisant mieux <sup>2</sup>	<b>13</b>	3	<b>12</b>	4

<sup>1</sup> Erreur médiane et RMSE en degrés

<sup>2</sup> Nombre de sujets ayant une erreur médiane plus faible avec le filtre

TABLE 7.2 – Comparaison HRTFs vs Panning en conditions dynamiques

## 7.4.2 Discussions

Les résultats des tests de localisation de sons naturels statiques (erreur médiane  $3.6^\circ$ ) sont en phase avec la littérature si on se compare avec les études sur la localisation sonore dans la zone frontale (Blauert [31]) bien que la méthodologie soit ici différente. Pour les sons spatialisés, parmi les rares études utilisant le *nose pointing* avec suivi de tête, Brungart et Wenzel relèvent une précision d'environ  $17^\circ$  [39] quand Brungart et Kordin [23] obtiennent environ  $8^\circ$ . Dans nos expériences, nous obtenons une précision sensiblement meilleure avec une erreur médiane de  $4^\circ$ . Cette précision pourrait s'expliquer par la faible latence du système, les performance du tracking ou la qualité des

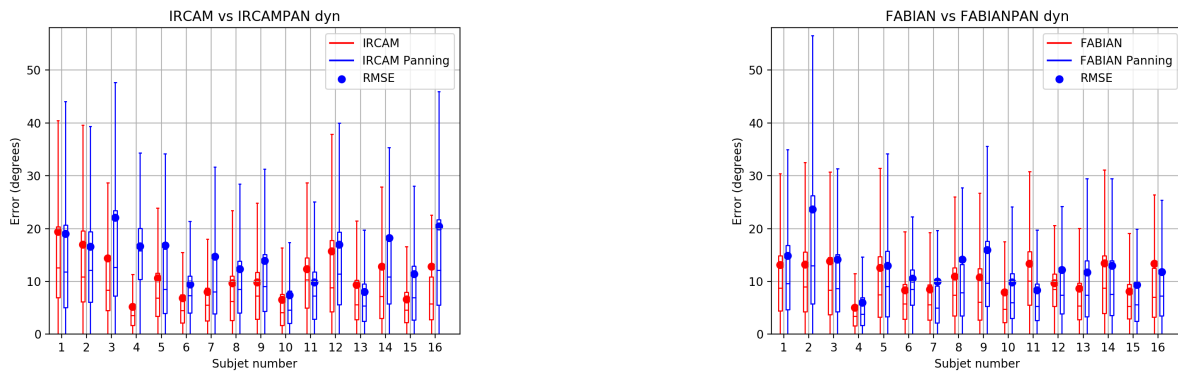


FIGURE 7.8 – Erreurs de pointage vers des sources dynamiques : HRTF vs Panning pour chaque sujet (IRCAM-Listen1040 à gauche, Fabian à droite).

interpolations.

Dans le cas des sons naturels, l'accroissement de l'erreur de positionnement pour suivre un mouvement rapide ( $> 10^\circ/s$ ) par rapport au cas statique n'est que de deux degrés. Cet écart est le même avec les HRTFs. Par contre le panning introduit des erreurs sensiblement plus importantes que les HRTF (et les sons naturels). Par ailleurs, le panning semble être particulièrement inefficace dans le cas des mouvements à dynamique rapide.

Si les performances obtenues avec les HRTF de Listen1040 et de Fabian sont extrêmement similaires, on peut observer une différence plus sensible entre les deux jeux de filtres de panning, les performances étant meilleures avec le panning Fabian. Une explication possible de ce résultat pourrait être la surestimation des ILD pour Fabian avec notre technique d'extraction des ILD.

## 7.5 Conclusion des expériences

D'une manière générale, pour les cas statiques comme pour les cas dynamiques, les performances de localisation des sons naturels sont légèrement meilleures que celles des sons spatialisés par HRTF mais la marge d'erreur des mesures limite la valeur de cette comparaison (d'autant qu'en raison des contraintes sanitaires le nombre de sujets participant aux expériences reste limité). Les filtres de panning permettent quant à eux une précision de localisation sensiblement inférieure. Ainsi, les résultats, à la fois sur les tests statiques et les tests dynamiques tendent à montrer que les indices spectraux jouent un rôle important dans les tâches d'orientation vers une source sonore sur le plan azimutal. On peut noter que les HRTF interpolées de Listen1040 permettent des performances très comparables à celle obtenues avec les HRTF à haute résolution spatiale du projet Fabian.

Les performances supérieures des HRTF par rapport aux filtres de panning justifient donc leur utilisation pour une application de guidage. En revanche, le choix des HRTFs et la question de l'individualisation mériteraient d'être approfondi par des études supplémentaires.



# Conclusion de la deuxième partie

Dans cette partie nous nous sommes successivement intéressés à la perception des sources sonores et à la façon de générer les stimuli spatialisés efficacement. Grâce à un protocole expérimental adapté, nous avons également pu montrer, que sur le plan azimutal, la direction des sources sonores virtuelles était perçue avec une acuité comparable à celle obtenue pour des sources sonores naturelles, que les sources soient statiques ou en mouvement. Ce dernier point est particulièrement important pour l'application que nous envisageons, dans laquelle l'utilisateur sera en mouvement rapide. Nous avons pu en outre montrer que l'utilisation de HRTFs améliorerait sensiblement la perception de la direction par rapport au panning.

Ces constats ouvrent la voie au guidage audio par des sources virtuelles spatialisées et dynamiquement adaptées en temps-réel grâce à un système de suivi de la tête et de la position. Pour une mise en œuvre efficace, cette approche temps-réel a nécessité une compréhension approfondie des sources de latence impliquées dans la boucle perception-action-stimuli, qu'elles soient liées au head-tracking, au sous-système audio des OS modernes et aux algorithmes de traitement du signal mis en œuvre. Ainsi, nous avons implémenté des algorithmes efficaces pour la spatialisation sonore temps-réel sur des plateformes embarquées disposant de peu de ressources. En particulier nous sommes capable d'effectuer des rendus binauraux sur plateforme ARM pour des sources multiples avec des latences sensiblement inférieures aux seuils habituellement audibles.



## **Troisième partie**

# **Vers un dispositif de guidage pour la pratique sportive**





# Introduction

Aujourd'hui de nombreuses personnes non-voyantes pratiquent la course à pied ou le roller sur piste avec un guide voyant, mais cette activité se heurte à différents obstacles pratiques et logistiques. En effet il est souvent difficile pour les clubs sportifs de mettre à disposition un guide valide pour chaque athlète déficient visuel, et par ailleurs, dans le contexte du sport de compétition, il est difficile de trouver des guides avec des capacités sportives suffisantes.

Habituellement, les athlètes déficients visuels et les guides accompagnateurs sont reliés ensemble par un lien physique comme un bâton de relais ou un élastique, mais certains sportifs non voyants pratiquent également le roller, la course à pied ou la natation avec un guidage sonore. L'athlète suit alors son guide au bruit de ses pas ou de ses rollers. Nous nous sommes inspirés de cette pratique pour développer un système de guidage où l'athlète doit suivre continuellement une source sonore virtuelle spatialisée par synthèse binaurale.

Nous avons montré précédemment que les sujets étaient capables de s'orienter vers une source sonore spatialisée de façon comparable à une source sonore naturelle, sans introduire de délai dans la perception du mouvement (hors latence liée au traitement numérique) et avec une précision peu dégradée. Le son binaural permet une présentation de l'information spatiale à transmettre de façon immédiatement perceptible, sans introduire des délais liés à un encodage particulier (comme la voix par exemple, où on doit attendre que l'information soit énoncée). Nous montrerons dans cette partie qu'en s'appuyant sur les pratiques des personnes déficientes visuelles et en utilisant des stimuli binauraux, il est possible de guider des personnes de façon efficace et naturelle y compris dans le contexte de la pratique sportive et de la recherche de performance.



## Chapitre 8

# Navigation et locomotion des personnes DV

### 8.1 Cognition spatiale

La perception de l'espace par les individus, sa représentation mentale, la localisation des éléments qui la composent et les mécanismes mentaux mis en œuvre dans la navigation ont principalement été étudiés dans le cadre de la psychologie cognitive à partir du milieu du 20<sup>ème</sup> siècle [62]. En 1948, le psychologue Edward Tolman [215] introduit le concept de *carte cognitive* pour décrire la représentation mentale qu'un individu se fait de l'organisation de l'espace dans lequel il se trouve. Tolman fait alors l'hypothèse que les individus sont capables d'établir des relations entre les différents points de l'espace sans avoir eu une expérience directe. Pour lui, les individus s'appuient sur leur mémoire pour organiser des relations entre les lieux. Ces relations prennent alors la forme de liens, de distances ou de directions à la manière d'un graphe. Plus tard, la biologie a tenté d'identifier les structures cérébrales supports de ces capacités. Ainsi, les *cellules de lieux* de l'hippocampe [175] seraient le support de ces cartes cognitives alors que les *cellules de grilles* du cortex entorhinal joueraient un rôle dans la localisation dans l'espace [201]. La découverte de ces deux types de cellules a valu le prix Nobel de médecine à John O'Keefe, Edvard et May-Britt Moser en 2014.

#### 8.1.1 Sources d'informations lors d'un déplacement

Lors d'un déplacement, un ensemble d'informations est perçu et utilisé pour les tâches de navigation. Ces informations peuvent être réparties en deux catégories :

- Les *informations allothétiques* correspondent à l'ensemble des informations issues de l'environnement extérieur. Ces informations sont typiquement de nature visuelles, olfactives, sonores ou tactiles.
- Les *informations idiothétiques* proviennent de l'individu lui-même via les systèmes vestibulaires, proprioceptifs ou moteurs. Elles renseignent donc l'individu sur sa position, son orientation et éventuellement sa vitesse grâce à la perception du corps et de ses mouvements. L'intégration de ces informations peut permettre d'estimer sa position en l'absence d'informations allothétiques. Ces informations sont donc particulièrement importantes pour les personnes déficientes visuelles.

### 8.1.2 Représentation de l'espace et référentiels

On utiliserait simultanément deux types de référentiels pour se représenter l'espace [44]. Un référentiel égocentré, qui permet à un individu de connaître la position des éléments qui l'entourent et un référentiel allocentré (centré sur l'environnement) qui lui permet de se situer par rapport à des objets ou des lieux de référence.

Selon divers auteurs [62, 175], les tâches de navigation s'appuient sur différentes représentations ou connaissances des objets géographiques. Tout d'abord l'individu a besoin de *points de repères* (allocentrés) qui sont des éléments de l'environnement aux caractéristiques reconnaissables (bâtiments, lieu associé à un stimulus particulier...). Pour relier ces points, la connaissance des *routes* (souvent présentée de façon égocentrée) permet alors de décrire les liens en termes d'orientations ou de distances entre les différents points de repères. Une représentation d'ensemble structurée d'une zone constituera alors une *carte* (allocentrée). Cette carte mentale permettrait alors d'établir des relations entre des lieux et permettra d'appréhender les lieux inexplorés de la zone.

## 8.2 Navigation globale et fine, locomotion

Après une phase préliminaire de planification qui permet à la personne de déterminer grossièrement le chemin à emprunter, la navigation active met en œuvre deux types de capacités :

- *La navigation globale* consiste à suivre le chemin planifié en reliant les différentes étapes du parcours.
- *La navigation fine* permet de demeurer sur le bon chemin tout en gérant les différents aléas du parcours. Chez la personne valide cette tâche est principalement assurée par la vue. Les personnes déficientes visuelles (DV), quant à elles, vont devoir mettre en œuvre des techniques particulières, dites de *locomotion*.

Pour une navigation efficace, les personnes DV vont devoir développer et mettre en œuvre des stratégies dans leurs déplacements. Il s'agit typiquement d'être certain de suivre un chemin identifié comme sûr par exemple en restant sur un trottoir, rester sur les graviers d'un chemin piéton, etc.

Les associations de déficients visuels ont mis en place des formations pour enseigner ces techniques de locomotion. Elles consistent entre autre à apprendre à utiliser les outils d'aide aux déplacements (comme la canne blanche), et à développer des capacités sensorielles et cognitives par des exercices en situation.

Ne pouvant s'appuyer sur sa vue, la personne déficiente visuelle doit utiliser toutes les informations qui lui sont fournies par ses autres sens. La *proprioception* lui permet de conserver une trajectoire et d'estimer les distances parcourues. Le *touché* permet d'analyser les textures du sol via la canne ou les semelles des chaussures. Les personnes déficientes visuelles interrogées expliquent également utiliser leur perception de la chaleur pour identifier des lieux (typiquement la proximité des fenêtres en intérieur). L'*odorat* peut dans une moindre mesure être utilisé pour identifier des lieux. Enfin, l'*audition* joue un rôle primordial, à la fois pour recueillir des indices sur les objets sonores (véhicules, personnes) ou de manière active à travers l'écholocalisation. Le sens auditif est généralement décrit comme prépondérant par les personnes déficientes visuelles que nous avons interrogées, certaines d'entre elles utilisant volontiers la formule "mes oreilles sont mes yeux". Tout dispositif d'aide au déplacement devra éviter de perturber la perception de ces informations au risque d'être contre productif. Ainsi, il est de première importance qu'un dispositif de guidage utilisant des informations auditives préserve la perception sonore de l'utilisateur.

## 8.3 Pratique sportive chez les déficients visuels

Une faible part des personnes DV pratiquent une activité physique ou sportive en club. En France, en 2017, sur 28 684 licenciés à la Fédération Française Handisport, 1841 étaient des déficients visuels [75]. Pourtant le nombre de personnes atteintes de déficience visuelle sévère approche le million de personnes. Le taux de pratiquant est donc de l'ordre de 0.2%.

Plusieurs facteurs peuvent expliquer le mode de vie sédentaire des personnes déficientes visuelles. On peut noter le manque de ressources et d'activités accessibles, la peur des blessures, le besoin d'un soutien particulier, l'âge ou le manque d'informations sur les activités disponibles.

Pour autant, l'intérêt de la pratique sportive et de l'activité physique en général d'un point de vue physique et psychologique n'est plus à démontrer. Pour les personnes DV plus spécifiquement, l'activité sportive peut être un moyen de promouvoir l'inclusion sociale. Par ailleurs le sport permet de stimuler la sensibilité proprioceptive et en conséquence de développer la motricité et la perception des membres dans l'espace lors des mouvements, ce qui est susceptible d'avoir un impact direct sur l'autonomie.

En compétition, les sportifs déficients visuels sont classés en trois catégories :

- La classe 1, B1 ou 11 (selon le sport) pour cécité majeure ou complète ;
- La classe 2, B2 ou 12 pour une cécité sévère ;
- La classe 3, B3 ou 13 pour une cécité légère.

Selon ces différentes catégories, des épreuves spécifiques sont mise en place. C'est le cas de la course à pied où la présence d'un guide est obligatoire pour la catégorie T11, possible pour la catégorie T12 et interdite pour la catégorie T13.

### 8.3.1 Les sports pratiqués par les personnes déficientes visuelles

Les personnes déficientes visuelles peuvent pratiquer à la fois les activités sportives communes (avec guide ou aménagements) mais aussi des sports spécifiques. Voici une liste non exhaustive des activités pratiquées ainsi que les aménagements possibles :

#### L'athlétisme

L'athlétisme regroupe plusieurs types d'activités comme la course à pied, les lancers et les sauts qui peuvent toutes être adaptées pour les personnes vivant avec un handicap visuel.

En ce qui concerne la course, qui est certainement l'activité la plus pratiquée, les personnes ayant peu ou pas de vision doivent être accompagnées d'un guide valide. Différentes techniques peuvent être utilisées mais la plus classique consiste à utiliser une cordelette d'environ 25 cm tenue par le guide et le coureur aveugle. Si l'athlète a une vision résiduelle, le guide peut porter un dossard très voyant et se placer de façon visible devant le coureur pour lui permettre de se diriger. Pour le saut en longueur, le guide est placé au niveau du sautoir et frappe dans ses mains pour indiquer où se situe l'espace de saut appelé *chaux*. Il s'agit d'un carré mesurant 1 mètre sur 1 mètre formé de chaux. La longueur du saut est alors mesurée en utilisant l'empreinte du pied. Le sportif doit alors connaître précisément le nombre de foulées à réaliser pour réussir son saut.

## **Aviron**

L'aviron est une activité facilement praticable pour une personne déficiente visuelle puisque ce sport se pratique à deux ou en équipe.

## **Équitation**

Il s'agit d'un sport assez populaire chez les personnes DV. Les cavaliers sont généralement accompagnés de personnes qui serviront de guides sonores, à la fois pour s'orienter et pour signaler les obstacles.

## **Escalade**

Le grimpeur DV utilise le toucher pour trouver les prises. Ce sport se pratique généralement accompagné d'un guide qui fournira des indications pour faciliter l'ascension.

## **Goalball**

Le goalball a été le premier sport se pratiquant en équipe créé spécifiquement pour les personnes DV. Ce sport a été inventé après la Seconde Guerre mondiale comme moyen de réadaptation pour les personnes devenues aveugles pendant la guerre. Plus tard les règles du jeu ont été formalisées et le goalball a fait son entrée aux Jeux paralympiques en 1976.

Au Goalball, deux équipes composées de trois joueurs s'affrontent et doivent marquer des buts avec un ballon contenant des grelots. Les joueurs vont faire rouler la balle sur le sol suffisamment vite pour que l'équipe adverse ne puisse pas la bloquer avant qu'elle n'entre dans leur but. Les joueurs en défense tentent alors d'arrêter le ballon en se couchant. Si le ballon entre dans le but, un point est marqué. Si l'un des joueurs en défense arrête le ballon, il peut le renvoyer ou le passer à un coéquipier.

## **Natation**

En général, une personne valide est positionnée à chaque extrémité de la piscine. Cette personne est dénommée *tapeur*. Son rôle consiste à signaler au nageur la proximité du bord de la piscine en le tapant à l'aide d'une perche recouverte d'une mousse. Le nageur pourra alors réaliser son virage. Pour permettre au nageur d'évoluer plus facilement en ligne droite, les couloirs de nage peuvent être rendus plus étroits. En eau vive le sportif peut nager avec un guide avec lequel il est relié par un lien d'environ 60cm.

## **Ski alpin et ski de fond**

Les skieurs DV sont guidés sur le chemin à suivre par des guides voyants qui utilisent des indications vocales par radio pour indiquer le chemin à suivre. En compétition, aucun contact physique entre le guide et le compétiteur n'est permis, par contre le guide peut être positionné n'importe où par rapport au skieur. Le guide et le sportif créent un code verbal généralement basé sur un langage très court pour transmettre rapidement les informations et réagir au plus vite ; par exemple « top » ou « oui » (virage), « plus » (se rapprocher) ...

### **Cécifoot**

Ce sport, qui s'apparente au football, se pratique par équipes de cinq joueurs sur une surface de la taille d'un terrain de handball. Les différents joueurs se localisent en émettant des sons (le possesseur du ballon ayant l'obligation de se signaler par la voix). Le ballon lui-même est sonorisé par des grelots. Lors des actions de tir, un assistant sonorise l'emplacement du but en frappant sur le cadre métallique.

### **Tir sportif**

Cette activité se pratique souvent à l'aide de pistolet laser. Un système de sonification permet au tireur de viser et indique également si le tir a été réussi ou non.

### **Judo**

Il s'agit d'un des rares sports pouvant se pratiquer sans adaptation ni assistance. Les pratiquants peuvent ainsi développer leur sens de l'équilibre, de l'orientation et les techniques de chutes : des capacités utiles dans la vie quotidienne.

## **8.3.2 Compétitions :**

Des compétitions au niveau international pour les personnes ayant une déficience visuelle ont lieu dans plusieurs disciplines. Le Cécifoot, le GoalBall, l'Athlétisme, l'Aviron, le Cyclisme, le Judo, la Natation et le Triathlon sont les sports qui seront présents aux jeux paralympiques d'été de 2024 pour les personnes ayant une déficience visuelle.

## **8.3.3 Guidage**

Suivant les disciplines et le niveau de handicap, les sportifs sont guidés en compétition et/ou à l'entraînement. Par exemple, en ce qui concerne l'athlétisme, tous les sportifs non-voyants (catégorie 11 en athlétisme) sont guidés à la fois à l'entraînement et en compétition. En revanche certains athlètes malvoyants peuvent effectuer leurs entraînements totalement ou partiellement seuls.

La relation qui lie le binôme guide-sportif est très importante pour permettre une bonne condition de pratique, en particulier pour la mise en confiance et la diminution du stress. Ils doivent également communiquer pour apprendre à se connaître et partager leurs ressentis. Cette relation est d'autant plus importante que le guide joue un rôle également dans un contexte extra-sportif. Il peut par exemple accompagner le sportif dans les lieux de compétition, et veiller à ce que les conditions réglementaires soient remplies (e.g. la mise en place du dossard). La relation guide-athlète étant particulière, le remplacement d'un guide est délicat et la Fédération Française Handisport (FFH) recommande de s'entraîner avec deux guides pour pallier à l'indisponibilité de l'un d'eux.

La question des performances du guide est également un sujet délicat pour l'organisation des compétitions et un frein potentiel à la pratique et aux performances. En compétition, pour l'athlétisme, la FFH estime que le niveau de performance du guide doit être sensiblement plus élevé que celui de l'athlète : réaliser 1/2 seconde de moins sur 100 m ; 1 seconde sur 200 m ; 3 à 4 secondes sur 400 m, etc. De fait, les athlètes non voyants ne peuvent donc pas dépasser les performances des athlètes valides.



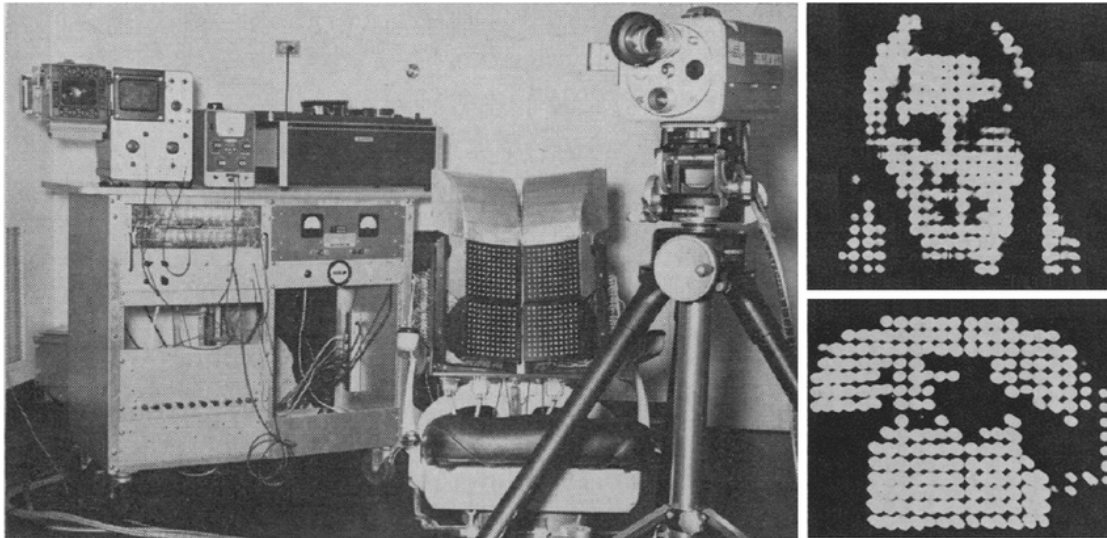


FIGURE 8.1 – Le dispositif de Bach-y-rita (extrait de [16])

## 8.4 Substitution et augmentation sensorielle

Le concept de substitution sensorielle (SS) a été introduit il y a 50 ans par Bach-y-Rita [16]. L'idée générale est de transmettre les données à un canal sensoriel différent de celui habituellement utilisé. Par exemple, en remplaçant l'information visuelle par un codage sonore ou tactile.

Bach-y-Rita, développa des systèmes pour permettre aux personnes aveugles de percevoir des informations visuelles par stimulation vibro-tactile. Son premier appareil captait les images d'une caméra et les représentait sur la peau du dos par le biais de vibrations utilisant une matrice de  $20 \times 20$  actionneurs haptiques, comme le montre la figure 8.1. Après un entraînement suffisant, les signaux utilisés sont supposés pouvoir être traités inconsciemment par l'utilisateur et ainsi permettre une perception intuitive, comme une nouvelle expérience sensorielle. En parallèle, des avancées en sciences cognitive sur la plasticité cérébrale ont conduit de nombreux chercheurs à s'intéresser à ces techniques. Cette approche a fait l'objet de nombreuses recherches, mais présente cependant d'importantes limitations [92, 155] et peu de dispositifs utilisant la SS ont trouvé de réels intérêts pratiques. Parmi les problèmes qu'ils induisent, on peut mentionner l'interférence avec les autres activités et le canal sensoriel utilisé, mais surtout ils nécessitent un apprentissage très important et induisent une charge cognitive supplémentaire.

Pour pallier ces difficultés, J.M. Loomis [150] propose une méthodologie pour concevoir des SSD plus efficaces via une approche en deux étapes. Tout d'abord, il propose d'identifier les informations réellement nécessaire pour effectuer une tâche. Par exemple, pour éviter un obstacle il n'est pas nécessaire de connaître tout l'environnement, puisqu'il faut simplement identifier les points à éviter. Dans la deuxième étape Loomis suggère de coupler cette information avec les modalités de substitution (typiquement toucher ou audition) en prenant en compte la *bande passante* sensorielle de la modalité de substitution. On peut en effet considérer qu'il est possible de transférer bien plus d'informations de nature spatiale avec la vision qu'avec l'audition ou le toucher. Ainsi, tenter d'utiliser une représentation de l'espace isomorphe à une image à partir d'une caméra vidéo vers les sens du toucher ou de l'audition impliquera une perte d'information ou de la latence et sans doute une surcharge sensorielle.

D'une manière générale on peut qualifier d'*augmentation sensorielle* [66, 118] les méthodes qui ne cherchent pas

à transmettre une information de substitution de manière isomorphe mais qui effectuent un traitement préalable de traduction ou de sélection de l'information à transmettre. La méthode d'orientation par audio binaural que nous décrirons plus bas peut s'inscrire dans cette approche.



## Chapitre 9

# Les dispositifs d'aide à la mobilité des personnes déficientes visuelles

### 9.1 Dispositifs classiques

La *canne blanche* est le dispositif d'aide à la locomotion le plus connu et le plus utilisé à travers le monde. Par un mouvement de balayage latéral, il permet à son utilisateur de rechercher des obstacles et d'identifier les textures du sol. Largement reconnu comme symbole de la déficience visuelle, elle constitue également un moyen pour la personne non-voyante de se signaler.

Le *chien guide d'aveugle* est un chien d'assistance formé pour guider les personnes aveugles et malvoyantes en évitant les obstacles qui se présentent sur son parcours. C'est l'homme qui fixe la direction, alors que le chien remplit principalement une tâche de navigation fine guidant son maître sur un trajet sûr. À noter que d'autres animaux peuvent remplir cette fonction, c'est le cas de poneys de petites tailles, également spécialement entraînés pour cette tâche.

### 9.2 Dispositif électroniques

#### 9.2.1 Historique et état de l'art

##### 9.2.1.1 Les premiers développements

Les dispositifs d'aide électroniques au déplacement (souvent désignés sous l'acronyme ETA pour *Electronic Travel Aid*) ont déjà une longue histoire. À la fin du 19e siècle, le chercheur polonais Kazimierz Noiszewski a développé un dispositif [46] appelé *Elektroptalm* utilisant une cellule photoélectrique de sélénium qui permettait aux personnes aveugles de distinguer la lumière du noir en faisant varier l'intensité d'un son. En 1912, d'Albe développa l'*exploring Optophone* [46], utilisant le même principe. Malheureusement, ce type d'informations reste d'un intérêt limité pour une personne déficiente visuelle. Des travaux se sont poursuivis au début du 20ème siècle pour fournir une information plus pertinente et utile pour les personnes. On a alors cherché à détecter des obstacles avec les technologies disponibles. Parmi ces tentatives, on peut citer l'anecdotique "Radiation Guider" [28], qui cherchait à utiliser les réflexions sur les obstacles de particules issues d'une source radioactive. Un compteur Geiger était alors utilisé pour détecter

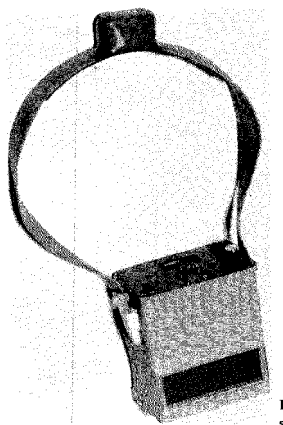


FIGURE 9.1 – Le *Pathsounder* de Russel (extrait de [85])

les particules réfléchies et signaler les obstacles. Plus raisonnablement, en 1963, Benham et Benjamin [25] proposent un système utilisant la réflexion d'un faisceau lumineux. Peu après, plusieurs systèmes utilisant des ultra-sons ont été développés. En 1965 Leslie Key développe l'Ultra Sonic Torch [127], un système performant permettant de balayer l'espace comme l'on ferait avec une lampe torche et qui fournit un retour sous la forme d'un son monophonique. En 1968, Russel propose le PathSounder [85, 197]) un autre système ultra-son qui se porte autour du cou (c.f. fig. 9.1) et doté de 5 capteurs ultra-son. Il propose également un retour sous forme sonore en mode tout-ou-rien (go-no-go). La portée de ces premiers systèmes restait limitée en raison des caractéristiques des capteurs (environ 2m). En 1972, le système *Mowat* [187] permettait, quant à lui, de détecter des objets jusqu'à 4m de distance. C'est un appareil de la taille d'une lampe de poche que l'on utilise par balayage et qui vibre avec une période qui varie en fonction de la distance à la cible.

Ainsi, les premiers systèmes étaient essentiellement des dispositifs d'évitement d'obstacle. À la fin des années 60, la nouvelle approche méthodologique de Paul Bach-y-Rita oriente de nombreuses recherche sur la voie de la substitution sensorielle et conduisent aux développement de plusieurs dispositifs. Parmi les plus récents, les plus notables sont :

- Le système *The voice* (1992) [162] qui propose de convertir des images en motifs sonores. Les représentations sonores sont multiplexées dans le temps pour effectuer une conversion image/son. Le système propose de représenter des images jusqu'à une résolution de 64 × 64 pixels avec 16 tons de gris par pixel.
- Le système *Videotac* (1996) [43] qui cherche à reproduire des images en temps réel via une matrice de 768 actionneurs. En 1998, l'équipe de Bach-y-Rita propose le *Tongue display* [17] fonctionnant sur le même principe mais avec la matrice d'actionneurs sur la langue de l'utilisateur.

### 9.2.1.2 L'approche actuelle

Aujourd'hui certains auteurs [9, 67] classent les ETA en trois catégories suivant leurs fonctionnalités : les systèmes de détection d'obstacles, les systèmes renseignant sur l'environnement et enfin les systèmes de guidage cherchant à orienter les utilisateurs. Dans la première catégorie, on cherche à proposer des solutions à la question de la navigation fine de façon analogue à la canne blanche en alertant les utilisateurs de la présence d'un obstacle proche. Ces systèmes ont été les premiers à être développés et sont également les plus utilisés au quotidien par les personnes déficientes visuelles. De nombreux systèmes ont connu des succès industriels importants comme la canne blanche électronique développée à l'Université Paris-Saclay par Farcy & al [74].

Les systèmes de la deuxième catégorie tentent d'aller plus loin en fournissant des informations plus précises et

complètes sur l'environnement (localisation des objets, couleurs, distances, etc.). Les dispositifs les plus rudimentaires sont constitués de capteurs ultra-son et fournissent les informations correspondant à la distance des objets encodée sous forme tactile ou auditive. Le *Binaural Sonar Electronic Travel Aid* [132] utilise un sonar binaural (bi-capteur ultra-son) et fournit de façon tactile des informations sur la distance des objets, les mouvements dans le champ et même les textures. Les dispositifs des plus récents sont souvent basés sur la technologie de vision par ordinateur qui fournissent des stimuli vibrotactiles [113] ou sonores [41].

La troisième classe d'ETA est constituée des systèmes d'aide à la navigation. Ces systèmes sont destinées à orienter et guider les personnes le long de toutes les étapes d'un parcours en les localisant en temps réel. Ces ETA peuvent également inclure des technologies d'analyse de l'environnement ou d'évitement d'obstacles pour une navigation sûre. De la même façon que les systèmes de navigation automobile, ils utilisent souvent les systèmes de localisation par satellites ainsi que des centrales inertielles pour informer l'utilisateur sur l'itinéraire à suivre, à l'aide de commentaires verbaux, de points de repère, ou d'instructions.

Les premiers systèmes de ce type sont apparus dans les années 90. Parmi les pionniers, *MoBIC* [182] était constitué d'un outil de planification d'itinéraire et d'un système de guidage en temps-réel. L'IHM était constituée d'un petit clavier portable et d'une voix artificielle. *Geotact* développé à l'université Paris-Sud fournissait lui des informations vocales sur les distances et les coordonnées, l'utilisateur était alors en mesure de sélectionner le chemin à suivre pour atteindre son objectif. *Makino* [156] fonctionnait lui sur téléphone portable dès 1996.

Aujourd'hui plusieurs systèmes ont pu arriver au stade industriel, on peut citer : le *GPS Braille Note*, le GPS pour non voyant *Trekkeur* (constitué d'un clavier, d'un haut parleur placé sur le col et d'une antenne placée sur l'épaule), le *Trekkeur Brezze* qui permet d'enregistrer un parcours pour le 'rejouer' ou le module *Kapten* qui utilise une IHM par reconnaissance vocale. Très récemment, *Orcam* propose un dispositif de type lunette intelligente qui analyse l'environnement, reconnaît les visages ou lit des textes.

La tendance actuelle est d'utiliser les téléphones portables qui fournissent à la fois le capteur GPS, l'unité de traitement et l'IHM (via la couche d'accessibilité incluse dans les téléphones modernes). On peut citer les applications *Navigon* qui est comparable au guidage automobile avec une fonction d'annonce vocale de la position (adresse), *Ariadne* qui donne des informations de cap, ou les application *Navi-urbain* et *Navi-rando* [190]. Enfin le système *GoSense Wizigo* [5] propose un GPS avec des indications sonores binaurales de points d'intérêts.

### 9.2.1.3 Tendance

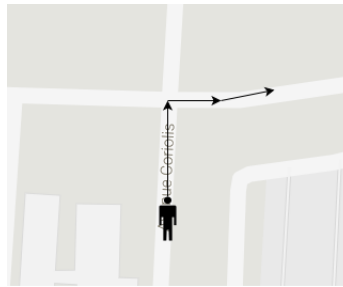
L'évolution des techniques a permis l'émergence de dispositifs de plus en plus complets. Pour autant, les systèmes d'évitement d'obstacles restent les plus utilisés, même si les dispositifs de guidage GPS sur téléphones sont aujourd'hui assez populaires. Du côté de la substitution sensorielle très peu de dispositifs sont utilisés dans la pratique par des personnes aveugles (voir [92] pour une étude complète sur les SSD tactiles et leurs utilisations) et la tendance actuelle est plutôt à une forme de d'augmentation sensorielle.

## 9.3 Le son spatialisé comme outil de guidage

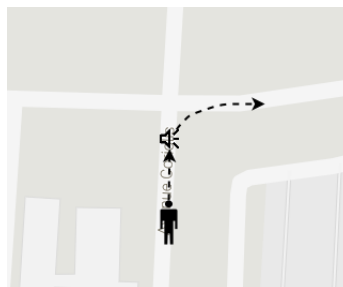
Si la plupart des dispositifs de navigation utilisent la synthèse vocale, d'autres ont cherché à exploiter les capacités des personnes à localiser des sons pour les orienter.

L'idée du guidage par audio spatialisé n'est pas nouvelle et plusieurs méthodes ont été proposées, généralement pour des applications de navigation urbaine pour les piétons.

- **Guidage par points d'intérêts spatialisés** : Cette méthode a été proposée par Loomis qui l'a mise en oeuvre dans le système *Personal Guidance System* (PGS) en 1998 [149]. Il s'agit d'un système de navigation et d'orientation basé sur le GPS mais l'information est cette fois restituée par sonification binaurale. Dans ce système, lorsque la personne aveugle se déplace dans l'environnement, elle entend les noms des bâtiments, des intersections de rue, etc., prononcés par un synthétiseur de parole, provenant des endroits appropriés comme si des haut-parleurs étaient situés en ces lieux. Le dispositif incluait donc une forme de head-tracker (uniquement un magnétomètre), un générateur de son 3D utilisant un DSP (l'*Alphatron* de Crystal River Engineering) un ordinateur et un casque audio. L'utilisateur effectue alors une navigation point-par-point et peut conserver une image mentale de la localisation des différents points d'intérêt qui l'entourent. Loomis obtient des résultats encourageants avec son dispositif, même s'il note des difficultés, comme la perception de la distance qui reste peu fiable, l'occlusion des oreilles de l'utilisateur ou les difficultés dans la réponse temps-réel du traitement binaural avec les mouvements de la tête. Cette méthode a notamment été utilisée dans les projets de recherche GpsTunes en 2006 [209], Ontrack en 2008 [114], Swan en 2007 [233] ou encore Navig en 2012 [126].



- **Guidage adaptatif binaural temps réel (ou guidage "fée clochette")** : il consiste à guider les personnes avec une source sonore virtuelle qui précède l'utilisateur à une distance constante. Mais, la méthode présente plusieurs difficultés. En effet, si on veut obtenir un guidage précis, la source sonore doit précéder de peu l'utilisateur (par exemple un mètre). Dans ce cas, il est nécessaire de localiser une personne de façon très précise, car un déplacement même très léger doit pouvoir être perceptible et restitué dans le système audio. Par ailleurs le moteur audio doit également avoir une faible latence. Cette méthode a notamment été utilisée dans le cadre du projet Argus [48], avec un système GPS. Toutefois la précision et la latence introduite par le système GPS limitait la précision et le confort du guidage.



### 9.3.1 Le suivi sonore comme tâche écologique pour la personne déficiente visuelle

Nous avons interrogé des personnes malvoyantes ou aveugles pour savoir dans quelles activités elles utilisaient la localisation de sources sonores. Ces questions ont été posés à l'occasion des tests de nos dispositifs ou lors de d'entretiens téléphoniques postérieurs :

- Au quotidien, vous arrive-t'il de vous diriger vers un point de repère sonore fixe (par exemple un magasin bruyant ou un feu de signalisation sonore) ? Choix possibles : Dès que possible, Souvent, Rarement, Jamais.
- Vous est il difficile de vous diriger vers ce type de son (fixe) ? Choix possibles : Très peu difficile, peu difficile, moyennement difficile, difficile, impossible. Êtes vous capable d'effectuer confortablement d'autres tâche en même temps (discuter, utiliser une canne blanche) ?
- Au quotidien, vous arrive-t'il de suivre des sources sonores en mouvement. Il s'agit par exemple de suivre une personne traînant une valise sur un trottoir ou de suivre les bruits de pas d'une personne ? (mêmes choix que pour les sons fixes).
- Vous est-il difficile de suivre un tel son en mouvement (personne qui tire une valise) ? (même choix que pour les sons fixes). Êtes vous capable d'effectuer confortablement d'autres tâche en même temps ?

Chacune de ces questions a donné lieu à des discussions où les personnes pouvaient décrire les circonstances où elles sont amenées à suivre des sons dans leur vie quotidienne. Les réponses à ce questionnaire sont récapitulées dans le tableau 9.1.

Il est apparu qu'elles utilisaient à la fois la localisation de sources sonores fixes (lieu habituellement bruyant, feux de signalisation sonorisés, etc.) pour s'orienter et la localisation de sources sonores mobiles – typiquement le suivi de personnes – pour se diriger. Ces deux tâches écologiques sont analogues aux techniques de guidage présentées ci-dessus. En terme d'efficacité, les personnes déficientes visuelles plébiscitent le suivi de source(s) sonore(s) dynamique(s). Les cas d'usage évoqués par les personnes interrogées sont assez variés : suivre un guide en intérieur ou en extérieur, suivre une personne traînant une valise, suivre un groupe, suivre un coach sportif, etc. Nous les avons également interrogées sur le niveau de concentration nécessaire pour suivre une source mobile afin de savoir si toute l'attention doit être portée sur la tâche de localisation de la source sonore. D'une manière générale, les personnes ont évalué la concentration nécessaire de faible à moyenne et dans tous les cas, la charge cognitive était suffisamment faible pour permettre des tâches annexes comme utiliser une canne blanche ou poursuivre une conversation.

### 9.3.2 Guidage sonore et pratique sportive

Plusieurs sports adaptés à la pratique des personnes aveugles utilisent un guidage sonore. C'est en particulier le cas du ski (indications vocales) ou de l'équitation. Si pour la course à pied le lien physique est la norme, elle peut être pratiquée avec un guide muni d'un haut parleur qui précède l'athlète. Pour le roller, le lien physique n'est pas adapté et le guide précède souvent le sportif avec un dispositif sonore, du moins en intérieur. Ce type de guidage, pratiqué par exemple par l'association MCV (MixHandi Cap sur la Vie) est analogue au guidage adaptatif présenté plus haut mais avec des sons naturels et a été une source d'inspiration pour les travaux présentés dans cette thèse.



Sujet	Handicap	Sons fixes		Sons mobiles	
		Pratique	Difficulté	Pratique	Difficulté
N.	malvoyant	Souvent	Très peu difficile	Rarement	Ne sait pas
B.	malvoyant	Souvent	Très peu difficile	Rarement	Très peu difficile
Ty.	malvoyant	Souvent	Très peu difficile	Dès que possible	peu à moyennement difficile
Al.	malvoyant	Souvent	Très peu difficile	Dès que possible	peu difficile
S.	aveugle	Souvent	Très peu difficile	Souvent	Très peu difficile
A.	aveugle	Souvent	Très peu difficile	Dès que possible	Très peu difficile
M.	aveugle	Souvent	Très peu difficile	Dès que possible	Très peu difficile
Th.	aveugle	Souvent	Très peu difficile	Rarement	peu à moyennement difficile

Notes :

Choix possibles pour la *Pratique* : Dès que possible, Souvent, Rarement, Jamais.

Choix possibles pour la *Difficulté* : Très peu difficile, peu difficile, moyennement difficile, difficile, impossible.

TABLE 9.1 – Suivi de sons dans la vie quotidienne.

## Chapitre 10

# Un dispositif de guidage par son spatialisé pour la pratique sportive

Notre objectif est de fournir à l'utilisateur une source sonore virtuelle qui lui indique continuellement la direction à suivre en le précédant sur un parcours prédéfini. Pour cela, nous devons localiser le plus précisément possible l'utilisateur et créer un son spatialisé devant lui, dans la direction à suivre, qui doit être restitué en tenant compte de l'orientation de sa tête. En effet, l'idée principale derrière un dispositif de guidage binaural est de permettre à la personne de suivre un son naturel. Pour ce faire, il est essentiel de donner à l'utilisateur un sentiment réaliste et cohérent avec ce qui serait perçu dans une situation réelle.

Il s'agit donc de transposer la tâche écologique de suivi de source sonore dans un guidage par son spatialisé (c.f. fig 10.1).

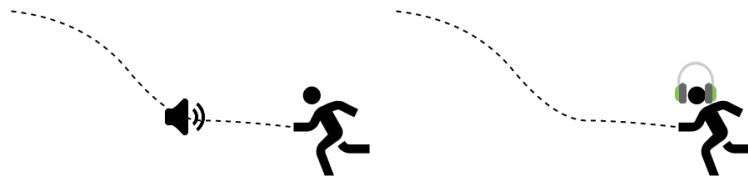


FIGURE 10.1 – Suivi sonore naturel (à gauche) et guidage par son spatialisé (à droite).

De plus, pour s'adapter à la pratique sportive, le dispositif mis en œuvre devra être portable, léger et robuste.

### 10.1 Conception générale du système

Le dispositif complet, qui est présenté dans la figure 10.2, comprend un head-tracker, un système de localisation en temps réel et une unité de traitement.

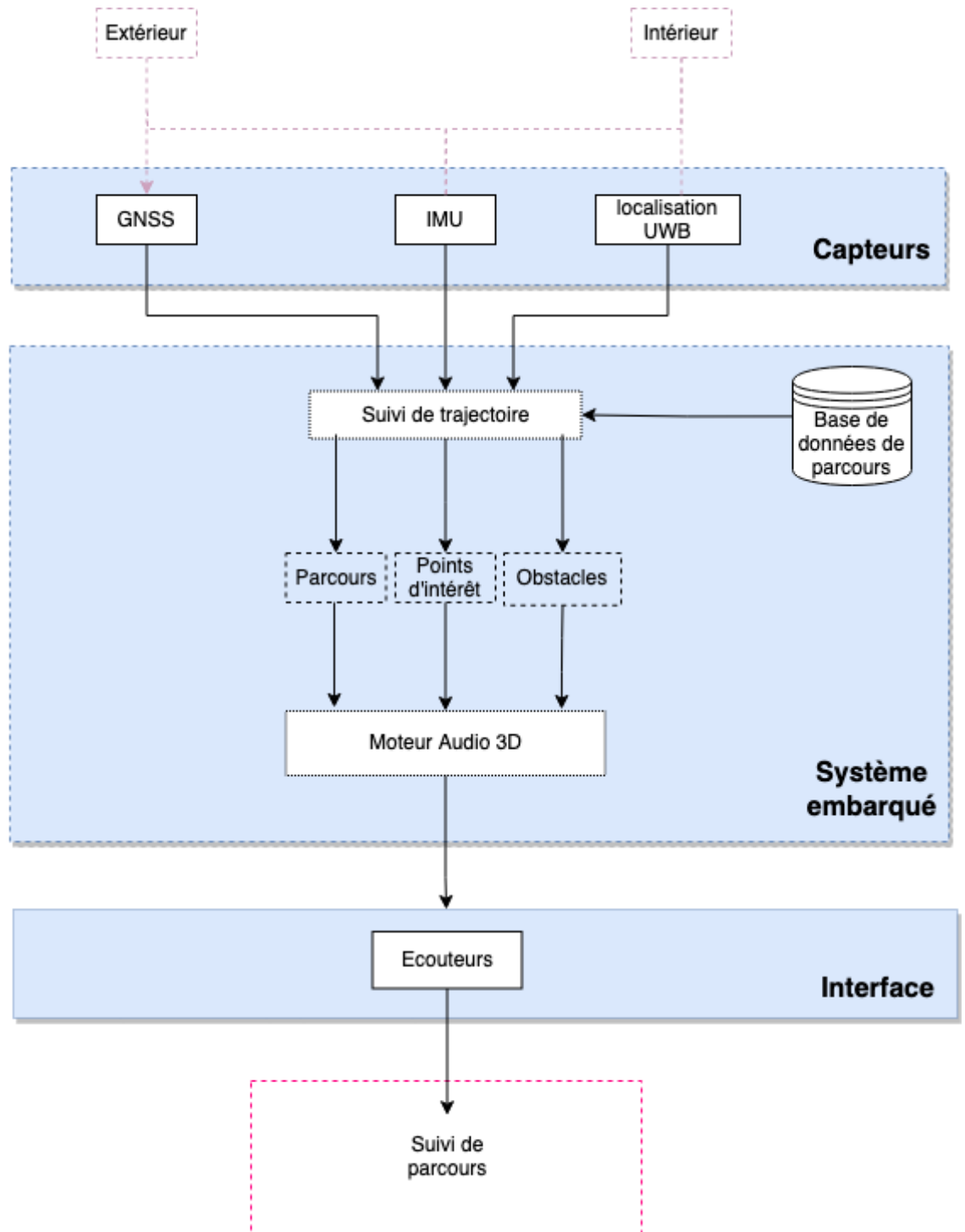


FIGURE 10.2 – Fonctionnement général du système proposé.

## 10.2 Architecture

### 10.2.1 Conception logicielle

Nous avons cherché à rendre le logiciel modulaire tout en minimisant la latence associée aux communications entre les différentes couches logicielles et matérielles. Pour cela, nous avons conçu une architecture multi-tiers où chaque composant logique échange des données via une API REST [83].

Le programme principal agit comme un serveur REST et gère les communications avec les clients. Il comprend également l'interface de communication avec les capteurs (head-tracker et système de positionnement) et les systèmes de suivi et de filtrage.

Le moteur binaural est un client qui obtient en permanence des données de localisation du serveur via des requêtes HTTP REST. Le serveur REST agit comme une couche d'abstraction pour les capteurs, de sorte que le moteur binaural obtient les informations de localisation indépendamment de la technologie de localisation (GPS, systèmes RTLS, etc) et quel que soit son taux de rafraîchissement.

Un deuxième client REST assure le suivi du chemin, la visualisation et l'enregistrement des données. Il est à noter qu'avec cette architecture REST basée sur HTTP, les clients peuvent être hébergés sur un ordinateur distant ; en particulier, la visualisation ou l'analyse des données pourraient bien être transférées sur une autre machine.

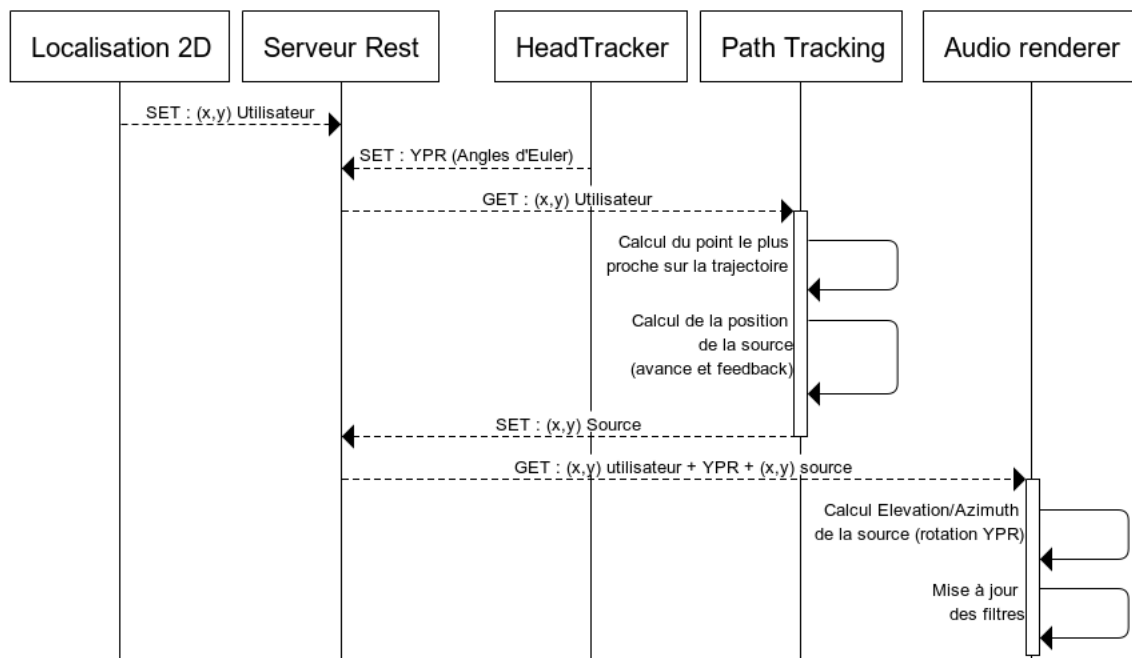


FIGURE 10.3 – Échanges de messages entre modules logiciels.

La figure 10.3 montre les échanges de messages et l'organisation des différents processus autour du serveur Rest. En séparant chaque tâche en différents processus, le logiciel est capable d'utiliser efficacement des processeurs embarqués multi-coeurs.

## 10.2.2 Conception matérielle

### — La carte mère :

Le processeur devant effectuer simultanément les tâches de traitement audio et de localisation en temps réel, nous utilisons un ordinateur à carte unique haute performance basé sur ARM. La modularité de la conception logicielle du système permet l'utilisation efficace d'un système multicœur et l'architecture big.LITTLE des processeurs ARM v7 est bien adaptée à notre objectif. Le moteur audio nécessite de stocker un grand nombre de filtres HRTF (paire de filtres tous les 2° couvrant une sphère complète). Pour des raisons de performances, tous ces filtres sont pré-chargés en mémoire au démarrage, ce qui, selon la taille des filtres, peut représenter jusqu'à 10 Mo de stockage. Cela disqualifie un grand nombre de processeurs de traitement du signal (DSP) et nous avons donc choisi d'utiliser des processeurs ARM d'usage général, mais les cartes Raspberry PI 4b et autres cartes de développement ARM peuvent également convenir.

### — Le module d'alimentation :

La batterie doit fournir l'énergie pour tous les modules, soit environ 15 W en état de marche. Nous utilisons une batterie LiPo de 3000 mA·h 2 cellules (7,4V) connectée à un convertisseur DC-DC.

### — Interface capteurs et périphériques d'entrées :

Les capteurs (GNSS, head-tracker), de même qu'une carte d'interface à microcontrôleur dotée de boutons sont interfacés en USB ce qui permet d'utiliser un matériel standard. Un inconvénient du bus USB est la latence de réception des messages série qui est de l'ordre de 1ms.

Le système est doté de plusieurs périphériques d'entrée qui permettent le pilotage et la prise en main du dispositif.

- Des boutons poussoirs, facilement accessibles aux personnes déficientes visuelles permettent de déclencher les différentes phases de fonctionnement (enregistrement de parcours, guidage, etc.).
- Pour la mise au point et les tests, une dizaine d'interrupteurs (dip switch) permet de sélectionner les paramètres de guidage (type de feedback, signalisation des virages, etc.). L'ergonomie de ce système n'est pas optimale et il pourrait être remplacé, par exemple par des menus audio.
- Une prise en main totale du dispositif est possible via le protocole SSH et un réseau wifi ad hoc en utilisant un ordinateur portable. Le réseau wifi peut permettre également d'interagir directement avec le serveur Rest embarqué, ce qui ouvre la possibilité de créer des interfaces sur smartphones ou sur ordinateur.

Le dispositif est connecté en USB aux capteurs de localisation. Pour la **localisation en extérieur** nous utilisons une paire de récepteurs RTK-GNSS qui nous permet d'obtenir une précision décimétrique ou un récepteur GNSS multifréquence (ublox f9P), plus simple à mettre en œuvre, mais avec une précision métrique.

- Pour le système RTK, le plus précis, nous utilisons les modules U-blox M8P ou F9P. La base (visible fig. 10.5) est installée en condition *open-sky*. Elle est configurée pour recevoir toutes les constellations disponibles. Elle est ainsi laissée en position parfaitement statique pendant au minimum quelques dizaines de minutes de façon à obtenir un positionnement parfaitement stable. Une fois la solution suffisamment stable, la base émet les informations de correction sous forme de messages RTCM 10402.3 (Radio Technical Commission for Maritime) via un lien radio 433MHz (pour le M8P) ou ZigBee (IEEE 802.15.4) à 2.4GHz pour le F9P. Ces corrections RTCM sont envoyées à une fréquence de 1Hz (indépendamment de la fréquence de sortie du récepteur GNSS mobile). Le récepteur mobile (*Rover*) est lui doté d'une antenne GNSS plus portable et d'une autre antenne destinée à recevoir les corrections RTCM. La portée de réception de ces corrections est de quelques centaines de mètres en champ libre pour ces deux systèmes.

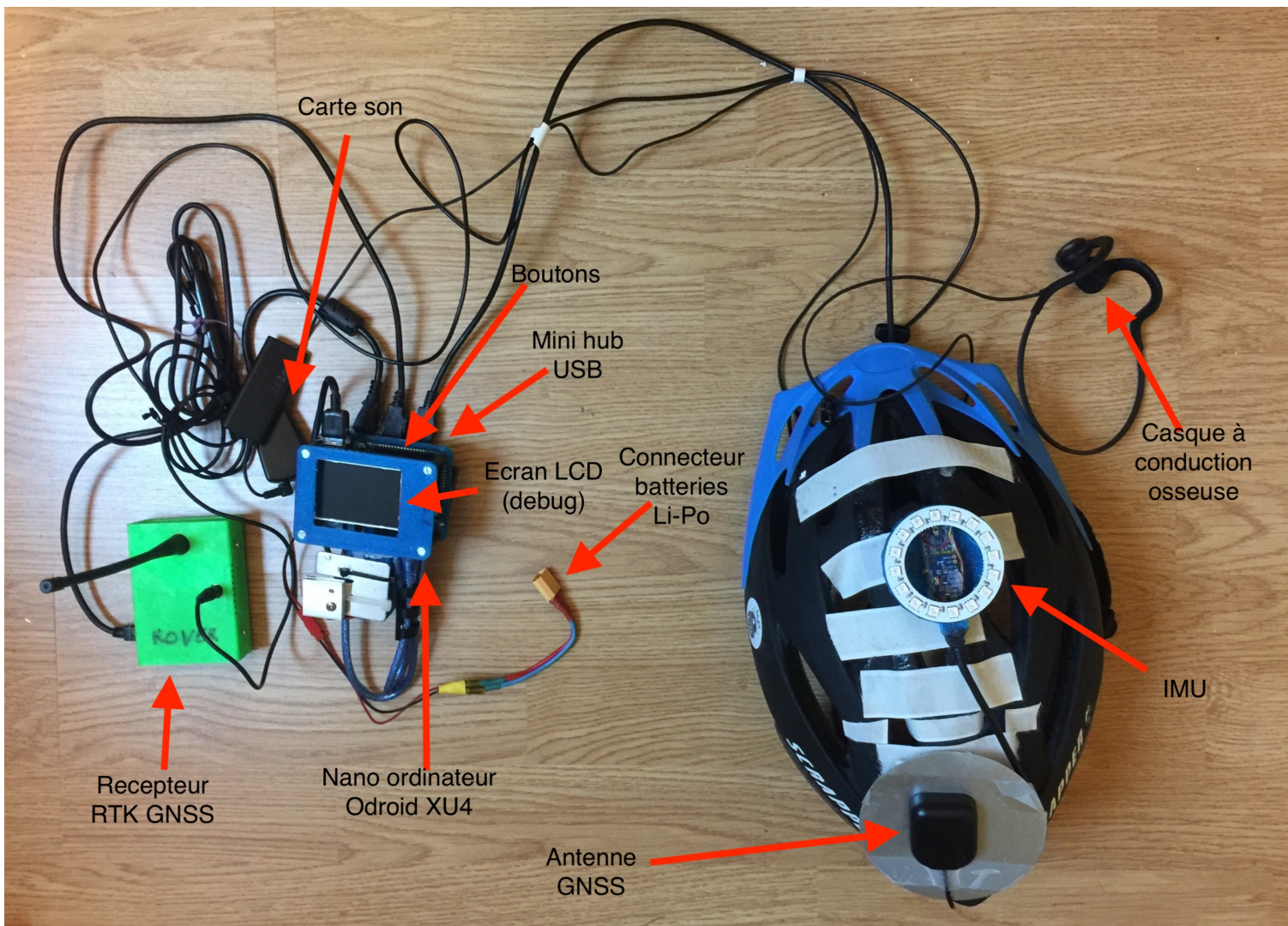


FIGURE 10.4 – Le prototype au complet (sans les batteries).

— Dans le cas d'un récepteur GNSS *stand-alone*, nous avons expérimenté un récepteur Ublox F9P en mode bi-fréquence L1/L2 GPS+E1/E5b Galileo + SBAS/EGNOS offrant une précision métrique dans nos cas d'usage (ciel largement visible, peu de trajets multichemin).

Pour la **localisation en intérieur**, nous utilisons des réseaux de capteurs UWB avec le protocole décrit en première partie qui permet une précision décimétrique et une bonne robustesse en intérieur. Le transceiver UWB mobile est là encore connecté via le port USB et vu par le système comme un périphérique série.

**Encombrement** : Le dispositif a été conçu pour être transportable et suffisamment robuste pour une utilisation en contexte sportif. Nous n'avons pas cherché à pousser la miniaturisation du dispositif, ce qui sortirait du champ de ce travail. Hors casque, le dispositif est donc constitué de l'unité de calcul qui représente une boîte d'environ 10cm x 8cm x 5cm ( 300g), du récepteur GNSS de taille comparable ( 200g) et d'une batterie dont la taille et le poids dépendent





FIGURE 10.5 – Récepteur *Base* RTK, doté de son antenne dôme et d'un récepteur avec émetteur radio 433GHz pour l'envoi des corrections RTCM

de l'autonomie nécessaire (typiquement 200g). L'ensemble est logeable dans un sac à dos (sac de trail qui peut être maintenu près du corps).

### 10.3 Restitution audio

Lors des tests préliminaires utilisant un casque audio fermé ou semi-ouvert les utilisateurs non-voyants ont indiqué un inconfort lié à la perte d'informations sonores habituellement perçues. D'une part ils étaient isolés en percevant mal les conversations ou les consignes et d'autre part ils étaient privés des indices sonores les renseignant sur l'environnement (réverbération, sons ambiants, écho-localisation). Pour éviter ces désagréments, nous nous sommes orientés vers des casques à conduction osseuse grand public (dit à conduction osseuse).

Ces casques sont des dispositifs non invasifs dotés de vibrateurs généralement mis en contact avec les tempes de l'utilisateur. La conduction sonore se fait alors en partie par l'intermédiaire de l'os temporal mais également par la peau et par voix aérienne sur les modèles grands publics utilisés. Certains dispositifs utilisés comme aide auditive sont eux placés sur le mastoïde. Les vibrateurs osseux sont similaires dans leur fonctionnement aux transducteurs à conduction aérienne comme des haut-parleurs. La principale différence entre ces deux groupes de transducteurs est l'impédance du milieu face au transducteur. Afin d'assurer une transmission efficace de l'énergie entre le transducteur et la structure de la tête, les impédances des deux éléments doivent être proches et c'est pourquoi la surface des vibrateurs est généralement assez dure.

La perception spatiale du son à travers un casque à conduction osseuse est un sujet qui fait l'objet de discussions (voir ci dessous). En effet, deux problèmes se posent : d'une part (1) la restitution sonore des casques à conduction osseuse est souvent imparfaite, d'autre part (2) le son peut être conduit à travers la tête, et un vibrateur risque donc de stimuler les deux cochlées en même temps.

1. L'apparition de casques à conduction osseuse pour le grand public a permis le développement de cette technologie avec des améliorations sensibles en matière de restitution sonore et de confort. Ces casques sont conçus

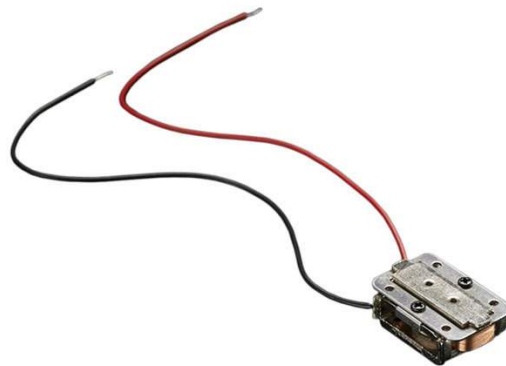


FIGURE 10.6 – Exemple de vibreur utilisé en conduction osseuse (Adafruit technology)

pour l'écoute musicale et présentent, pour certains, une réponse en fréquence comparable à des oreillettes traditionnelles. Nous avons sélectionné le casque *AfterShokz Sportz titanium*. Sa réponse en fréquence est relativement plate ce qui est susceptible de le rendre adapté à la technologie binaurale (c.f. Fig. 10.7 courbe de réponse en fréquence du modèle AfterShokz Sport 3 vs oreillettes Invisio X5) et par ailleurs c'est l'un des rares modèles filaires du marché. En effet, les modèles sans fil sont à proscrire en raison de la latence introduite par la transmission audio Bluetooth bufferisée. À noter que, comme la plupart des modèles grand public, c'est en réalité un casque à conduction hybride aérienne/osseuse puisqu'en l'absence de contact avec la peau, le son est perçu à un niveau largement audible.

2. Le problème de la stimulation bilatérale des oreilles internes par chaque vibreur est susceptible de limiter les performances d'une écoute spatialisée. Le problème est particulièrement visible pour les dispositifs à conduction osseuse stimulant le mastoïde, c'est le cas des aides auditives fixées sur cette partie du crâne. Dans le cas des dispositifs non invasifs placées face à l'os temporal, l'atténuation et le retard transcrâniens peuvent être suffisants pour permettre à l'auditeur de séparer les signaux. L'atténuation du son du vibreur contralatéral par rapport aux vibrations du transducteur ipsilatéral a été évalué à environ 10-15dB à 2kHz [206]. Sur les performances de localisation, plusieurs études démontrent clairement la possibilité de localiser des sons spatialisés en utilisant la conduction osseuse. MacDonald et al. [152] ont montré que les capacités de localisation dans le plan azimutal sont très comparables entre un casque à conduction aérienne et un casque à conduction osseuse. Barde et al. [19] notent un effet de compression dans la localisation verticale (perception concentrée sur  $-15$  à  $15^\circ$  et une bonne sensation d'externalisation chez les sujets. Bujacz et al. [40] utilisent des écouteurs à conduction osseuse pour l'orientation de personnes aveugles et notent une légère perte de précision par rapport à un casque classique.

**Stimuli :** Les stimuli utilisés doivent être (1) appropriés pour être localisés efficacement et (2) confortables pour la pratique d'une activité sportive. Pour une bonne efficacité de localisation, le son doit permettre de fournir de bons indices



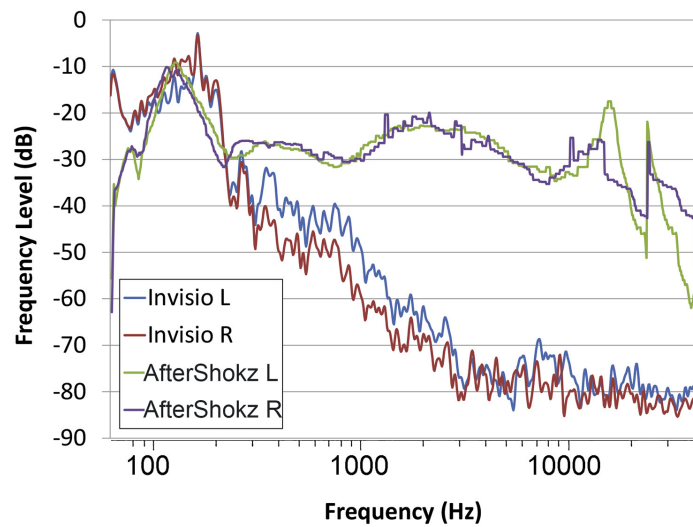


FIGURE 10.7 – Réponse en fréquence du modèle AfterShokz Sport 3 vs oreillettes Invisio X5 (d'après [159])

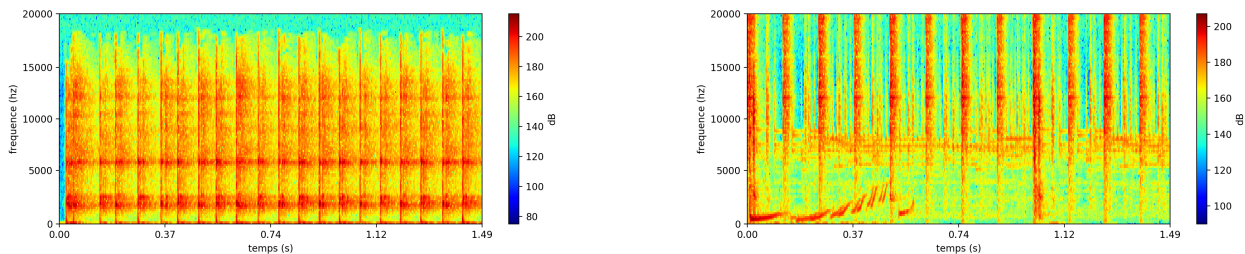


FIGURE 10.8 – Gauche : spectrogramme d'un son de castagnettes, Droite : son hybride rythmé et répétitif créé par un ingénieur du son.

de localisation binauraux et monauraux. Ainsi le signal, doit autant que possible, avoir un spectre complexe [45, 107] et être large bande. Nous avons cependant considéré qu'il ne fallait pas sélectionner les stimuli uniquement sur des critères fonctionnels. Il est important, en lien avec les utilisateurs de choisir des sons à la fois attractifs et surtout non fatigants. Les recherches sur la *ludification* [112, 161] ont montré que pour des applications de santé ou de réhabilitation, l'utilisation de concepts ou de stimuli attractifs ou amusants contribuait à améliorer la motivation, l'engagement des utilisateurs et finalement l'efficacité des applications, et ceci d'autant plus que les expériences ou les exercices peuvent être répétitifs et potentiellement ennuyeux. Nous avons testé des stimuli musicaux, a priori confortables par nature, mais les utilisateurs ont fait remarquer qu'ils captaient trop leur attention et les empêchaient de se concentrer suffisamment sur l'orientation. Suite à des tests informels et des discussions avec les utilisateurs, les stimuli répétitifs et rythmés comportant un ou plusieurs types de sons ont semblé moins fatigants que les sons continus. Nous avons utilisé des sons de percussions de façon satisfaisante. Le spectrogramme d'un son de castagnette que nous avons utilisé est donné en figure 10.8. Il comporte les caractéristiques spectrales adaptées à une bonne localisation. En revanche, ce type de sons n'est pas très attractif. Nous avons donc également utilisé un son composite créé par un ingénieur du son comportant des motifs rythmés et répétitifs rappelant des bruits de chaîne de vélo, de percussions et un métronome.

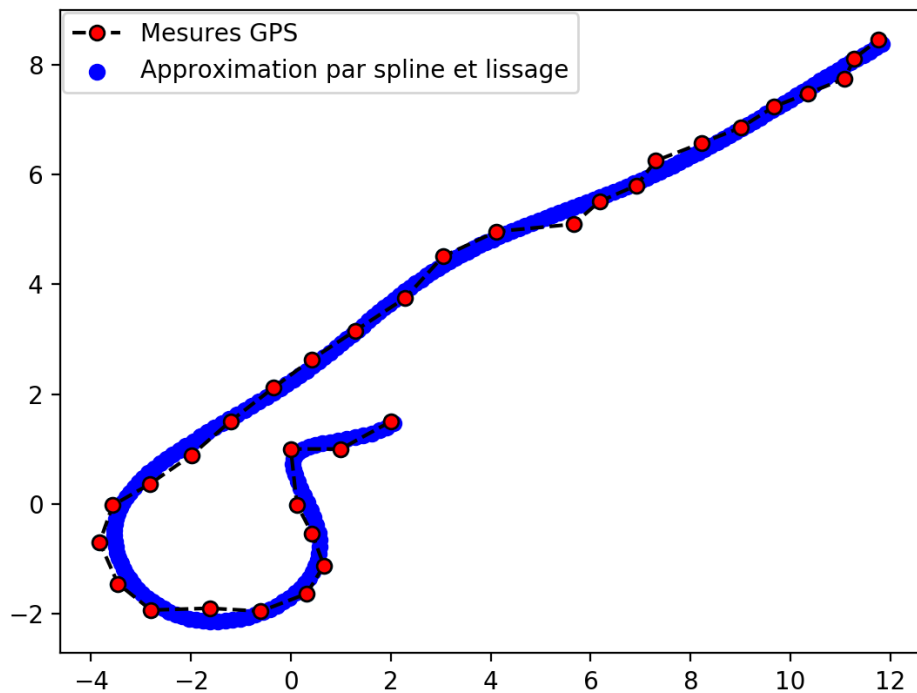


FIGURE 10.9 – Mesure GPS (RTK) et approximation sous forme de splines (mesures en m)

## 10.4 Gestion du parcours

### 10.4.1 Enregistrement du parcours

Le parcours à réaliser (tour de piste, déplacement en intérieur) est enregistré au préalable par une personne valide. Lors de cette étape, qui se fait généralement à la vitesse de la marche, les coordonnées sont enregistrées typiquement à une cadence de 1Hz. Le programme permet alors de gérer des parcours elliptiques ou de formes quelconques. Les parcours elliptiques sont adaptés aux terrains ouverts sans chemin tracé (par ex. roller dans un gymnase, course sur un terrain de foot). Dans ce cas on effectuera un fitting de la trajectoire par méthode des moindres carrés. On obtiendra alors un parcours très régulier même avec un enregistrement approximatif. Si le parcours est d'une géométrie quelconque (typiquement suivre un chemin, évoluer sur une piste), les données sont approchées par des splines ce qui permettra à la fois de les lisser et de permettre une interpolation efficace du parcours. La figure 10.9 montre des mesures issues d'un GPS RTK sous forme brute (sans filtrage) ainsi qu'un lissage et une approximation par des splines.

### 10.4.2 Suivi du parcours

La méthode de suivi de parcours consiste à trouver le point le plus proche de la position de l'utilisateur sur la courbe paramétrée issue du fitting de la trajectoire enregistrée. On peut alors minimiser la distance à la courbe par une fonction d'optimisation, ou chercher le point le plus proche après discrétisation. Cette méthode est efficace en pratique dans la mesure où le pas de discrétisation reste raisonnablement large (nous utilisons typiquement un pas de 10cm). Une fois le point trouvé, on peut alors chercher un point en avance d'une distance  $d$  sur le parcours où sera effectivement placée la source sonore. La valeur de  $d$  ne doit pas être trop petite car une brusque accélération pourrait conduire à dépasser la source, ni trop longue car cela pourrait conduire à gommer ou décaler certaines variations du parcours. Cette distance peut être soit fixe (typiquement deux à trois mètres), soit dynamique. Dans une approche dynamique, on cherchera à rapprocher la source de l'athlète dans les virages et l'éloigner dans les lignes droites. Cette approche présente plusieurs avantages :

- Le moteur audio prenant en compte la distance, l'augmentation de volume permet d'avertir l'utilisateur d'une zone de virage ;
- le guidage sera plus précis dans la zone de virage ;
- dans une moindre mesure, le système sera plus robuste à une accélération dans une zone de ligne droite.

### 10.4.3 Courbure

L'évaluation de la courbure du parcours permet aussi d'identifier les zones rectilignes et les zones de virages ce qui permet d'adapter la stratégie de guidage au fur et à mesure du parcours. En particulier, en tenant compte de cette information on pourra alors jouer sur le type de feedback, l'éloignement de la source et la signalisation des virages.

La figure 10.10 montre un parcours comportant des zones rectilignes et des virages. On calcule la courbure en chaque point et on identifie des zones contiguës de courbure faible ou importantes. On obtient alors une segmentation du parcours en zone *rectilignes* et en zone *courbées*.

## 10.5 Stratégie de suivi du parcours

Dans nos expériences, à chaque instant nous mesurons la position de l'utilisateur et estimons la position idéale qu'il devrait avoir sur le parcours. L'erreur courante correspondant à la distance entre ces deux positions est notée  $e_t$ .

### 10.5.1 Preuve de concept, suivre un parcours inconnu

Dans une première expérience (figure 10.11), une personne voyante aveuglée doit suivre un trajet sinueux non connu, avec des virages doux (typiquement une somme de sinusoides dont on fait varier phases et fréquences à chaque essai). Dans cette expérience, l'utilisateur évolue dans un espace de 8mx12m. Dans ces conditions, sans entraînement préalable, la personne est capable de suivre le trajet, mais ses déplacements restent limités à une faible vitesse (marche lente, environ 3-4km/h). Il apparaît que suivre un trajet très sinueux de géométrie inconnue est difficile et nécessite une importante concentration. Dans une perspective de prise de vitesse, nous serons donc amenés à proposer soit des parcours plus réguliers (tourner sur un stade), soit des mécanismes destinés à mieux anticiper les virages. Il est à noter que l'apprentissage joue ici un rôle important mais difficile à quantifier expérimentalement.

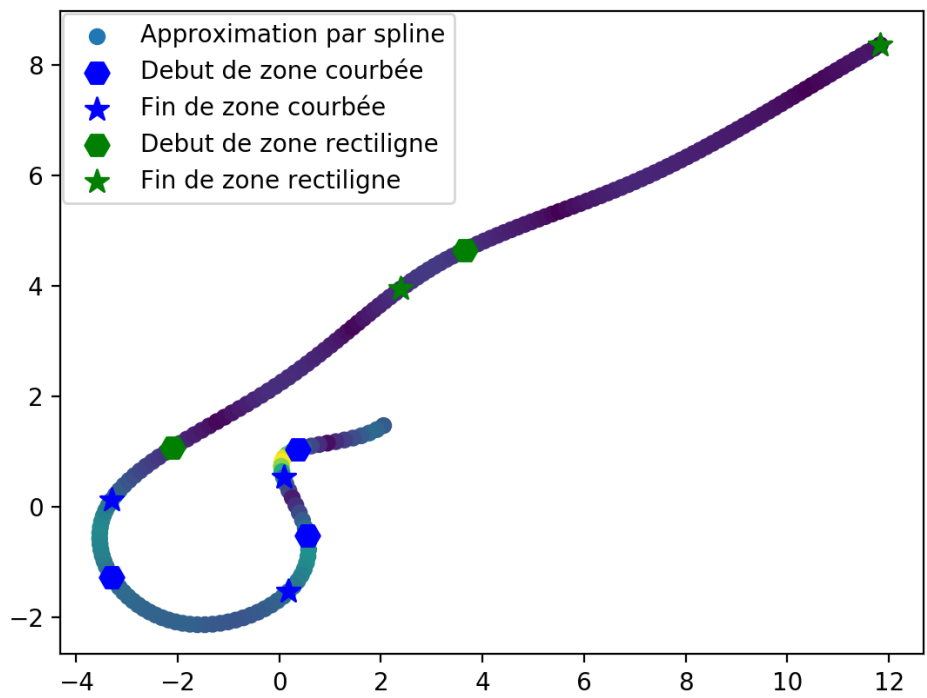


FIGURE 10.10 – Identification des zones sur le trajet.

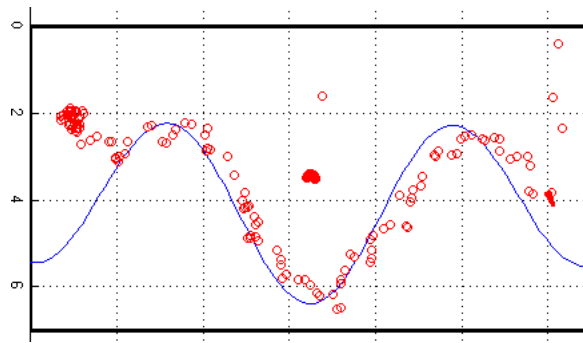


FIGURE 10.11 – Déplacement le long d'un trajet inconnu. Le parcours à suivre est matérialisé par la sinusoïde bleu et les cercles rouges représentent la position de l'utilisateur. Départ en haut à gauche.

### 10.5.2 Contrôle en boucle fermée

On peut supposer, qu'en cas de déplacement sur un parcours sinueux, l'utilisateur n'a pas toujours assez de temps pour localiser le son avec précision, ce qui pourrait expliquer l'augmentation de l'erreur quand l'utilisateur se déplace plus rapidement. Pour résoudre ce problème, nous introduisons une rétroaction proportionnelle.

Pour une erreur de positionnement  $e_t$ , nous introduisons un feedback :

$$F_t = K_p e_t$$

qui va déplacer la source de  $F_t$  dans la direction opposée à l'erreur de la position théorique sur le parcours (voir Fig. 10.12). Le paramètre  $K_p$  est le coefficient de rétroaction proportionnel et peut être ajusté empiriquement.

Par ailleurs, dans certains cas, l'utilisateur suit le chemin sans jamais l'atteindre (car la précision de la perception des sources sonores est limitée). Pour résoudre ce problème, à la manière d'un contrôleur PID [124], on peut aussi ajouter un terme qui tient compte de l'accumulation de l'erreur sur une fenêtre temporelle (semblable à celle du contrôleur proportionnel-intégral bien connu) :

$$F_t = K_p e_t + K_i \sum_j e_j.$$

Il semble raisonnable de n'activer ce type de feedbacks que dans les zones à faible courbure.

## 10.6 Tests du prototype

L'évaluation d'un dispositif de guidage soulève plusieurs problèmes. D'une part elle ne peut se faire qu'après la construction d'un prototype complet donc assez tard dans le projet, et d'autre part l'utilisateur doit être immergé dans une scène naturelle complexe qui rend difficile l'ajustement et l'évaluation des modes de guidage.

Pour résoudre certains de ces problèmes, des travaux ont cherché à reproduire une tâche de navigation par le déplacement d'un avatar dans un environnement virtuel en 3D [128, 137]. Cependant, dans de telles configurations, les retours sensoriels de l'utilisateur sont limités. Un test en condition semi-statique ne permet pas de prendre en compte la perception sensori-motrice et ne permet pas un couplage action-perception totalement efficace. C'est pourquoi nous

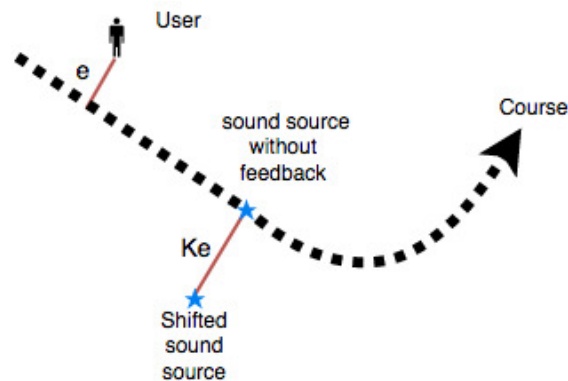


FIGURE 10.12 – Illustration du système de feedback proportionnel. Il dépend de la position de l'utilisateur par rapport au parcours à suivre. La source sonore est alors déplacée à une distance  $K_e$  à l'opposé de la position de l'utilisateur par rapport au parcours afin de l'inciter à corriger son erreur plus rapidement.

avons opté pour une approche en trois étapes. Dans un premier temps nous avons évalué l'efficacité des stimuli binauraux dans une situation semi-statique (partie 2 de ce document). En effet, le déplacement de la source sonore est par beaucoup d'aspects comparable aux tâches d'orientation lors d'un parcours. Dans un deuxième temps nous pré-évaluons plus spécifiquement nos techniques de guidage (contrôle de trajectoire, signalisation des virages, etc.) encore une fois en conditions semi-statiques (c.f. paragraphe suivant). Dans un troisième temps enfin, nous évaluerons notre dispositif dans des conditions de guidage réelles mettant alors en œuvre l'ensemble des boucles action-perception.

**Méthodologie pour les pré-tests :** Les pré-tests consistent à parcourir virtuellement un parcours (qui peut être un parcours pré-enregistré grâce à une trace GPS) en étant équipé d'un casque de guidage audio et d'un head-tracker. Pour simuler sa progression le long du parcours, l'utilisateur se déplace virtuellement avec une vitesse constante. Suivant le mode de fonctionnement choisi, soit il peut orienter sa trajectoire à gauche ou à droite à l'aide d'un clavier, ou bien on adapte la trajectoire en fonction de l'orientation de sa tête. La deuxième solution permet un fonctionnement plus naturel, mais oblige l'utilisateur à pointer vers le son ce qui n'est pas toujours le cas dans la réalité. Ces pré-tests permettent en particulier d'évaluer les stratégies de guidage sans mettre en place un dispositif en extérieur qui nécessite une logistique plus importante.

## 10.6.1 Expériences de guidage en condition réelles

### 10.6.1.1 Ligne droite

Sans vision, les indices allothétiques (proprioception, perception sensori-motrice) peuvent permettre de se déplacer de façon assez rectiligne. Cependant, sans référence extérieure, les personnes sont souvent sujettes à une dérive assez sensible. Pour illustrer la nécessité de guider des personnes, y compris sur les zones rectilignes, nous avons demandé à quelques sujets, valides placés en situation de cécité, de se déplacer en ligne droite sur environ 25m (Fig. 10.13 gauche). La plupart des sujets, voyants ou non, commencent à dévier de plusieurs mètres après seulement 10 à 15m. Il est donc tout à fait utile et important de pouvoir guider les personnes aveugles dans les lignes droite pour éviter cette dérive.

Pour tenter d'expliquer ce problème de dérive (appelé *veering* dans la littérature) de nombreuses études (citées

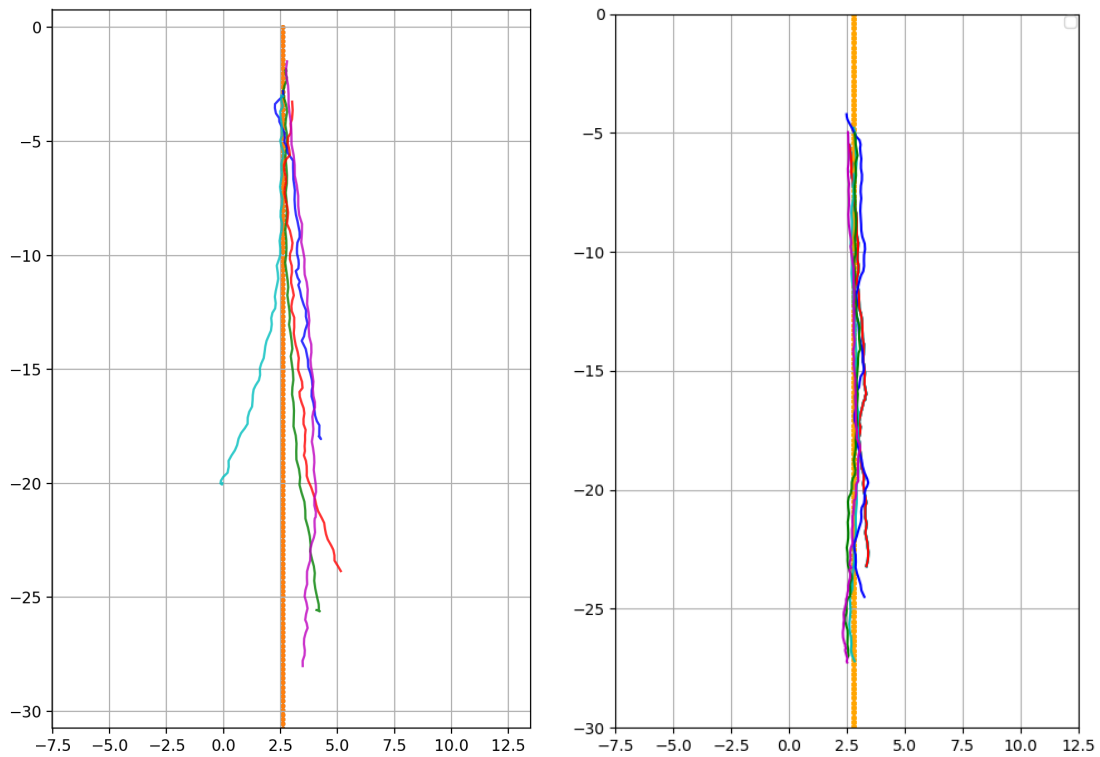


FIGURE 10.13 – A gauche, des sujets aveuglés se déplacent en ligne droite. A droite, le résultat d'un guidage (feedback  $p=0.5$ ).

par [97]) mettent en avant les différences de force, de longueur des jambes et les asymétries biomécaniques lors de la marche. Mais la capacité à marcher en ligne droite semble également liée à la qualité des informations sensorielles de direction [184] et à la capacité à exécuter des mouvements dans une direction voulue. Les indices vestibulaires, proprioceptifs ou kinesthésiques sont utiles pour détecter une rotation et donc une déviation par rapport à la trajectoire rectiligne [50].

Nous avons effectué des tests en ligne droite en mettant en œuvre le guidage avec 5 sujets préalablement familiarisés avec le dispositif. Dans tous les cas la source sonore est placée en avance de deux mètres sur le trajet. Pour chaque sujet nous utilisons trois méthodes différentes : sans feedback, avec un feedback proportionnel ( $K_p = 0.5$ ) et avec un feedback proportionnel et intégral (fenêtre de 3m et  $K_p = 0.5$  et  $K_i = 0.5$ ). Le tableau 10.1 récapitule les performances pour 5 sujets. D'une manière générale le guidage permet de conserver une trajectoire rectiligne y compris pour les utilisateurs qui en étaient incapables sans guidage.

Sujet	Sans feedback		Feedback P		Feedback P+I	
	RMS	Max	RMS	Max	RMS	Max
Sujet 1	1.24	2.57	0.67	1.29	<b>0.37</b>	<b>0.64</b>
Sujet 2	1.54	2.37	<b>0.52</b>	1.35	0.72	<b>1.25</b>
Sujet 3	0.92	1.36	<b>0.46</b>	<b>0.99</b>	0.73	1.22
Sujet 4	0.5	1.51	0.36	0.61	<b>0.30</b>	0.71
Sujet 5	0.61	1.65	0.32	0.71	<b>0.29</b>	<b>0.56</b>

TABLE 10.1 – Déplacement sur une trajectoire rectiligne de 5 utilisateurs.

L'influence du feedback proportionnel sur la précision du guidage est assez visible. L'effet du feedback I est lui moins significatif, il apparaît néanmoins dans des situations plus spécifiques quand l'utilisateur longe la trajectoire idéale.

### 10.6.1.2 Parcours ayant une géométrie connue : apprentissage et performance

Pour permettre aux utilisateurs d'atteindre des vitesses plus élevées, compatibles avec une activité physique ou sportive, nous avons testé le dispositif sur une boucle fermée de forme elliptique.

Après un phase d'apprentissage préalable de quelques minutes, les utilisateurs sont invités à se déplacer plus ou moins rapidement selon leur niveau de confiance dans le dispositif.

Les figures 10.14 et 10.15 présentent le cas d'un utilisateur à roller, avec et sans feedback, mettant ici en évidence une précision supérieure avec feedback. Le tableau 10.2 montre les résultats obtenus par plusieurs utilisateurs, aveugles ou mis en condition de cécité, à pied ou en roller sur un parcours comportant plusieurs tours (quelques centaines de mètres). Dans tous les cas le feedback a permis un gain de précision notable.

Il faut noter que nous n'avons pas de méthode pour déterminer les coefficients optimaux. Les coefficients ont volontairement été fixés à des niveaux assez faibles pour éviter les oscillations qui se manifestent par un changement de direction de la source qui devient rapidement très perceptible et gênant.



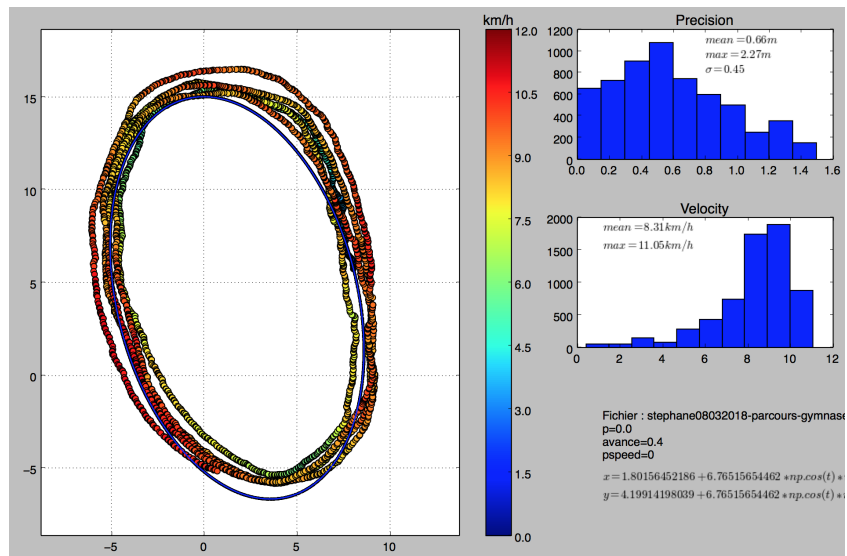


FIGURE 10.14 – Roller sur un parcours en boucles elliptiques (Personne aveugle),  $K_p = 0$ .

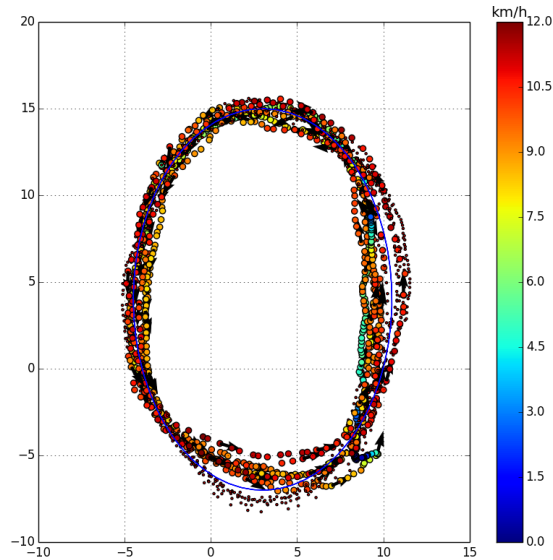


FIGURE 10.15 – Roller sur un parcours en boucles elliptiques feedback proportionnel-integral activé  $K_p = 0.7$  et  $K_i = 0.1$  (Personne aveugle).

Sujet	Activité	Err. Moy. sans PI	Err. Moy. avec PI
Personne aveugle #1	roller	0.66 m	0.5 m
Personne aveugle #2	marche	0.37 m	0.16 m
Personne valide aveuglée #1	marche rapide	0.8 m	0.7 m
Personne valide aveuglée #2	marche	0.88 m	0.56 m

TABLE 10.2 – Comparaison des erreurs mesurées lors de guidage sur un parcours elliptique répétitif (plusieurs tours effectués à chaque fois), avec feedback activé ou non.



FIGURE 10.16 – Vitesse instantanée mesurée lors d’une course réalisée par un athlète mal voyant aveuglé sur une piste d’athlétisme (distance parcourue légèrement supérieure à 800m). Enregistrement du parcours et résultats obtenus par l’application de suivi sportif *runkeeper* sur téléphone portable (distance parcourue en abscisse). La vitesse moyenne était supérieure à 12km/h avec deux accélérations au delà de 14km/h.

Par rapport aux expériences sur un parcours sinueux et inconnu du sujet, on voit que la prise de vitesse est possible puisqu’un sujet a pu atteindre la vitesse de 12km/h. Par ailleurs l’erreur de suivi reste très contenue, le même utilisateur ayant obtenu une erreur moyenne d’environ 0,5 m (maximum 1,3 m,  $\sigma = 0,43$ ) lors d’une séance sportive où il a parcouru plusieurs centaines de mètres.

Dans une optique de recherche de performances, nous avons accompagné un athlète très mal voyant (pratiquant habituellement la course à pied et le triathlon handisport) sur une piste d’athlétisme de 400m. L’athlète s’est placé en situation de cécité (yeux fermés) et nous avons progressivement augmenté le rythme lors d’une séance d’apprentissage très progressive d’environ une heure. Les derniers kilomètres ont été parcourus à plus de 12km/h de moyenne avec plusieurs pointes au delà de 14km/h. Un graphique des vitesses sur les deux derniers tours (800m) est donné en Figure 10.16. Ce test a été réalisé sans activer de méthode de feedback, l’athlète était accompagné de guides valides (en retrait de chaque côté) pour signaler d’éventuels situations de danger ou en cas de déviation importante.

## 10.6.2 Techniques de guidage avancées

### 10.6.2.1 Parcours sinueux avec courbes et portions rectilignes

La plupart des parcours sont constitués de zones rectilignes et de zones de virage. C’est typiquement le cas sur une piste d’athlétisme. Pour ce type de parcours, nous proposons de moduler les paramètres de guidage (feedback, distance à la source) en fonction de la zone. La figure 10.17 illustre cette approche. Il s’agit d’un parcours sur une piste de roller comportant plusieurs virages de rayon de courbure différents. Notre programme a alors automatiquement identifié les différentes zones. Le dispositif de guidage a été paramétré pour activer le feedback intégral et éloigner la source sonore dans les lignes droite.

Le résultat de guidage d’une personne voyante mise en situation de cécité est visible sur cette figure. La personne, entraînée a alors pu atteindre une vitesse de l’ordre de 10-12km/h dans les lignes droites.

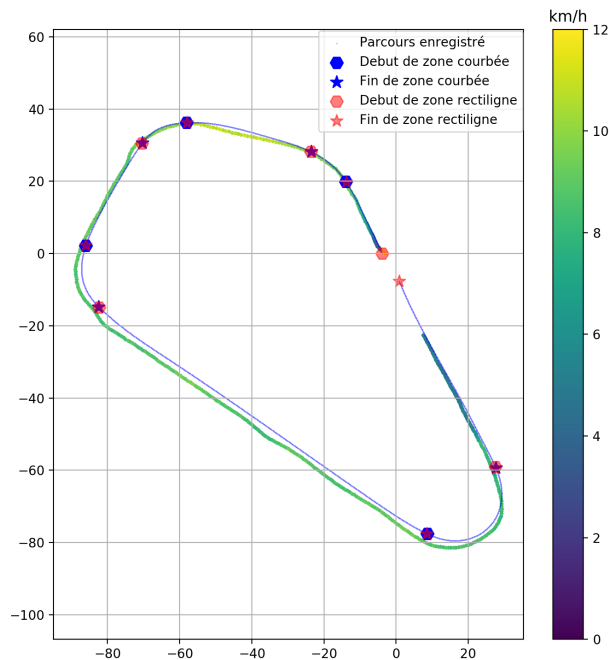


FIGURE 10.17 – Parcours sur une piste de roller (Breuillet 91). Le parcours à réaliser sur la piste (matérialisé par une ligne bleu) est analysé pour identifier les zones rectilignes et les virages. Un parcours réalisé à vélo par une personne valide aveuglée est visible sous forme de trace colorée en fonction de la vitesse.

### 10.6.2.2 Signalisation des zones de courbure

Sur piste, les coureurs valides et les athlètes DV accompagnés de guides, conservent une vitesse globalement constante dans les lignes droites et les virages (i.e. ils ralentissent peu dans les virages). Dans le cas d'un guidage sonore, nos tests ont montré que l'utilisateur va abaisser sa vitesse dans les courbes et en particulier en début de virage, le temps de bien en percevoir la courbure. Par contre, dans les lignes droites, l'utilisateur sera plus enclin à prendre de la vitesse s'il sait qu'il ne s'approche pas d'un virage. Ainsi, nous avons cherché à signaler l'imminence d'un virage, à la fois pour rassurer l'utilisateur dans les lignes droites et pour attirer l'attention avant la courbe.

Outre l'imminence d'un virage, il peut également être utile de donner des indications sur son rayon de courbure et sa direction. Pour cela nous proposons deux approches :

- Signaler l'approche ou le début d'un virage avec un ou plusieurs sons (cloche, etc.). Sur un parcours comme une piste d'athlétisme ou le type de virage est connu cette approche est suffisante pour fournir l'information utile à l'utilisateur.
- Placer une source sonore (spatialisée) matérialisant le virage en avance de quelques mètres. La direction de cette source sonore permettra alors de connaître la direction du virage et son azimuth afin d'estimer son rayon de courbure initial.

Nous avons testé ces deux approches et il s'est avéré que la seconde entraînait des confusions entre le son à suivre et le son de signalisation. En revanche, un son très ponctuel, éventuellement différent pour signaler chaque type de virage est peu perturbant. Ce système permet alors à un utilisateur qui ne disposerait d'aucun autre indice de savoir où il se situe dans un parcours répétitif.

### 10.6.3 Retour des utilisateurs

Les sujets déficients visuels ont été interrogés pour mieux comprendre leur expérience avec le dispositif. Il s'agissait d'évaluer leur impression générale, les difficultés rencontrées et l'influence du dispositif sur leur perception de l'environnement. Le tableau 10.3 récapitule les réponses de façon synthétique. Les échanges post-expérience étaient largement libres, mais certaines questions étaient posées systématiquement :

— **Avez vous éprouvé des difficultés à localiser et à suivre le son ? Cela nécessitait-il un effort de concentration ?**

D'une manière générale les sujets DV comme les sujets valides ont éprouvé peu de difficultés à s'orienter et à se diriger vers le son. Un sujet a été gêné par le head-tracking. Plusieurs sujets n'ont pas éprouvé d'externalisation sur la zone frontale sans qu'ils trouvent le problème pénalisant pour l'orientation.

— **Était-il plus difficile ou plus facile de suivre ce son plutôt que de suivre une personne bruyante dans la rue (personne qui traîne une valise, bruits de talons, etc.) ?**

La majorité des sujets n'a pas trouvé plus difficile de s'orienter vers le son que dans un contexte naturel. Un seul sujet, gêné par le head-tracking, a trouvé qu'il devait se concentrer plus que dans une situation naturelle. Au contraire, la majorité des sujets a trouvé que, comparativement aux situations de la vie quotidienne, il était plus facile de suivre le son du dispositif, en raison de la régularité du son et du volume sonore (au quotidien, les sons peuvent être masqués ou de niveau trop faible).

— **Votre perception habituelle de l'environnement (en particulier écholocalisation) était-elle perturbée par le dispositif ?**

Tous les participants ont apprécié le casque à conduction osseuse et aucun n'a été gêné pour percevoir les sons extérieurs. Les participants qui maîtrisent les techniques d'écholocalisation n'ont pas été gênés dans leur pratique de cette technique.

Sujet	Handicap	Difficulté à suivre le son (vs son naturel)	Perception de l'environnement
B.	malvoyant	peu difficile	Non altérée
Ty.	malvoyant	peu/moyennement difficile	Peu altéré
Al.	malvoyant	très peu difficile (pas plus difficile)	Non altérée <sup>1</sup>
S.	aveugle	peu difficile (pas plus difficile)	Non altéré <sup>1</sup>
A.	aveugle	peu difficile (pas plus difficile)	Non altéré <sup>1</sup>
M.	aveugle	peu/moyennement difficile (plus difficile)	Peu altérée <sup>1</sup>
T.	aveugle	très peu difficile (pas plus difficile)	Non altérée <sup>1</sup>

Notes :

Choix possibles pour la *difficulté* : Très peu difficile, peu difficile, moyennement difficile, difficile, impossible.

Choix possibles pour la *perception de l'environnement* : Non altérée, Peu altérée, Très Altérée, Totale-ment altérée.

<sup>1</sup> Y compris les perceptions liées à l'écholocalisation.

TABLE 10.3 – Difficultés rencontrées par des utilisateurs déficients visuels lors de l'utilisation du dispositif.

## 10.7 Conclusion de la troisième partie

Le dispositif de guidage que nous avons développé est inspiré de la pratique habituelle des personnes non voyantes qui ont souvent l'habitude de s'orienter et de se guider avec des sources sonores dans leur vie quotidienne et, pour certaines d'entre elles, lors d'activités sportives. Il est constitué de capteurs de position (typiquement GNSS en extérieur ou triangulation par réseau de capteur UWB en intérieur) et d'orientation de la tête, d'un casque audio et d'une unité de traitement portable faisant fonctionner les algorithmes de suivi de parcours et de spatialisation sonore. Pour s'adapter à la pratique sportive, il a été conçu pour être portable, relativement léger et robuste. Très tôt dans le projet, nous avons travaillé en lien avec des personnes déficientes visuelles, ce qui a permis de se concentrer sur les besoins et les difficultés des utilisateurs plus que sur les aspects purement technologiques (c'est par exemple ce qui a poussé au choix du casque osseux, qui préserve la perception de l'environnement, alors qu'il est susceptible de dégrader l'externalisation du son binaural). Nous avons expérimenté ce système avec une dizaine de personnes déficientes visuelles pour la pratique de différents sport, en particulier la marche, le roller et la course à pied, toujours en environnement maîtrisé. Il permet de guider des personnes avec précision, et par exemple de courir sur une piste d'athlétisme sans en sortir mais pas encore de rester sur un couloir unique. Les retours utilisateurs ont été très largement positifs, saluant l'aspect naturel de la méthode. Les utilisateurs ont, dans leurs grande majorité, pu appréhender le système presque immédiatement et se sentir à l'aise après seulement quelques minutes. Avec un entraînement un peu plus conséquent, et lorsque la sécurité était assurée par un accompagnement de proximité, les utilisateurs ont pu atteindre des vitesses de 12 à 14km/h correspondant à une activité physique très soutenue. En l'absence de détection d'obstacles ou d'alertes en cas de défaillance de localisation GNSS (qui reste à réaliser, mais qui sort du champ de ce travail), le dispositif ne permet pas de pratiquer en autonomie totale sur piste en toute sécurité. En revanche, il pourrait permettre à plusieurs personnes déficientes visuelles de pratiquer avec un seul accompagnateur qui n'est par ailleurs pas obligé d'être à proximité immédiate.

## **Quatrième partie**

# **Conclusion et Perspectives**



## Conclusion générale

Ce travail présente différents aspects du développement d'un dispositif électronique pour le guidage des personnes déficientes visuelles. À la base de ce travail nous avons fait l'hypothèse, directement inspirée par les discussions avec des personnes aveugles, que le son binaural pouvait constituer un moyen de guidage efficace et naturel pour la pratique sportive. Pour cela nous nous sommes appuyés sur deux constats. Tout d'abord, dans leur vie quotidienne, les personnes aveugles s'orientent souvent en utilisant des sources sonores. Elles sont donc bien entraînées à cette pratique qui pour elles est très naturelle. Par ailleurs, fort de l'expérience en audio de l'équipe, et des connaissances actuelles dans le domaine de la perception sonore, nous avons pensé que le son spatialisé permettait de présenter une information spatiale de façon efficace pour orienter les personnes en temps réel.

Ces développements nous ont conduit à nous intéresser à des thématiques variées en localisation, acoustique et systèmes embarqués.

Dans le domaine de la localisation et du suivi nous avons implémenté et développé différentes méthodes. Tout d'abord nous avons développé une méthode de calibration de réseau de capteurs et de latération temps réel robuste par balises UWB utilisant un filtre de Kalman (UKF). Nous avons également développé une méthode originale de localisation par réseau de radars Doppler à onde continue non modulée. Nous avons montré qu'il était possible d'utiliser l'amplitude du signal Doppler pour estimer la distance à un objet mobile. Nous avons alors implémenté un filtre particulière qui permet la localisation en temps réel par hybridation des données de distance, des mesures de vitesse radiale Doppler et du cap fourni par une centrale inertielle. Ces développements ont fait l'objet de communications en conférences internationales [79, 82].

Dans le domaine de l'acoustique et de l'audio binaural, nous avons cherché à mieux comprendre les capacités des personnes à localiser et à suivre un son en mouvement. Pour cela nous avons mené des expériences en utilisant des sons naturels et des sons spatialisés par audio binaural. Ces expériences ont pu montrer, que sur le plan azimutal, les stimuli audio spatialisés permettaient une localisation assez comparable (bien que légèrement inférieure) aux sons naturels, y compris avec des HRTFs non individualisées et interpolées. Par ailleurs, nous avons pu montrer que même sur le plan azimutal, les stimuli obtenus par convolution de HRTF étaient supérieurs au panning (ITD+ILD) pour les sons fixes et pour les sons en mouvement. Les résultats, encore partiels de cette étude ont fait l'objet d'une communication à *International Congress on Acoustics* [80]. Toujours dans le domaine audio, en nous appuyant largement sur les travaux antérieurs de l'équipe, nous avons implémenté des algorithmes efficaces pour la spatialisation sonore temps-réel sur des plateformes embarquées disposant de peu de ressources. Pour une mise en œuvre efficace, cette approche temps-réel a impliqué une compréhension approfondie des sources de latence qu'elles soient liées au head-tracking ou au sous-système audio des OS modernes.

Finalement nous avons mis en œuvre ces méthodes de localisation et ces techniques audio pour construire un dispositif fonctionnel de guidage. Ce dispositif a pour principe la technique de guidage adaptatif de type *fee clochette* où une source sonore précède continuellement la personne pour lui indiquer le chemin à suivre. Ce dispositif a été conçu en lien avec des personnes déficientes visuelles, dans une démarche itérative et avec une approche centrée sur les besoins utilisateurs. Nous avons mené différentes expériences de guidage avec des personnes aveugles en lien avec nos partenaires associatifs qui ont permis d'évaluer différentes stratégies de contrôle et de signalisation lors de séances de tests. Ces expériences ont ainsi pu confirmer que le son spatialisé constitue un outil efficace pour guider des personnes aveugles sans induire de charge cognitive pénalisante pour la pratique sportive. Ces travaux ont fait l'objet de plusieurs communications [78, 79, 81]. D'un point de vue pratique, le dispositif permet à des personnes aveugles de pratiquer la marche sportive, la course à pied ou le roller en autonomie partielle, y compris dans un contexte de



recherche de performance (pour l'amateur).

## Perspectives

### Les approches anti-déviations (anti-veering), guidage en couloir

Les personnes non-voyantes sont naturellement sujettes aux déviations incontrôlées de trajectoire de façon très comparable aux personnes voyantes [119], cependant la marche en ligne droite peut être perfectionnée par l'entraînement [96] et fait d'ailleurs partie des enseignements de locomotion couramment pratiqués.

Des dispositifs d'aide à l'apprentissage ont été développés, comme l'application *Walking straight* [181] ou un système de couloir sonifié [158]. Ils fonctionnent selon deux principes : soit (1) ils signalent la sortie d'un couloir, soit (2) ils indiquent si l'utilisateur se déplace dans la mauvaise direction. Aucun stimulus n'est utilisé lorsque, (1) la position ou (2) la trajectoire satisfont les critères de précision fixés.

Le premier type d'approche peut être implémenté de façon directe avec notre dispositif, soit dans un but d'apprentissage, soit pour le guidage dans les lignes droites. En effet, plutôt que d'orienter continuellement un utilisateur, l'information directionnelle ne pourrait être proposée que lors de la sortie d'un couloir virtuel. Les tests préliminaires ont rapidement montré qu'un fonctionnement en mode binaire (son présent/absent) était peu confortable. Nous avons donc mis en place un fonctionnement plus doux où le niveau sonore augmente de façon assez rapide et non-linéaire lorsque l'utilisateur sort du couloir de guidage. Ainsi le son est correctement perçu et efficacement localisé sans pour autant ressentir d'effet de seuil désagréable et déroutant. Une des expériences réalisées est visible sur la Figure 10.18.

Nous pensons que ce type d'approche n'est pas forcément pertinent en guidage (bien qu'il semble utilisé par certaines applications) car il entraîne un effet de va et vient à l'intérieur du couloir peu efficace et perturbant pour un utilisateur peu entraîné. Cependant cette méthode, associée à notre dispositif pourrait sans doute être mise à profit pour l'apprentissage de la marche en ligne droite lors des exercices de locomotion. Cette approche reste à valider par des tests avec des instructeurs de locomotion.

### Guidage de personnes autistes

Le guidage pourrait être utile dans le cadre du *sport adapté* (c'est-à-dire les activités sportives destinées aux personnes présentant un handicap psychique ou mental). En effet, les personnes présentant des formes sévères de Troubles Envahissant du Développement (TED) et pratiquant la course à pied peuvent avoir de grandes difficultés pour identifier un parcours et le suivre. Elles doivent alors être guidées comme le serait une personne déficiente visuelle. Dans un rapport de 2010, la Haute Autorité de Santé (HAS) [100] recommandait spécifiquement la pratique de l'activité physique chez les personnes autistes dans la mesure où elle peut constituer un moyen de développer à la fois les capacités sensori-motrices et les capacités sociales. Très spécifiquement, les exercices aérobies intenses (comme la course à pied) seraient susceptibles de diminuer les stéréotypies [138, 183]. Plusieurs études révèlent également une amélioration des comportements sociaux [172, 204]. Un test préliminaire avec les athlètes du club *Savigny Athlétisme 91* laisse penser que les stimuli binauraux, qui ont été bien acceptés, pourraient constituer un moyen de guidage adapté. Avec des difficultés, lors d'une seule séance, un athlète a été capable d'identifier la direction de la source sonore et a parfois pu s'orienter dans la bonne direction. Pour autant, les difficultés de l'athlète ont rendu impossible un réel guidage. Les tests doivent être poursuivis, l'apprentissage étant assurément long avec des personnes présentant ce type de troubles.

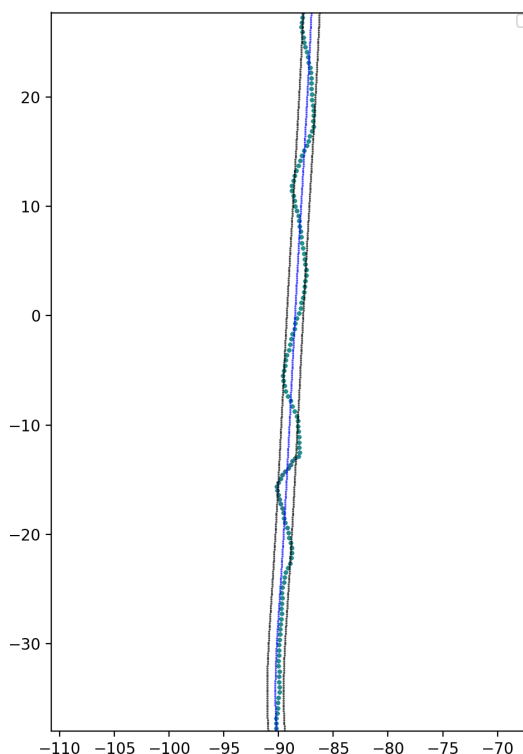


FIGURE 10.18 – Trajectoire d'une personne (voyante aveuglée), non entraînée, guidée dans une ligne droite avec déclenchement du guidage sonore à chaque sortie d'un couloir virtuel.

## Echolocation et guidage

Les personnes DV pratiquant une activité sportive sur un parcours qu'elles connaissent bien utilisent tous les indices sonores disponibles pour estimer leur position. En effet, savoir se localiser sur un parcours connu, même en étant guidé, permet une meilleure anticipation des virages et constitue un élément rassurant. Plusieurs des personnes que nous avons interrogées expliquent qu'elles cherchent à identifier la réverbération caractéristique de certaines zones de leurs parcours : par exemple pour identifier les gradins dans un gymnase. Comme dans un environnement naturel, une réalité sonore augmentée pourrait permettre d'utiliser des indices liés à la réverbération ou d'autres liés à des sources sonores fixes dans l'environnement. Par ailleurs, le sens des masses perçu par certaines personnes DV est également lié à la réverbération et il permet entre autres de repérer des obstacles massifs comme les murs. Pour mieux appréhender cette perception l'équipe audio du CMAP à d'ailleurs suivi une formation à la pratique de l'écholocation avec un instructeur spécialisé.

D'un point de vue acoustique, lorsqu'un son direct est réfléchi par un obstacle proche, le son perçu correspond à la combinaison du son direct et du son réfléchi. Il en résulte à la fois une modification du volume sonore perçu dans la direction de l'objet réfléchi ainsi qu'une coloration du son perçu. Ce phénomène de coloration du son est couramment appelé *repetition pitch* [30].

On peut alors se demander si cette perception des échos pourrait être utilisée à des fins de guidage. Pour cela nous avons mené quelques expériences exploratoires en créant des murs virtuels que les personnes pourront longer ou éviter. Ces expériences sont décrites en annexe D. Elles semblent montrer que les indices d'écholocation pourraient

constituer un complément aux techniques de guidage adaptatif décrites dans ce document en constituant une forme de réalité augmentée où le sujet utiliserait sa capacité à identifier des obstacles tels que des murs pour se diriger.

## Sécurité et gestion des incidents de parcours

Dans tous nos travaux, nous avons supposé que les personnes évoluaient dans un environnement totalement sûr et maîtrisé. L'analyse de l'environnement pour la gestion des obstacles et des différents aléas constitue un sujet en soi qui n'a pas été traité ici. Ainsi, en l'état actuel du prototype, une personne valide (qui peut être un accompagnateur extra-sportif) doit être présente pour assurer la sécurité, même si elle n'effectue pas de guidage. Le dispositif tel qu'il est présenté répond donc seulement à une partie des besoins exprimés par les clubs : il peut permettre à plusieurs personnes DV de courir avec un même guide/accompagnateur et le rôle de l'accompagnateur est par ailleurs largement facilité sur le plan sportif. Pour autant, il s'agit d'un frein important à l'utilisation des technologies développées dans ce document. Pour aller plus loin, on peut considérer deux approches :

1. *Permettre la détection d'obstacles sur la trajectoire* : Si on considère que la personne évolue dans un environnement parfaitement cartographié (par exemple une piste d'athlétisme) une pratique totalement autonome serait envisageable en couplant le dispositif à un système de détection d'obstacle. Ce système permettrait à la fois d'éviter d'éventuels objets ou personnes mais constituerait également une sécurité en cas de défaut du système de localisation.
2. *Identifier les éléments d'un environnement mal connu* : Une approche plus fine consiste à identifier les différents éléments de l'environnement, typiquement par un système de vision par ordinateur. Cette approche pourrait permettre de distinguer les obstacles et les coureurs : une demande courante lors de la pratique collective et en compétition.

La première approche pourrait être traitée avec des capteurs à faisceaux larges. Les plus simples et courants étant les capteurs ultra-son qui peuvent être disposés pour couvrir une surface plus ou moins importante. Ces capteurs ont pour inconvénient d'être relativement lents et sensibles aux conditions environnementales. De façon plus originale, on pourrait envisager d'utiliser des capteurs radar Doppler (tel que présenté en section 1) pour cette tâche.

## Navigation multi-échelle, atteinte de cible, saisie d'objet

Le guidage binaural peut être considéré à plusieurs échelles, à la fois pour le guidage sur un parcours (longue distance) mais aussi pour pointer et finalement atteindre un objet (courte distance). La visée des cibles et l'atteinte des objets peuvent nécessiter des corrections de mouvement plus précises et rapides qui n'ont pas été étudiées dans ce travail et des schémas de codage de substitution audio sont sans doute à explorer. Des applications sportives sont là encore possibles, en effet, le tir à l'aveugle est une discipline paralympique. Il est aujourd'hui pratiqué avec des dispositifs sonifiés qui n'utilisent pas de son 3D mais plutôt une approche chaud-froid.

## Approche multimodale

Enfin, nous nous sommes essentiellement intéressés au guidage par audio spatialisé, mais d'autres approches sont possibles comme la sononification ou la stimulation vibro-tactile. Des systèmes vibro-tactiles sont souvent présents dans les dispositifs destinés aux personnes DV (bracelets, vibreurs sur la canne) mais d'autres formats comme les

ceintures tactiles permettent de fournir des informations directionnelles. La multimodalité haptique et sonore constitue une piste de recherche intéressante : la redondance de l'information peut renforcer la robustesse matérielle et faciliter la perception et l'intégration de l'information.



## Annexe A

# Communications, publications

- An electronic travel aid device to help blind people playing sport, S Ferrand, F Alouges, M Aussal, IEEE Instrumentation Measurement Magazine 23 (4), 14-21, 2020.
- An Augmented Reality Audio Device Helping Blind People Navigation, S Ferrand, F Alouges, M Aussal, International Conference on Computers Helping People with Special Needs, 28-35, Linz 2018.
- Binaural Spatialization Methods for Indoor Navigation, S Ferrand, F Alouges, M Aussal, Audio Engineering Society Convention 142, Berlin 2017.
- A Real-Time Indoor Localization Method with Low-Cost Microwave Doppler Radar Sensors and Particle Filter, S Ferrand, F Alouges, M Aussal International Conference on Computers Helping People with Special Needs,(Lecture Notes in Computer Science volume 12376), Springer, 2020.
- HRTF and panning evaluations for binaural audio guidance, S Ferrand, F Alouges, M Aussal 23rd International Congress on Acoustics, Aachen, 2019.
- Guidage de personnes déficientes visuelles par audio spatialisé, S Ferrand, M Aussal, F Alouges 14ème Congrès Français d'acoustique, Le Havre 2018.



## Annexe B

# Filtrage, méthodes d'estimation d'états, fusion de données

### B.1 Filtrage bayésien

Dans le filtrage bayésien, on cherche à estimer l'état d'un système à partir d'une série d'observations. Dans les modèles les plus simples, le système est supposé Markovien, et son état  $\mathbf{x}_k$  et l'observation  $\mathbf{z}_k$  à l'instant  $k$  évoluent récursivement selon les équations :

$$\begin{aligned}\mathbf{x}_k &= \mathbf{f}_k(\mathbf{x}_{k-1}, \mathbf{u}_k) \\ \mathbf{z}_k &= \mathbf{h}_k(\mathbf{x}_k, \mathbf{v}_k).\end{aligned}\tag{B.1}$$

(B.2) La première équation s'appelle habituellement l'équation d'état, et  $\mathbf{f}_k$  est la fonction de transition entre les temps  $k-1$  et  $k$ , tandis que la seconde est l'équation d'observation, donnant l'évolution de la mesure à partir de l'état et d'une fonction d'observation  $\mathbf{h}_k$ . Les deux paramètres  $\mathbf{u}_k$  et  $\mathbf{v}_k$  sont des aléas sur le modèle et sur la mesure supposés indépendants.

Si l'on note  $\mathbf{z}_{1:k} = \{\mathbf{z}_1, \dots, \mathbf{z}_k\}$ , l'objectif de la méthode est de calculer récursivement la distribution a posteriori  $p(\mathbf{x}_k | \mathbf{z}_{1:k})$  permettant d'estimer l'état du système  $\mathbf{x}_k$  à partir des observations  $\mathbf{z}_{1:k}$ .

On pose habituellement comme hypothèse que les observations  $\mathbf{z}_k$  sont indépendantes conditionnellement au processus d'état, que leur distribution ne dépend que de l'état au même instant et que le processus  $x_k$  est Markovien. L'état initial  $x_0$  présente une distribution de probabilité supposée connue.

En connaissant  $p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1})$ , le caractère markovien du processus permet d'écrire (étape dite *de prédiction*) :

$$\begin{aligned}p(\mathbf{x}_k | \mathbf{z}_{1:k-1}) &= \int p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{z}_{1:k-1}) p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}) d\mathbf{x}_{k-1} \\ &= \int p(\mathbf{x}_k | \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}) d\mathbf{x}_{k-1}.\end{aligned}\tag{B.3}$$



L'application du théorème de Bayes et de la loi de la probabilité totale conduit à l'expression :

$$p(\mathbf{x}_k | \mathbf{z}_{1:k}) = \frac{p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{z}_{1:k-1})p(\mathbf{x}_k | \mathbf{z}_{1:k-1})}{p(\mathbf{z}_k | \mathbf{z}_{1:k-1})} = C p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{z}_{1:k-1})p(\mathbf{x}_k | \mathbf{z}_{1:k-1}), \quad (\text{B.4})$$

dans laquelle  $C$  est une constante de normalisation indépendante de  $\mathbf{x}_k$  donnée par

$$C = \frac{1}{p(\mathbf{z}_k | \mathbf{z}_{1:k-1})}.$$

Ce cadre général permet, en principe, le calcul récursif de la distribution a posteriori. Connaissant  $p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1})$ , on peut calculer, grâce à (B.1) et (B.3) la densité  $p(\mathbf{x}_k | \mathbf{z}_{1:k-1})$ . En reportant dans (B.4) et en utilisant (B.2) pour estimer  $p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1})$ , on obtient finalement  $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ , la constante  $C$  étant ajustée de façon à obtenir une densité de probabilités.

## B.2 Filtre de Kalman

Le calcul direct des expressions précédentes n'est pas possible en général, mais peut être effectué dans plusieurs cas particuliers. Par exemple, dans le cas de modèles linéaires et de bruits normalement distribués, le filtre de Kalman [120–122] permet ce calcul et fournit l'estimateur optimal pour des modèles linéaires.

Plus précisément, on se place dans le cas Le filtre de Kalman s'applique dans le cas :

— d'un système dynamique linéaire bruité qui peut s'écrire sous la forme :

$$\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{w}_k \quad (\text{B.5})$$

où l'évolution de l'état du système est donnée par la matrice de transition  $\mathbf{F}_k$  et le bruit additif  $w_k$  est supposé gaussien et de moyenne nulle.

— Les mesures sont également liées linéairement à l'état du système, et bruitées de façon additive et gaussienne :

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k. \quad (\text{B.6})$$

On note  $\hat{\mathbf{x}}_{m|n}$  l'estimation de l'état du système à l'instant  $m$  (c'est-à-dire  $\mathbf{x}_m$ ), obtenue à partir des observations  $\mathbf{z}_{1:n}$ , avec  $n \leq m$ . Comme précédemment, le filtre de Kalman est un estimateur qui fonctionne en deux phases, une étape de **prédiction** qui permet de proposer une estimation  $\hat{\mathbf{x}}_{k|k-1}$ , et une étape de **mise à jour** qui utilise l'observation  $\mathbf{z}_k$  pour proposer l'estimation  $\hat{\mathbf{x}}_{k|k}$ .

En notant  $\mathbf{P}_{m|n}$  la matrice de covariance de l'erreur  $\hat{\mathbf{x}}_{m|n} - \mathbf{x}_m$ , et  $\mathbf{Q}_k$  la matrice de covariance du bruit  $\mathbf{w}_k$  du processus, on peut écrire la phase de prédiction sous la forme la plus naturelle :

$$\begin{aligned} \hat{\mathbf{x}}_{k|k-1} &= \mathbf{F}_k \hat{\mathbf{x}}_{k-1|k-1} && \text{(Prédiction de l'état)} \\ \mathbf{P}_{k|k-1} &= \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^T + \mathbf{Q}_k && \text{(Estimation de la matrice de covariance)} \end{aligned}$$

L'étape de mise à jour prend en compte la mesure  $\mathbf{z}_k$  pour corriger l'état prédit. Si on note  $\mathbf{R}_k$  la matrice de covariance du bruit de mesure, on peut écrire cette étape sous la forme :

$$\begin{aligned}
\tilde{\mathbf{y}}_k &= \mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1} && \text{(innovation)} \\
\mathbf{S}_k &= \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k && \text{(covariance de l'innovation)} \\
\mathbf{K}_k &= \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{S}_k^{-1} && \text{(gain de Kalman "optimal")} \\
\hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \tilde{\mathbf{y}}_k && \text{(état mis à jour)} \\
\mathbf{P}_{k|k} &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} && \text{(covariance mise à jour)}.
\end{aligned}$$

Le filtre de Kalman fait l'hypothèse de bruits additifs et gaussiens dont les composantes sont totalement décorré- lées. Il est par ailleurs limité aux systèmes linéaires. Si ces conditions sont satisfaites, le filtre de Kalman est le filtre optimal pour l'estimation de l'état du système. Il garantit que l'erreur sur l'estimation est en moyenne nulle et minimise l'incertitude sur cette erreur.

Différentes variantes non-optimales ont été développées pour étendre la technique aux systèmes non linéaires. En particulier le filtre de Kalman étendu (EKF pour Extended Kalman Filter) utilise des développements de Taylor du premier ordre pour linéariser les fonctions de transition d'état et d'observation. L'hypothèse de bruit gaussien est en revanche conservée. L'EKF est très largement utilisé et est souvent considéré comme un standard de fait en navigation.

Une autre approche consiste à utiliser des techniques d'échantillonnage déterministe. Dans l'Unscented Kalman Filter (UKF) [115, 117, 222] la variable aléatoire gaussienne est approchée par un nombre limité de points d'échantillonnage soigneusement choisis pour capturer convenablement la moyenne et la covariance. Ces échantillons sont alors propa- gés par le véritable système non linéaire. Julier et al [116] ont montré que ce filtre était généralement plus robuste que les EKF pour une complexité du même ordre et que l'estimation de la variance et de la moyenne était correcte jusqu'au second ordre. Par ailleurs, par rapport à l'EKF, l'UKF présente l'avantage pratique de ne pas nécessiter de calcul de matrice jacobienne.

## B.3 Filtres particuliers (PF)

Quand la dynamique du système est fortement non linéaire ou que le bruit est non gaussien, les techniques dérivées du filtre de Kalman échouent à obtenir une estimation précise de l'état. La technique du filtrage particulière peut alors permettre d'obtenir de bons résultats. Le filtre particulière (parfois appelé Méthode de Monte-Carlo séquentielle) est un algorithme itératif permettant d'estimer l'état d'un système dynamique bruité, par des méthodes de type Monte-Carlo.

### B.3.1 Introduction aux méthodes de Monte-Carlo

Les techniques d'intégration de Monte Carlo (MC) offrent une évaluation approximative des expressions présentées dans la section B.1, en partant du principe qu'une espérance peut être approchée par une moyenne si la distribution sous-jacente peut être échantillonnée et que ces échantillons sont indépendants et de même distribution. Les premières méthodes de Monte-Carlo sont généralement attribuées à Metropolis et Ulam à la fin des années 40 [165]. Les méthodes séquentielles ont ensuite vu le jour dans les années 70 [98] avant de connaître un développement impor- tant dans la communauté "tracking" dans les années 90 avec le développement des *bootstrap filters* [91] et des *filtres particuliers* [61].

Soit une variable aléatoire  $X$  dans  $\mathbb{R}$  de densité de probabilité  $p(x)$ . Pour toute fonction  $\phi$  bornée de  $\mathbb{R}$  dans  $\mathbb{R}$ , l'espérance de  $\phi(X)$  peut s'écrire :

$$\mathbb{E}(\phi(X)) = \int \phi(x)p(x)dx = \bar{\phi}. \quad (\text{B.7})$$

Soit maintenant  $X^i, i = 1, 2, \dots, N$  des variables aléatoires indépendantes de même loi que  $X$ , la loi des grands nombres assure que la moyenne empirique tend presque sûrement vers l'espérance :

$$\bar{\phi}^N = \frac{1}{N} \sum_{i=1}^N \phi(X^i) \xrightarrow{N \rightarrow +\infty} \bar{\phi} \text{ p.s.} . \quad (\text{B.8})$$

La variance de  $\bar{\phi}^N$ , qui mesure la vitesse de convergence de la moyenne empirique précédente vers la moyenne peut s'écrire :

$$\text{var}(\bar{\phi}^N) = \frac{\sigma^2}{N} \quad (\text{B.9})$$

où

$$\sigma^2 = \int (\phi(x) - \bar{\phi})^2 p(x) dx. \quad (\text{B.10})$$

Dans l'approximation de Monte-Carlo l'erreur a l'énorme avantage de ne dépendre ni de la dimension de l'état, ni de la régularité de la fonction  $\phi$ .

### B.3.2 Échantillonnage préférentiel

Dans la pratique, il peut être difficile d'échantillonner directement depuis une fonction de densité arbitraire, ici  $p$ . Plusieurs techniques d'échantillonnages ont été proposées [193] pour résoudre cette question comme la technique d'échantillonnage préférentiel (*Importance sampling*) qui consiste à calculer l'espérance en échantillonnant suivant une autre loi  $q$ , dite *loi d'importance*, et en pondérant cet échantillon de manière adéquate. Le support de la distribution  $q$  doit alors contenir le support de  $p$  On peut alors écrire l'espérance de la façon suivante :

$$\mathbb{E}(\phi(X)) = \int \phi(x)p(x)dx = \int \phi(x) \frac{p(x)}{q(x)} q(x) dx = \mathbb{E}_q \left( \phi(Y) \frac{p(Y)}{q(Y)} \right) \quad (\text{B.11})$$

où  $Y$  est une variable aléatoire de densité  $q$ .

Considérant maintenant  $Y^i, i = 1, 2, \dots, N$  un échantillon de variables aléatoires indépendantes et de même loi  $q$  que  $Y$ , on aura alors

$$\mathbb{E}(\phi(X)) \approx \frac{1}{N} \sum_{i=1}^N \phi(Y^i) \frac{p(Y^i)}{q(Y^i)}. \quad (\text{B.12})$$

Lorsque la densité  $p$  n'est connue qu'à un facteur de normalisation près, on peut définir les *ponds d'importance*  $w^i$  sous forme normalisée

$$w^i = \alpha \frac{p(Y^i)}{q(Y^i)} \text{ avec } \alpha = \frac{1}{\sum_{i=1}^N \frac{p(Y^i)}{q(Y^i)}}, \quad (\text{B.13})$$

puis écrire

$$\mathbb{E}(\phi(X)) \approx \sum_{i=1}^N w^i \phi(Y^i). \quad (\text{B.14})$$

### B.3.3 Filtrage particulaire

Utilisant les techniques présentées précédemment, le filtre particulaire cherche à modéliser l'état d'un système par un ensemble de particules qui vont permettre d'explorer l'espace indépendamment selon le modèle dynamique. Un mécanisme de pondération et de sélection est implémenté pour évaluer l'adéquation de chaque particule en fonction de la mesure.

L'algorithme SIS (*Sequential Importance Sampling*) à la base du filtre particulaire consiste à utiliser l'échantillonnage préférentiel précédent et propager récursivement les échantillons (particules) en tenant compte de la dynamique du système  $p(x_k|x_{k-1})$  donné par (B.1) et de la vraisemblance  $p(y_k|x_k)$  donné par (B.2).

Considérons le modèle suivant :

$$\begin{aligned} \mathbf{x}_k &\sim p(\mathbf{x}_k|\mathbf{x}_{k-1}) \\ \mathbf{z}_k &\sim p(\mathbf{z}_k|\mathbf{x}_k) \end{aligned} \quad (\text{B.15})$$

On considère  $\mathbf{x}_k^i$  pour  $i = 1, \dots, N$  un échantillon de  $N$  variables (les particules), prises à l'instant  $k$  et l'on souhaiterait que

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) \sim \sum_{i=1}^N w_k^i \delta(\mathbf{x}_k - \mathbf{x}_k^i). \quad (\text{B.16})$$

Il s'agit de donner l'évolution des particules  $(\mathbf{x}_k^i)_i$  et des poids  $(w_k^i)_i$  en fonction de  $k$ .

Or, le théorème de Bayes, l'hypothèse de Markov et les calculs donnés en section B.1, montrent que la loi  $p(\mathbf{x}_{1:k}|\mathbf{z}_{1:k})$  peut se calculer récursivement par

$$\begin{aligned} p(\mathbf{x}_{1:k}|\mathbf{z}_{1:k}) &\propto p(\mathbf{z}_k|\mathbf{x}_{1:k}, \mathbf{z}_{1:k-1})p(\mathbf{x}_{1:k}|\mathbf{z}_{1:k-1}) \\ &= p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{x}_{1:k-1}, \mathbf{z}_{1:k-1})p(\mathbf{x}_{1:k}|\mathbf{z}_{1:k-1}) \\ &= p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{x}_{1:k}|\mathbf{z}_{1:k-1}). \end{aligned}$$

En supposant que les variables  $(\mathbf{x}_{1:k}^i)_i$  suivent une loi d'importance  $q(\mathbf{x}_{1:k}|\mathbf{z}_{1:k})$  connue, on aura le résultat souhaité, à condition que les poids  $(w_k^i)_i$  vérifient

$$\begin{aligned} w_k^i &\propto \frac{p(\mathbf{x}_{1:k}^i|\mathbf{z}_{1:k})}{q(\mathbf{x}_{1:k}^i|\mathbf{z}_{1:k})} \\ &\propto \frac{p(\mathbf{z}_k|\mathbf{x}_k^i)p(\mathbf{x}_k^i|\mathbf{x}_{k-1}^i)p(\mathbf{x}_{1:k}^i|\mathbf{z}_{1:k-1})}{q(\mathbf{x}_{1:k}^i|\mathbf{z}_{1:k})}. \end{aligned}$$

En formant récursivement la loi d'importance par

$$q(\mathbf{x}_{1:k}|\mathbf{z}_{1:k}) = q(\mathbf{x}_k|\mathbf{x}_{1:k-1}, \mathbf{z}_{1:k})q(\mathbf{x}_{1:k-1}|\mathbf{z}_{1:k-1}), \quad (\text{B.17})$$

on peut également calculer les poids récursivement grâce à

$$\begin{aligned}
 w_k^i &\propto \frac{p(\mathbf{z}_k | \mathbf{x}_k^i) p(\mathbf{x}_k^i | \mathbf{x}_{k-1}^i) p(\mathbf{x}_{1:k}^i | \mathbf{z}_{1:k-1})}{q(\mathbf{x}_{1:k}^i | \mathbf{z}_{1:k})} \\
 &= \frac{p(\mathbf{z}_k | \mathbf{x}_k^i) p(\mathbf{x}_k^i | \mathbf{x}_{k-1}^i) p(\mathbf{x}_{1:k}^i | \mathbf{z}_{1:k-1})}{q(\mathbf{x}_k^i | \mathbf{x}_{1:k-1}^i, \mathbf{z}_{1:k}) q(\mathbf{x}_{1:k-1}^i | \mathbf{z}_{1:k-1})} \\
 &\propto \frac{p(\mathbf{z}_k | \mathbf{x}_k^i) p(\mathbf{x}_k^i | \mathbf{x}_{k-1}^i)}{q(\mathbf{x}_k^i | \mathbf{x}_{1:k-1}^i, \mathbf{z}_{1:k})} w_{k-1}^i.
 \end{aligned} \tag{B.18}$$

La fonction d'importance  $q$  doit être choisie pour minimiser la variance des poids, mais pour des raisons pratiques, il peut être commode d'utiliser directement la dynamique du système lorsqu'elle est connue [91], de sorte que

$$q(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{z}_k) = p(\mathbf{x}_k | \mathbf{x}_{k-1}). \tag{B.19}$$

En substituant cette expression dans B.18, on peut alors exprimer le poids de façon simplifiée :

$$w_k^i \propto w_{k-1}^i p(\mathbf{z}_k | \mathbf{x}_k^i) \tag{B.20}$$

L'algorithme SIS présente l'inconvénient d'être sensible au problème de dégénérescence des poids. En effet, la variance des poids a tendance à augmenter à chaque itération [65]. Un poids va alors converger vers 1 et les autres vers zéro. Une solution consiste à utiliser une étape de rééchantillonnage, c'est le principe de l'algorithme *Sequential Importance Resampling* (SIR). Le rééchantillonnage augmentant la variance, il n'est généralement pas pratiqué à chaque étape, on va plutôt estimer un nombre de particules effectives  $N_{\text{eff}}$  et on rééchantillonnera quand il passera sous un certain seuil :

$$N_{\text{eff}} \approx \frac{1}{\sum_{i=1}^N (w_k^i)^2}. \tag{B.21}$$

L'estimation peut alors être obtenue en calculant la moyenne a posteriori du vecteur d'état pondéré par les poids.

L'algorithme SIR se décompose ainsi, en l'enchaînement d'une étape de prédiction, de mise à jour et de rééchantillonnage :

1. **Initialisation** : Création aléatoire d'un jeu de particule (représentant l'état)

$$\mathbf{x}_0^i \sim p(\mathbf{x}_0) \text{ et } w_0^i = 1/N.$$

2. **Prédiction** de l'état des particules depuis la distribution d'importance

$$\mathbf{x}_k^i \sim q(\mathbf{x}_k | \mathbf{x}_{0:k-1}, \mathbf{z}_{1:k}).$$

3. **Mise à jour** du poids des particules suivant eq. (B.18) et renormalisation.

4. Calcul de particules effectives  $N_{\text{eff}}$  d'après (B.21).

Le cas échéant : **Rééchantillonnage**.

5. Calcul de l'estimation (moyenne pondérée).

## Annexe C

# Exemple de fusion de donnée flot optique et UWB

Il est possible de coupler le système flot optique + AHRS avec de la localisation par proximité ou avec une technologie de mesure de distance. À titre d'exemple nous avons effectué une hybridation avec la mesure de distance d'une balise UWB unique. La fonction de mesure du filtre de Kalman prend alors la forme :

$$h(x_k) = \begin{pmatrix} \sqrt{\dot{x}^2/\dot{y}^2} \\ \arctan \frac{\dot{y}}{\dot{x}} \\ \sqrt{(x - a_{x,1})^2 + (y - a_{y,1})^2} \end{pmatrix}.$$

Dans les expériences de la figure C.1, un utilisateur se déplace dans un espace de  $6.5\text{m} \times 12.5\text{m}$  en suivant un parcours rectangulaire reproduit plusieurs fois. Les expériences sont ici réalisées dans un cas défavorable avec un sol peu texturé.

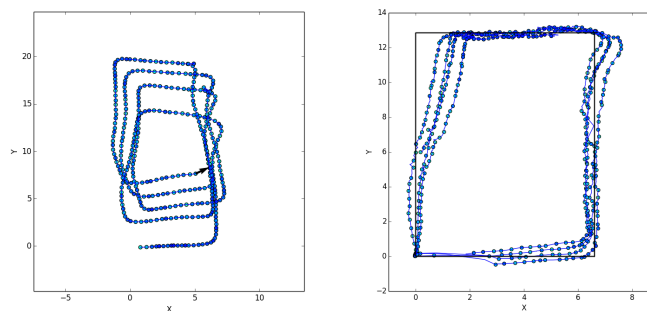


FIGURE C.1 – Positions estimées d'une personne se déplaçant sur un parcours rectangulaire (noir, à droite). À gauche, la position est estimée avec l'IMU et le capteur de flot. À droite, on ajoute une borne UWB positionnée à l'origine du repère.

Alors que sans balise UWB les erreurs d'orientation et de distance s'accroissent rapidement, avec une référence de distance, on peut limiter ces décalages pour les parcours comportant des boucles. Dans cette expérience l'erreur RMS reste inférieure à 1m au cours du temps.



## Annexe D

# Echolocation et guidage : expériences exploratoires

Afin de savoir si la perception des échos d'obstacles virtuels pourraient être utilisé à des fins de guidage, nous avons mené les expériences exploratoires décrites ci-dessous.

Dans un premier temps, il a été demandé à trois personnes, non voyantes et maîtrisant l'écholocation, d'écouter des sons dans un environnement virtuel simulant des pièces fermées ou des espaces semi-ouverts avec 1, 2 ou 3 murs. Pour cela nous avons utilisé les méthodes d'auralisation décrites en partie deux. Il est apparu que ces personnes étaient capables d'identifier la présence et la direction approximative d'un ou deux murs mais éprouvaient beaucoup de difficultés dans les environnements comportant plus de deux murs. Elles rencontraient en particulier beaucoup de confusions avant/arrière (peut être liées à l'utilisation de HRTF non-individualisées). Nous avons alors invité une personne aveugle à participer à deux expériences en situation de guidage dans un environnement ne comportant qu'un seul mur.

Pour permettre l'écholocation, nous plaçons une source sonore omnidirectionnelle devant la personne. Lorsque le sujet se déplace, la source se déplace aussi de façon à rester face à elle à environ un mètre. On cherche ici à reproduire une pratique d'écholocation utilisant le claquement des mains. Le stimulus est choisi en conséquence, c'est un claquement bref répétitif. Le sol est toujours présent mais peu réfléchissant (absorption 0.8). Les murs virtuels ont un coefficient de réflexion de 1 (sur tout le spectre) pour être très perceptibles dans ces expériences exploratoires.

**Se diriger vers un mur virtuel :** Équipé du dispositif, on explique à la personne qu'un mur est présent à quelque mètres d'elle, sans lui indiquer la distance ni la direction. Elle a alors pour instruction de se diriger droit vers le mur. Les images de la figure D.1 montrent deux exemples de parcours. Dans chaque cas, la personne est capable de se diriger vers le mur, presque perpendiculairement et d'une façon comparable à une situation réelle.

**Longer des murs et guider une personne** Dans une seconde expérience, on demande à la personne de longer les murs et de tourner à gauche quand elle rencontre un mur devant elle. Pour rendre la perception plus simple on fait un sorte qu'un seul mur soit audible à la fois. Pour cela on divise l'espace de la pièce virtuelle en quatre quadrants, et c'est donc le mur dans le quadrant où l'utilisateur se trouve qui est seulement audible.





FIGURE D.1 – On demande à un sujet de se diriger vers un mur virtuel de 10m de long et 3 m de haut créé par auralisation temps réel avec localisation du sujet et head-tracking. Les positions de départ sont matérialisées par une étoile et le parcours du sujet est en noir.

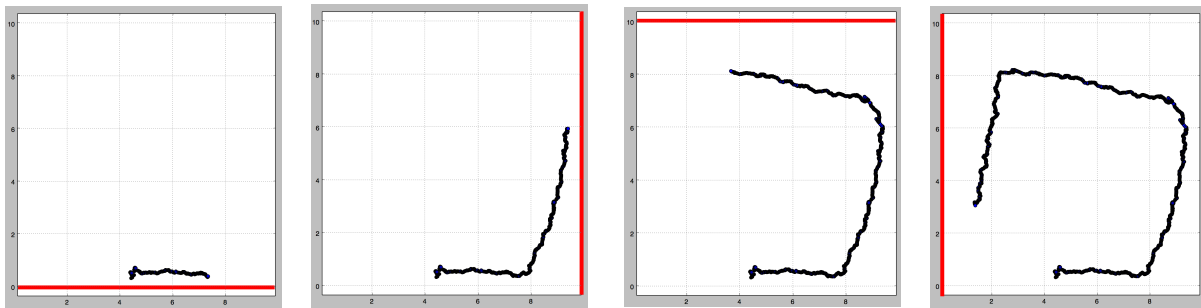


FIGURE D.2 – On demande à un sujet de suivre les murs et d'éviter les murs qui apparaissent devant lui (en tournant à gauche). Chaque image représente une étape de la progression du sujet.

La figure D.2 montre la progression de l'utilisateur, les murs audibles sont matérialisées par des lignes épaisses. La pièce virtuelle est de dimension 10mx10m. Nous avons été capables de diriger un utilisateur après un petit entraînement de quelques minutes, mais à très faible vitesse (max. 1.5km/h). Pour ce résultat, le stimulus est encore un claquement bref. D'autres utilisateurs non voyants ont préféré des bruits blancs hachés (100ms de bruit blanc, 200ms de silence) alors que les sons musicaux se sont révélés inefficaces.

Cette expérience constitue une première preuve de concept de guidage avec des murs virtuels qui pourrait être explorée plus avant pour être intégrée à un système de guidage. Ce type de guidage seul ne semble pas envisageable dans le contexte de la pratique sportive qui nous intéresse mais ces indices d'écholocation pourraient constituer un complément au guidage adaptatif.

# Liste des Acronymes

**AHRS** : Attitude and Heading Reference System

**API** : Application Programming Interface

**BRIR** : Binaural Room Impulse Response

**D-GNSS** : Differential Global Navigation Satellite System

**DV** : Personnes ayant une Déficience Visuelle

**EGNOS** : European Geostationary Navigation Overlay Service

**ETA** : Electronic Travel Aid

**FFT** : Fast Fourier Transform

**GBAS** : Ground-Based Augmentation System

**GDOP** : Geometric Dilution Of precision

**GNSS** : Géolocalisation et Navigation par un Systeme de Satellites/Global Navigation Satellite System

**GPIO** : General Purpose Input/Output

**GPS** : Global Positioning System

**HRIR** : Head-Related Impulse Response

**HRTF** : Head-Related Transfer Function

**ILD** : Interaural Level Difference

**IRBA** : Institut de Recherche Biomédicale des Armées

**IRCAM** : Institut de recherche et coordination acoustique/musique

**ITD** : Interaural Time Difference

**IMU** : Inertial Motion Unit

**LSE** : Least square Estimation

**MEMS** : Micro-Electro-Mechanical-Systems

**OMS** : Organisation mondiale de la santé

**OWR** : One Way Ranging

**PF** : Particles Filter

**PID** : Proportionnel, Intégral, Dérivé

**REST** : Representational State Transfer  
**RFID** : Radio-frequency identification  
**RIR** : Room Impulse Response  
**RMS** : Root Mean Square  
**RSS** : Received Signal Strength  
**RSS** : Received Signal Strength Indication  
**RTK** : Real Time Kinematic  
**SBAS** : Satellite-Based Augmentation Systems  
**SIR** : Sequential Importance Resampling  
**SIS** : Sequential Importance Sampling  
**SQUAL** : Surface Quality  
**SSD** : Sensory Substitution Devices  
**TDoA** : Time Difference of Arrival  
**TED** : Troubles Envahissant du Developpemen  
**TFD** : Transformé de Fourier Discrète  
**ToF** : Time of Flight  
**TWR** : Two Way Ranging  
**UERE** : User Equivalent Range Error  
**UKF** : Unscented Kalman Filter  
**UWB** : Ultra WideBand

# Bibliographie

- [1] Association Accessi Meca Sports. <https://www.lelouroux.com/asso/accessi-meca-sports/>.
- [2] Association Mixhandi-CapsurlaVie (MCV). <http://www.mixhandi-capsurlavie.fr/>.
- [3] Savigny athlétisme 91. <http://savigny-athlétisme91.athle.fr/>.
- [4] Spat, IRCAM forum online. <https://forum.ircam.fr/projects/detail/spat/>.
- [5] Wizigo - Application d'aide au déplacement des déficients visuels. <http://www.gosense.com/fr/wizigo/>.
- [6] ALGAZI, V., DUDA, R., THOMPSON, D., AND AVENDANO, C. The cipc hrtf database. In *Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics* (New Platz, NY, USA, 2001), pp. 99–102.
- [7] ALLEN, J. B., AND BERKLEY, D. A. Image method for efficiently simulating small-room acoustics. *The Journal of the Acoustical Society of America* 65, 4 (1979), 943–950.
- [8] ANDÉOL, G., SARAFIAN, D., PELLIEUX, L., ROUMES, C., AND GUILLAUME, A. Le son 3d comme interface homme/machine en aéronautique. *Acoustique et Techniques : trimestriel d'information des professionnels de l'acoustique* (2011).
- [9] ANDÒ, B. Electronic sensory systems for the visually impaired. *IEEE Instrumentation & Measurement Magazine* 6, 2 (2003), 62–67.
- [10] ARI. Base de données HRTFs ARI, Acoustics Research Institute. [https://www.kfs.oeaw.ac.at/index.php?option=com\\_content&view=article&id=608&catid=158&Itemid=606](https://www.kfs.oeaw.ac.at/index.php?option=com_content&view=article&id=608&catid=158&Itemid=606).
- [11] ARMELLONI, E., GIOTTOLI, C., AND FARINA, A. Implementation of real-time partitioned convolution on a dsp board. In *2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics* (Oct. 2003), pp. 71–74.
- [12] ATIENZA, C., ARCELO, R. F., BUENCONSEJO, G., VISITACION, A., PALCONIT, M. G., AND OBISO, J.-J. A wearable electronic guidance system for the blind person. *Journal of Science, Engineering and Technology* 6 (2018), 247–255.
- [13] AUSSAL, M. *Methodes Numeriques Pour La Spatialisation Sonore, de La Simulation a La Synthese Binaurale*. Phd thesis, Ecole Polytechnique X, 2014.
- [14] AUSSAL, M., AND ALOUGES, F. Gypsilab, un nouvel outil de calcul fem/bem pour l'acoustique numérique. In *Actes Du Congrès Français d'Acoustique, Le Havre* (2018).
- [15] AUSSAL, M., ALOUGES, F., AND KATZ, B. F. ltd interpolation and personalization for binaural synthesis using spherical harmonics. In *Audio Engineering Society UK Conference* (2012), p. 04.

- [16] BACH-Y-RITA, P., COLLINS, C. C., SAUNDERS, F. A., WHITE, B., AND SCADDEN, L. Vision substitution by tactile image projection. *Nature* 221, 5184 (1969), 963–964.
- [17] BACH-Y-RITA, P., AND KACZMAREK, K. A. Tongue placed tactile output device (brevet usa ref. us6430450b1), 2002.
- [18] BAHL, P., AND PADMANABHAN, V. Radar : An in-building rf-based user location and tracking system. In *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies* (Tel Aviv, Israel, 2000), vol. 2, pp. 775–784.
- [19] BARDE, A., LEE, G., HELTON, W. S., AND BILLINGHURST, M. Binaural spatialisation over a bone conduction headset : Elevation perception. In *The 22nd International Conference on Auditory Display* (2016).
- [20] BATTEAU, D. W. The role of the pinna in human localization. *Proceedings of the Royal Society of London. Series B. Biological Sciences* 168, 1011 (1967), 158–180.
- [21] BATTEAU, D. W., AND HUXLEY, H. E. The role of the pinna in human localization. *Proceedings of the Royal Society of London. Series B. Biological Sciences* 168, 1011 (1967), 158–180.
- [22] BATTENBERG, E., AND AVIZIENIS, R. Implementing real-time partitioned convolution algorithms on conventional operating systems. In *Proceedings of the 14th International Conference on Digital Audio Effects. Paris, France* (2011), pp. 248–235.
- [23] BEGAULT, D. R., WENZEL, E. M., AND ANDERSON, M. R. Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source. *Journal of the Audio Engineering Society* 49, 10 (2001), 904–916.
- [24] BENEDETTI, L. H., AND LOEB, M. A comparison of auditory monitoring performance in blind subjects with that of sighted subjects in light and dark. *Perception & Psychophysics* 11, 1 (1972), 10–16.
- [25] BENHAM, T., AND BENJAMIN, J. Active energy radiating systems : An electronic travel aid. In *Proceedings of the International Congress on Technology and Blindness* (1963), pp. 167–176.
- [26] BENKIC, K., MALAJNER, M., PLANINSIC, P., AND CUCEJ, Z. Using rssi value for distance estimation in wireless sensor networks based on zigbee. In *2008 15th International Conference on Systems, Signals and Image Processing* (2008), pp. 303–306.
- [27] BENKIC, K., MALAJNER, M., PLANINSIC, P., AND CUCEJ, Z. Using rssi value for distance estimation in wireless sensor networks based on zigbee. In *2008 15th International Conference on Systems, Signals and Image Processing* (2008), IEEE, pp. 303–306.
- [28] BEURLE, R. L. *Summary of Suggestions on Sensory Devices*, San Dunstan's, London. ed. 1947.
- [29] BÉZARD, P., AUSSAL, M., AND PARMENTIER, M. Using a binaural spatialization tool in current sound post-production and broadcast workflow. In *Audio Engineering Society Conference : 57th International Conference : The Future of Audio Entertainment Technology—Cinema, Television and the Internet* (2015), Audio Engineering Society.
- [30] BILSEN, F. A. Repetition pitch : Monaural interaction of a sound with the repetition of the same, but phase shifted, sound. *Acta Acustica united with Acustica* 17, 5 (1966), 295–300.
- [31] BLAUERT, J. *Spatial Hearing : The Psychophysics of Human Sound Source Localization*. Cambridge, MA : MIT Press, 1983.

- [32] BOMHARDT, R., DE LA FUENTE KLEIN, M., AND FELS, J. A high-resolution head-related transfer function and three-dimensional ear model database. In *Proceedings of Meetings on Acoustics 172ASA* (2016), vol. 29, ASA, p. 050002.
- [33] BORISH, J. Extension of the image model to arbitrary polyhedra. *The Journal of the Acoustical Society of America* 75, 6 (1984), 1827–1836.
- [34] BRANCH, M. A., COLEMAN, T. F., AND LI, Y. A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems. *SIAM Journal on Scientific Computing* 21, 1 (1999), 1–23.
- [35] BRINKMANN, F., LINDAU, A., WEINZIERL, S., GEISLER, G., AND VAN DE PAR, S. A high resolution head-related transfer function database including different orientations of head above the torso. In *Proceedings of the AIA-DAGA 2013 Conference on Acoustics* (2013).
- [36] BRUNATO, M., AND BATTITI, R. Statistical learning theory for location fingerprinting in wireless lans. *Computer Networks* 47, 6 (2005), 825–845.
- [37] BRUNGART, D. S., DURLACH, N. I., AND RABINOWITZ, W. M. Auditory localization of nearby sources. ii. localization of a broadband source. *The Journal of the Acoustical Society of America* 106, 4 (1999), 1956–1968.
- [38] BRUNGART, D. S., SIMPSON, B. D., AND KORDIK, A. J. The detectability of headtracker latency in virtual audio displays. In *Eleventh Meeting of the International Conference on Auditory Display* (2005), p. 7.
- [39] BRUNGART, D. S., SIMPSON, B. D., MCKINLEY, R. L., KORDIK, A. J., DALLMAN, R. C., AND OVENSHERE, D. A. The interaction between head-tracker latency, source duration, and response time in the localization of virtual sound sources. In *Proceedings of ICAD 04- Tenth Meeting of the International Conference on Auditory Display* (2004), Georgia Institute of Technology.
- [40] BUJACZ, M., KROPIDLOWSKI, K., IVANICA, G., MOLDOVEANU, A., SAITIS, C., CSAPÓ, Á., WERSÉNYI, G., SPAGNOL, S., JOHANNESSON, O. I., UNNTHORSSON, R., ROTNICKI, M., AND WITEK, P. Sound of vision - spatial audio output and sonification approaches. In *ICCHP* (2016).
- [41] BUJACZ, M., PEC, M., SKULIMOWSKI, P., STRUMILLO, P., AND MATERKA, A. Sonification of 3d scenes in an electronic travel aid for the blind. In *Advances in Sound Localization*. 2011, pp. 251–268.
- [42] BUJACZ, M., SKULIMOWSKI, P., AND STRUMILLO, P. Naviton—a prototype mobility aid for auditory presentation of three-dimensional scenes to the visually impaired. *Journal of the Audio Engineering Society* 60, 9 (2012), 696–708.
- [43] BULLINGER, H.-J., RGEN ZIEGLER, J., BULLINGER, H.-J., AND ZIEGLER, J. *Human-Computer Interaction : Communication, Cooperation, and Application Design*. CRC Press, 1999.
- [44] BURGESS, N., SPIERS, H. J., AND PALEOLOGOU, E. Orientational manoeuvres in the dark : Dissociating allocentric and egocentric influences on spatial memory. *Cognition* 94, 2 (2004), 149–166.
- [45] BUTLER, R. A., AND PLANERT, N. The influence of stimulus bandwidth on localization of sound in space. *Perception & Psychophysics* 19, 1 (1976), 103–108.
- [46] CAPP, M., AND PICTON, P. The optophone : An electronic blind aid. *Engineering Science and Education Journal* 9, 3 (2000), 137–143.
- [47] CARPENTIER, T., BAHU, H., NOISTERNIG, M., AND WARUSFEL, O. Measurement of a head-related transfer function database with high spatial resolution. In *7th Forum Acusticum (EAA)* (2014).

- [48] CARRASCO, E., LOYO, E., OTAEGUI, O., FÖSLEITNER, C., DUBIELZIG, M., OLMEDO, R., WASSERBURGER, W., AND SPILLER, J. Argus autonomous navigation system for people with visual impairments. In *International Conference on Computers for Handicapped Persons* (2014), Springer, pp. 100–107.
- [49] CHAN, R., MULLA, A., AND STOL, K. Characterisation of low-cost optical flow sensors. In *Proceeding of the Australasian Conference on Robotics and Automation* (2010), pp. 1–8.
- [50] CHANCE, S. S., GAUNET, F., BEALL, A. C., AND LOOMIS, J. M. Locomotion mode affects the updating of objects encountered during travel : The contribution of vestibular and proprioceptive inputs to path integration. *Presence* 7, 2 (1998), 168–178.
- [51] CHEN, Q., TAN, B., WOODBRIDGE, K., AND CHETTY, K. Indoor target tracking using high doppler resolution passive wi-fi radar. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (South Brisbane, Queensland, Australia, 2015), IEEE, pp. 5565–5569.
- [52] CHIANI, M., AND GIORGETTI, A. Coexistence between uwb and narrow-band wireless communication systems. *Proceedings of the IEEE* 97, 2 (2009), 231–254.
- [53] CHUMKAMON, S., TUVAPHANTHAPHIPHAT, P., AND KEERATIWINTAKORN, P. A blind navigation system using rfid for indoor environments. In *2008 5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology* (2008), vol. 2, pp. 765–768.
- [54] CNES. Guide egnos à l’usage des développeurs d’applications. Tech. rep., CNES, ESA, 2011.
- [55] COLEMAN, P. D. Failure to localize the source distance of an unfamiliar sound. *The Journal of the Acoustical Society of America* 34, 3 (1962), 345–346.
- [56] CUEVAS-RODRIGUEZ, M., GONZALEZ-TOLEDO, D., RUBIA-CUESTAS, E., GARRE, C., MOLINA-TANCO, L., REYES-LECUONA, A., POIRIER-QUINOT, D., AND PICINALI, L. An open-source audio renderer for 3d audio with hearing loss and hearing aid simulations. In *AES Convention 142*. (2017), p. 9.
- [57] DALLOS, P. Overview : Cochlear neurobiology. In *The Cochlea*, P. Dallos, A. N. Popper, and R. R. Fay, Eds., Springer Handbook of Auditory Research. Springer, New York, 1996, pp. 1–43.
- [58] DANIELS, J. *Daniels’ Running Formula, Second Edition*, the state university of new york at cortland ed. Human Kinetics, 2013.
- [59] DE PRETER, A., ANTHONIS, J., SWEVERS, J., AND PIPELEERS, G. Experiment design for ultra-wideband sensor node calibration. In *Proceedings of the 2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN)* (2018), IEEE.
- [60] DECAWAVE. Application note aps014 - antenna delay calibration of dw1000-based products and systems. Tech. rep., Decawave, 2014.
- [61] DEL MORAL, P. Non-linear filtering : Interacting particle resolution. *Markov processes and related fields* 2, 4 (1996), 555–581.
- [62] DENIS, M., Ed. *La psychologie cognitive*. Cogniprisme. Éditions de la Maison des sciences de l’homme, Paris, Nov. 2019.
- [63] DIAGNE, S., VAL, T., FAROTA, A. K., AND DIOP, B. Comparative analysis of ranging protocols for localization by uwb in outdoor. *Wireless Sensor Network* 10, 05 (2018), 103–117.

- [64] DILLE, M., GROCHOLSKY, B., AND SINGH, S. Outdoor downward-facing optical flow odometry with commodity sensors. In *Field and Service Robotics* (Berlin, Heidelberg, 2010), A. Howard, K. Iagnemma, and A. Kelly, Eds., Springer Tracts in Advanced Robotics, Springer, pp. 183–193.
- [65] DOUCET, A., GODSILL, S., AND ANDRIEU, C. On sequential monte carlo sampling methods for bayesian filtering. *Statistics and computing* 10, 3 (2000), 197–208.
- [66] DRAMAS, F. *Localisation d'objets Pour Les Non-Voyants : Augmentation Sensorielle et Neuroprothèse*. PhD thesis, Université de Toulouse, Université Toulouse III - Paul Sabatier, June 2010.
- [67] DUNAI, L. D., LENGUA, I. L., TORTAJADA, I., AND SIMON, F. B. Obstacle detectors for visually impaired people. In *2014 International Conference on Optimization of Electrical and Electronic Equipment (OPTIM)* (2014), IEEE, pp. 809–816.
- [68] EL ABOUSS, A., AND SAMANTA, N. A modeling of gps error distributions. In *2017 European Navigation Conference (ENC)* (2017), pp. 119–127.
- [69] EL-DIASTY, M., EL-RABBANY, A., AND PAGIATAKIS, S. Temperature variation effects on stochastic characteristics for low-cost mems-based inertial sensor error. *Measurement Science and Technology* 18, 11 (2007), 3321.
- [70] ELMEZAYEN, A., AND EL-RABBANY, A. Precise point positioning using world's first dual-frequency gps/galileo smartphone. *Sensors* 19, 11 (2019), 2593.
- [71] EUSTON, M., COOTE, P., MAHONY, R., KIM, J., AND HAMEL, T. A complementary filter for attitude estimation of a fixed-wing uav. In *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems* (2008), pp. 340–345.
- [72] FANG, S., AND LIN, T. Indoor location system based on discriminant-adaptive neural network in ieee 802.11 environments. *IEEE Transactions on Neural Networks* 19, 11 (2008), 1973–1978.
- [73] FANG, S.-H., WANG, C.-H., HUANG, T.-Y., YANG, C.-H., AND CHEN, Y.-S. An enhanced zigbee indoor positioning system with an ensemble approach. *IEEE Communications Letters* 16, 4 (2012), 564–567.
- [74] FARCY, R., LEROUX, R., JUCHA, A., DAMASCHINI, R., GRÉGOIRE, C., AND ZOGAGHI, A. Electronic travel aids and electronic orientation aids for blind people : Technical, rehabilitation and everyday life points of view. In *Conference & Workshop on Assistive Technologies for People with Vision & Hearing Impairments Technology for Inclusion* (2006), vol. 12.
- [75] FÉDÉRATION FRANCAISE HANDISPORT. Rapport sur les affiliations et licences saison sportive 2016-2017. Tech. rep., Fédération Francaise Handisport, 2018.
- [76] FELDMANN, S., KYAMAKYA, K., ZAPATER, A., AND LUE, Z. An indoor bluetooth-based positioning system : Concept, implementation and experimental evaluation. In *International Conference on Wireless Networks* (2003), p. 5.
- [77] FELIZ ALONSO, R., ZALAMA CASANOVA, E., AND GÓMEZ GARCÍA-BERMEJO, J. Pedestrian tracking using inertial sensors. *Journal of Physical Agent* (2009).
- [78] FERRAND, S., ALOUGES, F., AND AUSSAL, M. Binaural spatialization methods for indoor navigation. In *Audio Engineering Society Convention 142* (Berlin, 2017), p. 5.
- [79] FERRAND, S., ALOUGES, F., AND AUSSAL, M. An augmented reality audio device helping blind people navigation. In *Computers Helping People with Special Needs*, vol. 10897. Springer, 2018, pp. 28–35.



- [80] FERRAND, S., ALOUGES, F., AND AUSSAL, M. Hrtf and panning evaluations for binaural audio guidance. In *23rd International Congress on Acoustics* (Aachen, 2019), p. 7.
- [81] FERRAND, S., ALOUGES, F., AND AUSSAL, M. An electronic travel aid device to help blind people playing sport. *IEEE Instrumentation & Measurement Magazine* 23, 4 (2020), 14–21.
- [82] FERRAND, S., ALOUGES, F., AND AUSSAL, M. A real-time indoor localization method with low-cost microwave doppler radar sensors and particle filter. In *International Conference on Computers Helping People with Special Needs* (2020), Springer, pp. 467–474.
- [83] FIELDING, R. T., AND TAYLOR, R. N. *Architectural Styles and the Design of Network-Based Software Architectures*, vol. 7. University of California, Irvine Doctoral dissertation, 2000.
- [84] FOXLIN, E. Pedestrian tracking with shoe-mounted inertial sensors. *IEEE Computer graphics and applications* 25, 6 (2005), 38–46.
- [85] FREIBERGER, H. Mobility aids for the blind. *Bulletin of prosthetics research* (1974).
- [86] FRIIS, H. A note on a simple transmission formula. *Proceedings of the IRE* 34, 5 (1946), 254–256.
- [87] GAGEIK, N., STROHMEIER, M., AND MONTENEGRO, S. An autonomous uav with an optical flow sensor for positioning and navigation. *International Journal of Advanced Robotic Systems* 10, 10 (2013), 341.
- [88] GARDNER, M. B., AND GARDNER, R. S. Problem of localization in the median plane : Effect of pinnae cavity occlusion. *The Journal of the Acoustical Society of America* 53, 2 (1973), 400–408.
- [89] GARDNER, W. G. Efficient convolution without input/output delay. In *Audio Engineering Society Convention 97* (1994), Audio Engineering Society.
- [90] GEIER, M., AND SPORS, S. Spatial audio with the soundscape renderer. In *27th Tonmeistertagung* (2012), p. 10.
- [91] GORDON, N. J., SALMOND, D. J., AND SMITH, A. F. Novel approach to nonlinear/non-gaussian bayesian state estimation. In *IEE Proceedings F (Radar and Signal Processing)* (1993), vol. 140, IET, pp. 107–113.
- [92] GORI, M., CAPPAGLI, G., TONELLI, A., BAUD-BOVY, G, AND FINOCCHIETTI, S. Devices for visually impaired people : High technological devices with low user acceptance and no adaptability for children - sciencedirect. *Neuroscience & Biobehavioral Reviews* 69 (2016), 79–88.
- [93] GOZICK, B., SUBBU, K. P., DANTU, R., AND MAESHIRO, T. Magnetic maps for indoor navigation. *IEEE Transactions on Instrumentation and Measurement* 60, 12 (2011), 3883–3891.
- [94] GROVES, P. D. Principles of gnss, inertial, and multisensor integrated navigation systems, 2nd edition [book review]. *IEEE Aerospace and Electronic Systems Magazine* 30, 2 (2015), 26–27.
- [95] GUEGUEN, R. *Virtualisation Architecturale Visuelle et Auditive Du Théâtre Antique d'Orange*. These de doctorat, Sorbonne université, Dec. 2018.
- [96] GUTH, D. Why does training reduce blind pedestrians veering. *Blindness and brain plasticity in navigation and object perception* (2007), 353–365.
- [97] GUTH, D., AND LADUKE, R. The veering tendency of blind pedestrians : An analysis of the problem and literature review. *Journal of Visual Impairment & Blindness* (1994).
- [98] HANDSCHIN, J. E. Monte carlo techniques for prediction and filtering of non-linear stochastic processes. *Automatica* 6, 4 (1970), 555–563.

- [99] HARTMANN, W. M., AND RAKERD, B. On the minimum audible angle—a decision theory approach. *The Journal of the Acoustical Society of America* 85, 5 (1989), 2031–2041.
- [100] HAS. Autisme et autres troubles envahissants du développement état des connaissances hors mécanismes physiopathologiques, psychopathologiques et recherche fondamentale. Tech. rep., HAS, 2010.
- [101] HAVERINEN, J., AND KEMPPAINEN, A. Global indoor self-localization based on the ambient magnetic field. *Robotics and Autonomous Systems* 57, 10 (2009), 1028–1035.
- [102] HAZAS, M., AND HOPPER, A. Broadband ultrasonic location systems for improved indoor positioning. *IEEE Transactions on Mobile Computing* 5, 5 (2006), 536–547.
- [103] HENDRICKX, E., STITT, P., MESSONNIER, J.-C., LYZWA, J.-M., KATZ, B. F., AND DE BOISHÉRAUD, C. Influence of head tracking on the externalization of speech stimuli for non-individualized binaural synthesis. *The Journal of the Acoustical Society of America* 141, 3 (2017), 2011–2023.
- [104] HERTRICH, I., DIETRICH, S., MOOS, A., TROUVAIN, J., AND ACKERMANN, H. Enhanced speech perception capabilities in a blind listener are associated with activation of fusiform gyrus and primary visual cortex. *Neurocase* 15, 2 (2009), 163–170.
- [105] HOL, J. D., SCHÖN, T. B., AND GUSTAFSSON, F. Ultra-wideband calibration for indoor positioning. In *IEEE International Conference on Ultra-Wideband* (2010), vol. 2, pp. 1–4.
- [106] HUB, A., HARTTER, T., AND ERTL, T. Interactive tracking of movable objects for the blind on the basis of environment models and perception-oriented object recognition methods. In *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility* (2006), ACM, pp. 111–118.
- [107] IHLEFELD, A., AND SHINN-CUNNINGHAM, B. G. Effect of source spectrum on sound localization in an everyday reverberant room. *The Journal of the Acoustical Society of America* 130, 1 (2011), 324–333.
- [108] INNOSENT. Application note 3. IPM-165 a universal low cost k-band transceiver for motion detection in various applications.
- [109] INNOSENT. Datasheet ipm-165, low cost k-band transceiver, 2014.
- [110] ISLAM, T., ISLAM, M. S., SHAJID-UL-MAHMUD, M., AND HOSSAM-E-HAIDER, M. Comparison of complementary and kalman filter based data fusion for attitude heading reference system. *AIP Conference Proceedings* 1919, 1 (2017), 020002.
- [111] JANUSZEWSKI, J. Sources of error in satellite navigation positioning. *TransNav : International Journal on Marine Navigation and Safety of Sea Transportation* 11 (2017).
- [112] JOHNSON, D., DETERDING, S., KUHN, K.-A., STANEVA, A., STOYANOV, S., AND HIDES, L. Gamification for health and wellbeing : A systematic review of the literature. *Internet Interventions* 6 (2016), 89–106.
- [113] JOHNSON, L. A., AND HIGGINS, C. M. A navigation aid for the blind using tactile-visual sensory substitution. In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society* (2006), pp. 6289–6292.
- [114] JONES, M., JONES, S., BRADLEY, G., WARREN, N., BAINBRIDGE, D., AND HOLMES, G. Ontrack : Dynamically adapting music playback to support navigation. *Personal and Ubiquitous Computing* 12, 7 (2008), 513–525.
- [115] JULIER, S. J., AND UHLMANN, J. K. New extension of the kalman filter to nonlinear systems. In *Signal Processing, Sensor Fusion, and Target Recognition VI* (1997), vol. 3068, International Society for Optics and Photonics, pp. 182–193.

- [116] JULIER, S. J., AND UHLMANN, J. K. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE* 92, 3 (2004), 401–422.
- [117] JULIER, S. J., UHLMANN, J. K., AND DURRANT-WHYTE, H. F. A new approach for filtering nonlinear systems. In *Proceedings of 1995 American Control Conference-ACC'95* (1995), vol. 3, IEEE, pp. 1628–1632.
- [118] KACZMAREK, K. A. Sensory augmentation and substitution. *CRC handbook of biomedical engineering* (1995), 2100–2109.
- [119] KALLIE, C. S., SCHRATER, P. R., AND LEGGE, G. E. Variability in stepping direction explains the veering behavior of blind walkers. *Journal of experimental psychology. Human perception and performance* 33, 1 (2007), 183–200.
- [120] KALMAN, R. E. Contributions to the theory of optimal control. *Bol. soc. mat. mexicana* 5, 2 (1960), 102–119.
- [121] KALMAN, R. E. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering* 82(1) (1960), 35–45.
- [122] KALMAN, R. E., AND BUCY, R. S. New results in linear filtering and prediction theory. *Transactions of the ASME* (1961).
- [123] KANG, J., KIM, D., AND KIM, Y. Rss self-calibration protocol for wsn localization. In *2007 2nd International Symposium on Wireless Pervasive Computing* (2007).
- [124] KARL, J. Principles of feedback and control, pid control. cours caltech, online <https://www.cds.caltech.edu/~murray/courses/cds101/fa02/caltech/astrom-ch6.pdf>, 2002.
- [125] KATZ, B. F., AND NOISTERNIG, M. A comparative study of interaural time delay estimation methods. *The Journal of the Acoustical Society of America* 135, 6 (2014), 3530–3540.
- [126] KATZ, B. F. G., KAMMOUN, S., PARSEIHIAN, G., GUTIERREZ, O., BRILHAULT, A., AUVRAY, M., TRUILLET, P., DENIS, M., THORPE, S., AND JOUFFRAIS, C. Navig : Augmented reality guidance system for the visually impaired : Combining object localization, gnss, and spatial audio. *Virtual Reality* 16, 4 (2012), 253–269.
- [127] KAY, L. A new or improved blind aid (brevet uk ref. gb1098927a), 1968.
- [128] KHOO, W. L., SEIDEL, E. L., AND ZHU, Z. Designing a virtual environment to evaluate multimodal sensors for assisting the visually impaired. In *International Conference on Computers for Handicapped Persons* (2012), Springer, pp. 573–580.
- [129] KISTLER, D. J., AND WIGHTMAN, F. L. A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. *The Journal of the Acoustical Society of America* 91, 3 (1992), 1637–1647.
- [130] KLOBUCHAR, J. A. Ionospheric effects on earth-space propagation. Tech. rep., Air Force Geophysics Lab, 1983.
- [131] KOK, M., HOL, J. D., SCHÖN, T. B., GUSTAFSSON, F., AND LUINGE, H. Calibration of a magnetometer in combination with inertial sensors. In *2012 15th International Conference on Information Fusion* (2012), pp. 787–793.
- [132] KUC, R. Binaural sonar electronic travel aid provides vibrotactile cues for landmark, reflector motion and surface texture classification. *IEEE transactions on biomedical engineering* 49, 10 (2002), 1173–1180.
- [133] KUMAR, P., REDDY, L., AND VARMA, S. Distance measurement and error estimation scheme for rssi based localization in wireless sensor networks. In *2009 Fifth International Conference on Wireless Communication and Sensor Networks (WCSN)* (2009), IEEE, pp. 1–4.

- [134] LABBE, R. *Kalman and Bayesian Filters in Python*, en ligne : <https://github.com/rlabbe/kalman-and-bayesian-filters-in-python> ed. 2015.
- [135] LADD, A., BEKRIS, K., MARCEAU, G., RUDYS, A., WALLACH, D., AND KAVRAKI, L. Using wireless ethernet for localization. In *IEEE/RSJ International Conference on Intelligent Robots and System* (Lausanne, Switzerland, 2002), vol. 1, pp. 402–408.
- [136] LADETTO, Q., AND MERMINOD, B. Digital magnetic compass and gyroscope integration for pedestrian navigation. In *9th Saint Petersburg International Conference on Integrated Navigation Systems, Saint Petersburg, Russia* (2002).
- [137] LAHAV, O., SCHLOERB, D. W., AND SRINIVASAN, M. A. Rehabilitation program integrating virtual environment to improve orientation and mobility skills for people who are blind. *Computers & Education* 80 (2015), 1–14.
- [138] LANG, R., KOEGEL, L. K., ASHBAUGH, K., REGESTER, A., ENCE, W., AND SMITH, W. Physical exercise and individuals with autism spectrum disorders : A systematic review. *Research in Autism Spectrum Disorders* 4, 4 (2010), 565–576.
- [139] LANGLEY, R. B. Dilution of precision. *GPS world* 10, 5 (1999), 52–59.
- [140] LARCHER, V., AND JOT, J.-M. Techniques d'interpolation de filtres audio-numériques, application à la reproduction spatiale des sons sur écouteurs. In *Proc. CFA : Congrès Français d'Acoustique* (1997).
- [141] LENAY, C., GAPENNE, O., HANNETON, S., MARQUE, C., AND GENOUËLLE, C. Sensory substitution : Limits and perspectives. In *Touching for Knowing : Cognitive Psychology of Haptic Manual Perception*, Advances in Consciousness Research. John Benjamins Publishing Company, Amsterdam, Netherlands, 2003, pp. 276–292.
- [142] LESSARD, N., PARÉ, M., LEPORE, F., AND LASSONDE, M. Early-blind human subjects localize sound sources better than sighted subjects. *Nature* 395, 6699 (1998), 278–280.
- [143] LEWALD, J. Vertical sound localization in blind humans. *Neuropsychologia* 40, 12 (2002), 1868–1872.
- [144] LI, B., GALLAGHER, T., DEMPSTER, A. G., AND RIZOS, C. How feasible is the use of magnetic field alone for indoor positioning? In *International Conference on Indoor Positioning and Indoor Navigation (IPIN2012)* (2012).
- [145] LI, H. Channel capacity and channel estimation of ofdm ultra-wide-band systems. In *2012 International Conference on Computer Science and Electronics Engineering* (2012), vol. 2, pp. 152–156.
- [146] LI, J. Q., WELCHOWSKI, T., SCHMID, M., LETOW, J., WOLPERS, A. C., HOLZ, F. G., AND FINGER, R. P. Retinal diseases in europe. *Prevalence, Incidence and Healthcare Needs* (2017).
- [147] LIM, H., LEE, H., AND KIM, H. J. Onboard flight control of a micro quadrotor using single strapdown optical flow sensor. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems* (2012), pp. 495–500.
- [148] LIU, H., DARABI, H., BANERJEE, P., AND LIU, J. Survey of wireless indoor positioning techniques and systems. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)* 37, 6 (2007), 1067–1080.
- [149] LOOMIS, J. M., GOLLEDGE, R. G., AND KLATZKY, R. L. Navigation system for the blind : Auditory display modes and guidance. *Presence : Teleoperators and Virtual Environments* 7, 2 (1998), 193–203.
- [150] LOOMIS, J. M., KLATZKY, R. L., AND GIUDICE, N. A. -sensory substitution of vision : Importance of perceptual and cognitive processing. In *Assistive Technology for Blindness and Low Vision*. CRC Press, 2018, pp. 179–210.

- [151] LUCAS, B. D., AND KANADE, T. An iterative image registration technique with an application to stereo vision.
- [152] MACDONALD, J. A., HENRY, P. P., AND LETOWSKI, T. R. Spatial audio through a bone conduction interface : Audición espacial a través de una interfase de conducción ósea. *International Journal of Audiology* 45, 10 (2006), 595–599.
- [153] MADGWICK, S. O., HARRISON, A. J., AND VAIDYANATHAN, R. Estimation of imu and marg orientation using a gradient descent algorithm. In *2011 IEEE International Conference on Rehabilitation Robotics* (2011), pp. 1–7.
- [154] MADGWICK, S. O., HARRISON, A. J., AND VAIDYANATHAN, R. Estimation of imu and marg orientation using a gradient descent algorithm. In *2011 IEEE International Conference on Rehabilitation Robotics* (2011), pp. 1–7.
- [155] MAIDENBAUM, S., ABBOUD, S., AND AMEDI, A. Sensory substitution : Closing the gap between basic research and widespread practical visual rehabilitation. *Neuroscience & Biobehavioral Reviews* 41 (2014), 3–15.
- [156] MAKINO, H., ISHII, I., AND NAKASHIZUKA, M. Development of navigation system for the blind using gps and mobile phone combination. In *Proceedings of 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (1996), vol. 2, pp. 506–507 vol.2.
- [157] MAKOUS, J. C., AND MIDDLEBROOKS, J. C. Two-dimensional sound localization by human listeners. *The journal of the Acoustical Society of America* 87, 5 (1990), 2188–2200.
- [158] MANDANICI, M., RODÀ, A., CANAZZA, S., AND CAVAGNOLI, G. Following the cuckoo sound : A responsive floor to train blind children to avoid veering. In *International Conference on Smart Objects and Technologies for Social Good* (2017), Springer, pp. 11–20.
- [159] MANNING, C., MERMAGEN, T., AND SCHARINE, A. The effect of sensorineural hearing loss and tinnitus on speech recognition over air and bone conduction military communications headsets. *Hearing Research* 349 (2017), 67–75.
- [160] MARINS, J. L., YUN, X., BACHMANN, E. R., MCGHEE, R. B., AND ZYDA, M. J. An extended kalman filter for quaternion-based orientation estimation using marg sensors. In *Proceedings 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems. Expanding the Societal Role of Robotics in the the Next Millennium* (2001), vol. 4, pp. 2003–2011.
- [161] MCCALLUM, S. Gamification and serious games for personalized health. *pHealth* (2012).
- [162] MEIJER, P. B. An experimental system for auditory image representations. *IEEE transactions on biomedical engineering* 39, 2 (1992), 112–121.
- [163] MERSHON, D. H., AND KING, L. E. Intensity and reverberation as factors in the auditory perception of egocentric distance. *Perception & Psychophysics* 18, 6 (1975), 409–415.
- [164] MESSONNIER, J.-C., LYZWA, J.-M., DEVALLEZ, D., AND DE BOISHERAUD, C. Object-based audio recording methods. In *Audio Engineering Society Convention 140* (2016), Audio Engineering Society.
- [165] METROPOLIS, N., AND ULAM, S. The monte carlo method. *Journal of the American statistical association* 44, 247 (1949), 335–341.
- [166] MIDDLEBROOKS, J. C. Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency. *The Journal of the Acoustical Society of America* 106, 3 (1999), 1493–1510.
- [167] MILLER, J. D. Slab : A software-based real-time virtual acoustic environment rendering system. In *Proceedings of the 2001 International Conference on Auditory Display* (2001), p. 2.

- [168] MILLMAN G.H., G. *Modern Radar*, John Wiley & Sons, New York ed. 1965.
- [169] MILLS, A. W. On the minimum audible angle. *The Journal of the Acoustical Society of America* 30, 4 (1958), 237–246.
- [170] MINNAAR, P., PLOGSTIES, J., OLESEN, S. K., CHRISTENSEN, F., AND MØLLER, H. The interaural time difference in binaural synthesis. In *Audio Engineering Society Convention 108* (2000), Audio Engineering Society.
- [171] MOORE, S. K. Super-accurate gps coming to smartphones in 2018 [news]. *IEEE Spectrum* 54, 11 (2017), 10–11.
- [172] MOVAHEDI, A., BAHRAMI, F., MARANDI, S. M., AND ABEDI, A. Improvement in social dysfunction of children with autism spectrum disorder following long term kata techniques training. *Research in Autism Spectrum Disorders* 7, 9 (2013), 1054–1061.
- [173] NAM, J., KOLAR, M. A., AND ABEL, J. S. On the minimum-phase nature of head-related transfer functions. In *Audio Engineering Society Convention 125*. Audio Engineering Society. (2008), p. 8.
- [174] NIEMEYER, W., AND STARLINGER, I. Do the blind hear better? investigations on auditory processing in congenital or early acquired blindness ii. central functions. *Audiology* 20, 6 (1981), 510–515.
- [175] O’KEEFE, J., AND NADEL, L. *The Hippocampus as a Cognitive Map*. Oxford : Clarendon Press, 1978.
- [176] OPPENHEIM, A. V. *Signals and Systems*, 1st edition ed. Prentice Hall, Englewood Cliffs, N.J, 1982.
- [177] ORGANISATION MONDIALE DE LA SANTÉ. Global data on visual impairments. Tech. rep., 2010.
- [178] ORGANISATION MONDIALE DE LA SANTÉ. Recommandations mondiales sur l’activité physique pour la santé. Tech. rep., 2010.
- [179] P. SEDLACEK, SLANINA, M., AND KOVAC, D. An overview of indoor and outdoor positioning technologies with focus on their precision. *ElecktroRevue* 18, 6 (2016).
- [180] PAHLAVAN, K., AND LEVESQUE, A. H. *Wireless Information Networks*. John Wiley & Sons, 2005.
- [181] PANEELS, S. A., VARENNE, D., BLUM, J. R., AND COOPERSTOCK, J. R. The walking straight mobile application : Helping the visually impaired avoid veering. In *International Conference on Auditory Display* (Lodz, 2013).
- [182] PETRIE, H., JOHNSON, V., STROTHOTTE, T., RAAB, A., MICHEL, R., REICHERT, L., AND SCHALT, A. Mobic : An aid to increase the independent mobility of blind travellers. *British Journal of Visual Impairment* 15, 2 (1997), 63–66.
- [183] PETRUS, C., ADAMSON, S. R., BLOCK, L., EINARSON, S. J., SHARIFNEJAD, M., AND HARRIS, S. R. Effects of exercise interventions on stereotypic behaviours in children with autism spectrum disorder. *Physiotherapy Canada* 60, 2 (2008), 134–145.
- [184] PHILBECK, J. W., LOOMIS, J. M., AND BEALL, A. C. Visually perceived location is an invariant in the control of action. *Perception & Psychophysics* 59, 4 (1997), 601–612.
- [185] POPPER, A. N., FAY, R. R., AND POPPER, A. N. *Sound Source Localization*. Springer, 2005.
- [186] PRATAMA, A. R., WIDYAWAN, AND HIDAYAT, R. Smartphone-based pedestrian dead reckoning as an indoor positioning system. In *2012 International Conference on System Engineering and Technology (ICSET)* (2012), pp. 1–6.
- [187] PRESSEY, N. Mowat sensor. *Focus* 11, 3 (1977), 35–39.

- [188] RAKERD, B., AND HARTMANN, W. M. Localization of sound in rooms, ii : The effects of a single reflecting surface. *The Journal of the Acoustical Society of America* 78, 2 (1985), 524–533.
- [189] RANDELL, C., DJIALLIS, C., AND MULLER, H. Personal position measurement using dead reckoning. In *Seventh IEEE International Symposium on Wearable Computers, 2003. Proceedings.* (2003), pp. 166–173.
- [190] RASSENEUR, L., ZEGARRA, J., AND LINDER, N. Navi rando, l'application qui permet de randonner sans y voir ! *Lumen Magazine*, janv/fev 2016, 65–68.
- [191] RAYLEIGH, L. Xii. on our perception of sound direction. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 13, 74 (1907), 214–232.
- [192] RENAUDIN, V., AFZAL, M. H., AND LACHAPELLE, G. Complete triaxis magnetometer calibration in the magnetic domain. *Journal of Sensors* (2010).
- [193] RIPLEY, B. D. *Stochastic Simulation*, vol. 316. John Wiley & Sons, 2009.
- [194] ROBERTS, J. M., CORKE, P., AND BUSKEY, G. Low-cost flight control system for a small autonomous helicopter. In *Proceedings ICRA'03 IEEE International Conference on Robotics and Automation 2003* (2003), vol. 1, pp. 546–551.
- [195] ROBUSTELLI, U., BAIOCCHI, V., AND PUGLIANO, G. Assessment of dual frequency gnss observations from a xiaomi mi 8 android smartphone and positioning performance analysis. *Electronics* 8, 1 (2019), 91.
- [196] ROKEM, A., AND AHISSAR, M. Interactions of cognitive and auditory abilities in congenitally blind individuals. *Neuropsychologia* 47, 3 (2009), 843–848.
- [197] RUSSELL, L. Travel path sounder. In *Proceedings of Rotterdam Mobility Res. Conference* (1965).
- [198] SALMON, F., AUSSAL, M., HENDRICKX, E., MESSONNIER, J.-C., AND MILLOT, L. Optimization of interactive binaural processing. In *Audio Engineering Society Convention 143* (2017), Audio Engineering Society.
- [199] SANDEL, T. T., TEAS, D. C., FEDDERSEN, W. E., AND JEFFRESS, L. A. Localization of sound from single and paired sources. *The Journal of the Acoustical Society of America* 27, 5 (1955), 842–852.
- [200] SANTI, P. Topology control in wireless ad hoc and sensor networks. *Topology Control in Wireless Ad Hoc and Sensor Networks* 37, 2 (2005), 31.
- [201] SARGOLINI, F., FYHN, M., HAFTING, T., MCNAUGHTON, B. L., WITTER, M. P., MOSER, M.-B., AND MOSER, E. I. Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science* 312, 5774 (2006), 758–762.
- [202] SHAREEF, A., ZHU, Y., MORENO, V. M., AND PIGAZO, A. Localization using extended kalman filters in wireless sensor networks. *Kalman Filter Recent Advances and Applications* (2009), 297–320.
- [203] SHINN-CUNNINGHAM, B. G. Localizing sound in rooms. *ACM/SIGGRAPH and Eurographics Campfire : Acoustic Rendering for Virtual Environments* (2001), 1–6.
- [204] SOWA, M., AND MEULENBROEK, R. Effects of physical exercise on autism spectrum disorders : A meta-analysis. *Research in Autism Spectrum Disorders* 6, 1 (2012), 46–57.
- [205] STANDRING, S., AND GRAY, H., Eds. *Gray's Anatomy : The Anatomical Basis of Clinical Practice*, 40. ed., reprinted ed. Churchill Livingstone Elsevier, Edinburgh, 2009.
- [206] STENFELT, S., AND GOODE, R. L. Bone-conducted sound : Physiological and clinical aspects :. *Otology & Neurotology* 26, 6 (2005), 1245–1261.

- [207] STIRLING, R., COLLIN, J., FYFE, K., AND LACHAPPELLE, G. An innovative shoe-mounted pedestrian navigation system. In *Proceedings of European Navigation Conference GNSS (2003)*, vol. 110.
- [208] STITT, P., HENDRICKX, E., MESSONNIER, J. C., AND KATZ, B. F. The role of head tracking in binaural rendering. In *29th Tonmeistertagung, International VDT Convention (2016)*.
- [209] STRACHAN, S., ESLAMBOLCHILAR, P., MURRAY-SMITH, R., HUGHES, S., AND O'MODHRAIN, S. Gpstunes : Controlling navigation via audio feedback. In *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services - MobileHCI '05 (Salzburg, Austria, 2005)*, ACM Press, p. 275.
- [210] STRELOW, E. R., AND BRABYN, J. A. Locomotion of the blind controlled by natural sound cues. *Perception* 11, 6 (1982), 635–640.
- [211] TAKASU, T., AND YASUDA, A. Development of the low-cost rtk-gps receiver with an open source program package rtklib. In *International Symposium on GPS/GNSS (2009)*, International Convention Center Jeju Korea, pp. 4–6.
- [212] TANG, Y., AND LI, C. Wearable indoor position tracking using onboard k-band doppler radar and digital gyroscope. In *2015 IEEE MTT-S 2015 International Microwave Workshop Series on RF and Wireless Technologies for Biomedical and Healthcare Applications (IMWS-BIO) (Taipei, Taiwan, Sept. 2015)*, IEEE, pp. 76–77.
- [213] TEAM, G. P. *Global Positioning System (Gps) Standard Positioning Service (Sps) Performance Analysis Report*. Washington : Federal Aviation Administration, 2014.
- [214] TIME DOMAIN. Distributed calibration of timedomain uwb ranging radios. Tech. rep., Time domain, 2015.
- [215] TOLMAN, E. C. Cognitive maps in rats and men. *Psychological Review* 55, 4 (1948), 189–208.
- [216] TOPLISS, J. W., ZAPPI, V., AND MCPHERSON, A. Latency performance for real-time audio on beaglebone black. Tech. rep., Queen Mary University of London, 2014.
- [217] TROJER, T. Arduino-DW1000 : A library that offers functionality to use decawave's DW1000 chips/modules with arduino. <https://github.com/thotro/arduino-dw1000>.
- [218] US FEDERAL COMMUNICATION COMMISSION. Revision of part 15 of the commission's rules regarding ultra-wideband transmission systems. Tech. rep., US Federal Communication Commission, 2002.
- [219] VASCONCELOS, J. F., ELKAIM, G., SILVESTRE, C., OLIVEIRA, P., AND CARDEIRA, B. Geometric approach to strapdown magnetometer calibration in sensor frame. *IEEE Transactions on Aerospace and Electronic Systems* 47, 2 (2011), 1293–1306.
- [220] VOSS, P., LASSONDE, M., GOUGOUX, F., FORTIN, M., GUILLEMOT, J.-P., AND LEPORE, F. Early- and late-onset blind individuals show supra-normal auditory abilities in far-space. *Current Biology* 14, 19 (2004), 1734–1738.
- [221] WALLACH, H. On sound localization. *The Journal of the Acoustical Society of America* 10, 4 (1939), 270–274.
- [222] WAN, E., AND VAN DER MERWE, R. The unscented kalman filter for nonlinear estimation. In *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Lake Louise, Alta., Canada, 2000)*, pp. 153–158.
- [223] WANG, B., LIU, X., YU, B., JIA, R., AND GAN, X. Pedestrian dead reckoning based on motion mode recognition using a smartphone. *Sensors (Basel, Switzerland)* 18, 6 (2018).
- [224] WANG, J., TANG, Y., MUÑOZ-FERRERAS, J.-M., GÓMEZ-GARCÍA, R., AND LI, C. An improved indoor localization solution using a hybrid uwb-doppler system with kalman filter. In *2018 IEEE Radio and Wireless Symposium (RWS) (2018)*, pp. 181–183.



- [225] WANT, R., HOPPER, A., FALCÃO, V., AND GIBBONS, J. The active badge location system. *Transactions on Information Systems (TOIS)* (1992).
- [226] WARUSFEL, O. Listen hrtf database. *online, IRCAM and AK, Available : <http://recherche.ircam.fr/equipes/salles/listen/index.html>* (2003).
- [227] WATANABE, K., IWAYA, Y., SUZUKI, Y., TAKANE, S., AND SATO, S. Dataset of head-related transfer functions measured with a circular loudspeaker array. *Acoustical Science and Technology* 35, 3 (2014), 159–165.
- [228] WELCH, G., ILIE, A., ALLEN, B. D., AND BISHOP, G. Measurement sample time optimization for human motion tracking/capture systems. In *Trends and Issues in Tracking for Virtual Environments, Workshop at the IEEE Virtual Reality 2007 Conference.*, p. 6.
- [229] WENZEL, E. M., ARRUDA, M., KISTLER, D. J., AND WIGHTMAN, F. L. Localization using nonindividualized head-related transfer functions. *The Journal of the Acoustical Society of America* 94, 1 (1993), 111–123.
- [230] WIGHTMAN, F., AND KISTLER, D. Multidimensional scaling analysis of head-related transfer functions. In *Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics* (1993), pp. 98–101.
- [231] WIGHTMAN, F. L., AND KISTLER, D. J. Resolution of front–back ambiguity in spatial hearing by listener and source movement. *The Journal of the Acoustical Society of America* 105, 5 (1999), 2841–2853.
- [232] WILLIS, S., AND HELAL, S. Rfid information grid and wearable computing solution to the problem of wayfinding for the blind user in a campus environment. In *IEEE International Symposium on Wearable Computers (ISWC 05)* (2005).
- [233] WILSON, J., WALKER, B. N., LINDSAY, J., CAMBIAS, C., AND DELLAERT, F. Swan : System for wearable audio navigation. In *2007 11th IEEE International Symposium on Wearable Computers* (2007), pp. 91–98.
- [234] WINER, K. Simple and effective magnetometer calibration. <https://github.com/kriswiner/MPU6050/wiki/Simple-and-Effective-Magnetometer-Calibration>.
- [235] WOODWORTH, R., BARBER, S., AND SCHLOSBERG, H. Experimental psychology. *Oxford and IBH Publishing* (1954).
- [236] Y. WANG, X. YANG, Y. ZHAO, LIU, Y., AND CUTHBERT, L. Bluetooth positioning using rssi and triangulation methods. In *2013 IEEE 10th Consumer Communications and Networking Conference (CCNC)* (Las Vegas, NV, 2013), IEEE, pp. 837–842.
- [237] Y. ZHAO, L. DONG, JIANG WANG, B. HU, AND Y. FU. Implementing indoor positioning system via zigbee devices. In *2008 42nd Asilomar Conference on Signals, Systems and Computers* (Pacific Grove, CA, USA, 2008), IEEE, pp. 1867–1871.
- [238] YIU, S., DASHTI, M., CLAUSSEN, H., AND PEREZ-CRUZ, F. Wireless rssi fingerprinting localization. *Signal Processing* 131 (2017), 235–244.
- [239] YOST, W. A. Sound source localization identification accuracy : Level and duration dependencies. *The Journal of the Acoustical Society of America* 140, 1 (2016), EL14–EL19.
- [240] YOST, W. A. History of sound source localization : 1850-1950. In *173rd Meeting of Acoustical Society of America and 8th Forum Acusticum* (Boston, Massachusetts, 2017).
- [241] YOST, W. A. Sound source localization identification accuracy : Envelope dependencies. *The Journal of the Acoustical Society of America* 142, 1 (2017), 173–185.

- [242] YOST, W. A., AND ZHONG, X. Sound source localization identification accuracy : Bandwidth dependencies. *The Journal of the Acoustical Society of America* 136, 5 (2014), 2737–2746.
- [243] ZAHORIK, P. Auditory distance perception in humans : A summary of past and present research. *Acta acustica united with acustica* 91 (2005), 13.
- [244] ZAHORIK, P., BANGAYAN, P., SUNDARESWARAN, V., WANG, K., AND TAM, C. Perceptual recalibration in human sound localization : Learning to remediate front-back reversals. *The Journal of the Acoustical Society of America* 120, 1 (2006), 343–359.
- [245] ZHANG, C. *Improving Crane Safety by Agent-Based Dynamic Motion Planning Using UWB Real-Time Location System*. PhD thesis, Concordia University, 2010.
- [246] ZHANG, W., AND KOSECKA, J. Image based localization in urban environments. In *Third International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'06)* (2006), IEEE, pp. (pp. 33–40).
- [247] ZOTKIN, D. N., DURAISWAMI, R., AND DAVIS, L. S. Rendering localized spatial audio in a virtual auditory space. *IEEE Transactions on multimedia* 6, 4 (2004), 553–564.
- [248] ZWIERS, M. P., VAN OPSTAL, A. J., AND CRUYSBERG, J. R. M. A spatial hearing deficit in early-blind humans. *The Journal of Neuroscience* 21, 9 (2001), RC142–RC142.

**Titre :** Techniques de spatialisation binaurale pour le guidage de sportifs non-voyants

**Mots clés :** Binaural, déficiences visuelle, traitement du signal, localisation, systèmes embarqués

**Résumé :** Dans la continuité d'autres travaux de la communauté du son binaural, nous pensons que le son spatialisé peut constituer un outil efficace pour guider des personnes aveugles, y compris pour la pratique sportive. Un système destiné au guidage par audio binaural doit être suffisamment précis et réactif pour pouvoir prendre en compte chaque mouvement du sujet à guider, ce qui a impliqué le développement et la mise en œuvre d'un système de localisation temps réel et d'un logiciel de spatialisation binaural à faible latence. Enfin nous avons intégré l'ensemble dans un dispositif embarqué. Les techniques les plus avancées de navigation globale par satellite (augmenté et multibande) ne sont pas toujours disponibles en intérieur ou en environnement urbain. C'est pourquoi nous avons travaillé sur des méthodes alternatives de localisation et de suivi temps réel pour l'intérieur. Tout d'abord nous avons développé une méthode de calibration et de latération temps réel robuste par réseau de balises Ultra WideBand utilisant un filtre de Kalman (UKF). Nous avons également développé une méthode originale de localisation par réseau de radars Doppler à onde continue non modulée. Nous avons montré qu'il était possible d'utiliser l'amplitude du signal Doppler pour estimer la distance à un objet mobile. Nous avons alors implémenté un filtre particulière qui permet la localisation en temps réel par hybridation des données de distance, des mesures de vitesse radiale Doppler et du cap fourni par une centrale inertielle. Dans le domaine de l'acoustique et de l'audio binaural, nous avons cherché à mieux comprendre les capacités des personnes à localiser et à suivre un son en mouvement. Pour cela nous avons mené des expériences en utilisant des sons naturels et des sons spatialisés par audio binaural. Nous avons pu montrer, que sur le plan azimu-

tal, les stimuli audio spatialisés permettaient une localisation comparable aux sons naturels, y compris avec les HRTFs (head-related transfer function) non individualisées et interpolées. Par ailleurs, nous avons pu montrer que même sur le plan azimutal, les stimuli obtenus par convolution de HRTFs étaient supérieurs au panning (ITD+ILD) pour les sons fixes et pour les sons en mouvement. En nous appuyant sur les travaux antérieurs de l'équipe, nous avons implémenté des algorithmes efficaces pour la spatialisation sonore temps-réel sur des plateformes embarquées disposant de peu de ressources. Pour une mise en œuvre efficace, cette approche temps-réel a impliqué une compréhension approfondie des sources de latence qu'elles soient liées au head-tracking ou au sous-système audio des systèmes d'exploitation modernes. Finalement nous avons mis en œuvre ces méthodes de localisation et ces techniques audio pour construire un dispositif de guidage où la source sonore précède continuellement la personne pour lui indiquer le chemin à suivre. Il a été conçu en lien avec des personnes déficientes visuelles, dans une démarche itérative et avec une approche centrée sur les besoins utilisateurs. Nous avons alors mené des expériences de guidage avec des personnes aveugles, en lien avec nos partenaires associatifs, qui ont permis d'évaluer différentes stratégies de contrôle. Nous avons ainsi pu confirmer que le son spatialisé pouvait constituer un outil efficace pour guider des personnes aveugles sans induire de charge cognitive pénalisante pour des pratiques sportives comme la marche, la course à pied ou le roller en autonomie partielle, y compris dans un contexte de recherche de performances.

**Title :** Binaural spatialization methods for guidance of the visually impaired for sports activities

**Keywords :** Binaural, visual impairment, signal processing, real time localisation, embedded systems

**Abstract :** In continuity with other work of the binaural community we believe that spatialized sound can be an effective tool for guiding blind people, including for sports practice. A system for binaural audio guidance must be sufficiently accurate and responsive to be able to take into account every movement of the subject. This has led us to the development and implementation of a real-time localization system and a low-latency binaural spatialization software. Finally, we integrated the whole system into an embedded device. The advanced global satellite navigation techniques (augmented and multi-band) are not always available indoor or in urban environments. This is why we have been working on alternative methods of location and real-time tracking for indoor applications. First of all we have developed a robust real-time calibration and lateration method using an Ultra WideBand beacon network using a Kalman filter (UKF). We have also developed an original localization method using an unmodulated continuous wave Doppler radar array. We showed that it is possible to use the amplitude of the Doppler signal to estimate the distance to a moving object. We have then implemented a particle filter to allow real time localization by hybridization of distance data, measurements of radial Doppler velocity augmented by the heading provided by an inertial unit. In the field of acoustics and binaural audio, we have focused on understanding people's abilities to locate and track sound in motion. To do this we conducted experiments using natural and spatialized sounds using a binaural audio engine. These experi-

ments showed that, on the azimuthal plane, spatialized audio stimuli allow for a localization comparable to natural sounds, even with non-individualized and interpolated HRTFs (head-related transfer function). Moreover, we were able to show that even in the azimuthal plane, the stimuli obtained by HRTF convolution are superior to panning (ITD+ILD) both for fixed and moving sounds. Based on the previous work of the team, we have implemented efficient algorithms for the real-time sound spatialization on embedded platforms with limited resources. For an efficient implementation, this real-time approach implied a thorough understanding of the sources of latency whether they are related to head-tracking or to the audio subsystem of modern operating systems. Finally, we used these localization methods and audio techniques to build a guidance device where a sound source continuously precedes the person to indicate the path to follow. It has been designed in collaboration with visually impaired people, in an iterative process and with an approach focused on the user needs. We then carried out guiding experiments with blind people in conjunction with our associative partners, which made it possible to evaluate different control strategies. We have thus been able to confirm that spatialized sound could constitute an effective tool for guiding blind people, without inducing any penalizing cognitive load, for practicing sports such as walking, running or rollerblading in partial autonomy, including the context of performance.