



Détection de motifs temporels dans les environnements multi-perceptifs. Application à la classification automatique des Activités de la Vie Quotidienne d'une personne suivie à domicile par télémedecine

Anthony Fleury

► To cite this version:

Anthony Fleury. Détection de motifs temporels dans les environnements multi-perceptifs. Application à la classification automatique des Activités de la Vie Quotidienne d'une personne suivie à domicile par télémedecine. Sciences de l'ingénieur [physics]. Université Joseph-Fourier - Grenoble I, 2008. Français. NNT: . tel-00336400v2

HAL Id: tel-00336400

<https://theses.hal.science/tel-00336400v2>

Submitted on 4 Nov 2008

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.



Université de Grenoble 1 – Joseph Fourier

École Doctorale EEATS – ED 220

Électronique, Électrotechnique, Automatique et Traitement du Signal

THÈSE

pour l'obtention du grade de

DOCTEUR DE L'UNIVERSITÉ JOSEPH FOURIER

Spécialité : Traitement du Signal

Présentée et soutenue publiquement par

Anthony FLEURY

Le 24 Octobre 2008

Détection de motifs temporels dans les environnements multi-perceptifs. Application à la classification automatique des Activités de la Vie Quotidienne d'une personne suivie à domicile par télé médecine

Sous la direction de **Norbert NOURY** et **Michel VACHER**

JURY :

<i>Présidente</i>	Michèle ROMBAUT	Professeur des Universités, Université Joseph Fourier, Grenoble I
<i>Rapporteurs</i>	Olivier BASSET François CHARPILLET	Professeur des Universités, Université Claude Bernard, Lyon I Directeur de Recherche INRIA, LORIA, Nancy
<i>Examineurs</i>	Jacques DUCHÊNE Kamiar AMINIAN	Professeur des Universités, Université Technologique de Troyes Professeur, École Polytechnique Fédérale de Lausanne, Suisse
<i>Directeurs</i>	Norbert NOURY Michel VACHER	Professeur des Universités, Université Claude Bernard, Lyon I Ingénieur de Recherche CNRS, Laboratoire d'Informatique de Grenoble
<i>Invités</i>	Nicolas VUILLERME Vincent RIALLE	Maître de Conférences, Université Joseph Fourier, Grenoble I Maître de Conférences - Praticien Hospitalier, Université Joseph Fourier, Grenoble I

Table des matières

Table des matières	8
Abstract	9
Résumé	11
Remerciements	13
I Introduction	15
1 Introduction	17
1.1 La population en France et dans le monde : Intérêt des solutions pour personnes âgées	19
1.1.1 La démographie française : évolution	19
1.1.1.1 Les chiffres connus	19
1.1.1.2 Projections de population	20
1.1.2 Comparaison avec la démographie mondiale	21
1.1.3 Les solutions orientées personnes âgées : une obligation ?	22
1.1.3.1 Problèmes posés par l'évolution de la population	22
1.1.3.2 Solutions proposées	24
1.2 La Télémédecine et les Gérontechnologies : les enjeux et les nécessités	26
1.2.1 Le concept de gérontechnologie	26
1.2.2 La télémédecine	26
1.2.3 Les technologies pour la santé : une nécessité ?	27
2 État de l'art	29
2.1 L'actimétrie embarquée	31
2.1.1 Les types de capteurs	31
2.1.2 Les centrales équipées de capteurs cinématiques	32
2.1.3 Applications	33
2.1.3.1 Analyse du mouvement et dépense énergétique	33
2.1.3.2 La détection de la chute	34
2.2 La reconnaissance sonore	35
2.2.1 La reconnaissance de la parole pour l'assistance aux personnes	35
2.2.2 Application de la reconnaissance des sons et de la parole à l'habitat et à la domotique	36
2.3 La domotique au service de la santé et de l'assistance à domicile	37

2.3.1	Les dispositifs de secours et de téléassistance	37
2.3.2	La localisation de la personne	38
2.3.2.1	Suivi à l'intérieur du domicile	38
2.3.2.2	Les risques de fugue	39
2.3.3	La détection de situations anormales	40
2.3.4	Les habitats intelligents	41
2.3.4.1	Surveillance et suivi de personnes	41
2.3.4.2	L'assistance à domicile	44
2.3.4.3	Le suivi d'activités	45
2.3.4.4	La télémédecine	46
2.4	Éthique et habitats intelligents	47
2.5	Conclusion sur cet état de l'art	48
3	Objectifs des travaux de thèse	49
3.1	Historique des travaux du laboratoire	51
3.2	Nécessité d'une approche multi-paramétrique	52
3.3	Choix des capteurs	52
3.4	Les Activités de la vie Quotidienne (critères ADL)	53
3.4.1	Les ADLs de Katz	53
3.4.2	Les Activités Instrumentales de la vie quotidiennes (IADLs)	53
3.4.3	La grille AGGIR	55
3.4.4	Objectivation des critères	57
3.5	Organisation du manuscrit	57
II	La classification des activités dans le HIS : de la mise en place des capteurs et des algorithmes...	59
4	Actimétrie embarquée : Conception, algorithmes et tests de l'ACTIM 6D	61
4.1	Conception de la carte ACTIM6D	63
4.1.1	Description	63
4.1.2	Positionnement du capteur	64
4.1.3	Acquisition et enregistrement des données	65
4.1.4	Récupération des données enregistrées sur carte mémoire	66
4.1.5	Perturbations extérieures	68
4.2	Algorithmes	70
4.2.1	Les données accélérométriques	70
4.2.1.1	Description des données brutes	70
4.2.1.2	Pré-traitement des données accélérométriques	71
4.2.1.3	Identification des motifs temporels	71
4.2.2	Les données magnétométriques	72
4.2.2.1	Description des données brutes	72
4.2.2.2	Identification des motifs temporels	74
4.2.2.2.1	Motifs et dépendance par rapport à l'orientation du porteur	74
4.2.2.2.2	Passage dans le domaine des quaternions	75
4.2.3	Segmentation du signal	77
4.2.3.1	Critère choisi pour la segmentation	77
4.2.3.2	Motifs détectés	77
4.2.4	Classification	78
4.2.4.1	La transformée en ondelettes continue pour la classification	78
4.2.4.2	Réalisation des motifs de détection	81

4.2.4.2.1	Séquence Debout-Assis	81
4.2.4.2.2	Séquence Assis-Debout	83
4.2.4.2.3	Séquence Debout-Allongé	84
4.2.4.2.4	Séquence Allongé-Debout	85
4.2.5	Périodes de marche	85
4.2.5.1	Détection des périodes de marche	85
4.2.5.2	Périodes de marche et segmentation	87
4.3	Résultats sur quelques scénarios	88
4.4	Fichier de sortie et résultats finaux	90
4.5	Expérimentations et résultats	91
4.5.1	But de l'expérimentation et population	91
4.5.2	Protocole expérimental	92
4.5.3	Résultats	92
4.6	Conclusion et discussion sur l'implémentation de la carte ACTIM6D	94
4.6.1	Limitations matérielles	94
4.6.1.1	Bruit de mesure	94
4.6.1.2	Positionnement du capteur	94
4.6.2	Logiciel	95
4.6.2.1	Intérêt des ondelettes	95
4.6.2.2	Discussion sur la segmentation et la classification	95
4.6.2.3	Analyse temps-fréquences de la marche	96
4.6.2.4	Perspectives	96
4.6.2.4.1	Correction des incohérences des données	96
4.6.2.4.2	Système embarqué?	97
5	Les données sonores	99
5.1	Le son dans le HIS	101
5.1.1	Organisation globale du système de reconnaissance de sons et de la parole	102
5.1.2	Matériel utilisé	102
5.1.2.1	Microphones	102
5.1.2.2	Carte d'acquisition	102
5.1.2.3	Acquisition, traitement et stockage des données	102
5.1.3	Microphones : installation et réglages	103
5.1.3.1	Choix de l'emplacement des microphones	103
5.1.3.2	Réglage des gains	104
5.1.4	Acquisition des données sonores	105
5.1.5	Objectifs	105
5.2	Algorithmes de traitement des sons	106
5.2.1	Détection	106
5.2.2	Estimation du Rapport Signal sur Bruit (RSB)	107
5.2.3	Segmentation entre signal de parole et son de la vie courante	108
5.2.3.1	Modèles acoustiques	108
5.2.3.2	Segmentation parole/sons de la vie courante par classification .	109
5.2.4	Classification des sons de la vie courante	109
5.2.5	Reconnaissance automatique de la parole	110
5.3	Bases d'apprentissage	110
5.3.1	Corpus des sons de la vie courante	110
5.3.2	Paroles	111
5.3.2.1	Construction des modèles de langage et des modèles acoustiques	111
5.3.2.2	Base de tests	112
5.4	Résultats en laboratoire	112

5.4.1	Détection des sons par calcul d'un arbre d'ondelettes	112
5.4.2	Segmentation parole/sons de la vie courante	113
5.4.3	Classification des sons	113
5.4.4	Reconnaissance des mots clés de détresse	114
5.5	Implémentation : le logiciel AuditHIS	114
5.5.1	Description générale	114
5.5.2	Données extraites	116
5.6	Exploitation des données	117
5.6.1	Extraction des données	117
5.6.2	Règles de classification	117
5.6.2.1	Parole	117
5.6.2.2	Sons de la vie courante	118
5.6.3	Fichier de sortie	118
5.7	Expérimentations préliminaires	118
5.7.1	Expérimentation 1	119
5.7.1.1	Protocole	119
5.7.1.1.1	Reconnaissance sonore	119
5.7.1.1.2	Reconnaissance de la parole	120
5.7.1.2	Résultats expérimentaux	120
5.7.1.2.1	La parole	120
5.7.1.2.2	Les sons de la vie courante	121
5.7.1.2.3	Résultats globaux	121
5.7.2	Expérimentation 2	122
5.7.2.1	Protocole	122
5.7.2.1.1	Reconnaissance sonore	122
5.7.2.1.2	Reconnaissance de la parole	123
5.7.2.2	Résultats expérimentaux	123
5.7.3	Expérimentation 3	124
5.7.3.1	Protocole	124
5.7.3.2	Résultats expérimentaux	124
5.8	Analyse des résultats de la détection du son et de la parole	126
6	Un prélude à la classification de données : formatage et interprétation des données	27
6.1	Le HIS de Grenoble	130
6.1.1	Présentation générale	130
6.1.2	Positionnement des capteurs	130
6.1.3	Architecture, connectivité et acquisition des données	131
6.1.4	Matériel de la pièce technique	132
6.2	Les données actimétriques	133
6.2.1	Données brutes	133
6.2.2	Extraction des paramètres	133
6.3	Les données sonores	133
6.3.1	Données brutes	133
6.3.2	Extraction des paramètres	134
6.4	Les données de présence	134
6.4.1	Données brutes	134
6.4.2	Extraction des paramètres	135
6.5	Les données liées aux instruments de la vie quotidiennes	135
6.5.1	Données brutes	135
6.5.2	Extraction des paramètres	136
6.6	Les données environnementales	136

6.6.1	Données brutes	136
6.6.2	Extraction des paramètres	137
6.7	Conclusion	138
7	La classification de données : de l'apprentissage aux expérimentations	139
7.1	Les activités de la vie quotidienne	141
7.1.1	Définition	141
7.1.2	Critères sur les activités	141
7.1.3	Classification et multimodalité des activités de la vie quotidienne	142
7.2	Rappel sur les Séparateurs à Vaste Marge	143
7.2.1	Le problème de la classification	143
7.2.1.1	La classification : un problème ancien	143
7.2.1.2	Classification non supervisée ou supervisée	143
7.2.2	Les séparateurs à vaste marge pour la classification binaire	144
7.2.2.1	Classification avec séparation linéaire	144
7.2.2.2	Séparateur non-linéaire : Kernel Trick	147
7.2.3	Classification à N classes ($N > 2$)	149
7.2.3.1	Méthode Un-Contre-Tous	150
7.2.3.2	Méthode Un-Contre-Un	150
7.3	Apprentissage et tests dans le HIS	152
7.3.1	Indexation des données	152
7.3.2	Validation croisée	153
III	... Aux expérimentations en conditions réelles	155
8	Déroulement de l'expérimentation	157
8.1	But de l'expérimentation	159
8.2	Matériel, équipement	160
8.3	Protocole expérimental	160
8.3.1	Contraintes imposées au sujet	160
8.3.2	Activités à réaliser	161
8.3.2.1	Se reposer	161
8.3.2.2	Se préparer et prendre un petit-déjeuner	161
8.3.2.3	S'habiller et se déshabiller	161
8.3.2.4	Se détendre	162
8.3.2.5	Activité d'hygiène	162
8.3.2.6	Élimination	162
8.3.2.7	Communication avec l'extérieur	162
8.4	Conclusion	163
9	Résultats	165
9.1	Population	167
9.2	Méthode d'obtention des résultats	167
9.3	Résultats expérimentaux	168
9.3.1	Noyau polynômial	168
9.3.2	Noyau gaussien	170
9.3.3	Comparaison des performances	171
9.4	Discussion	172

IV Discussion et Conclusion	173
10 Discussion Générale	175
10.1 Apports de ces travaux	177
10.1.1 ACTIM6D	177
10.1.2 Système de reconnaissance de sons et de paroles	178
10.1.3 Classification des activités	179
10.1.3.1 Base d'apprentissage	179
10.1.3.2 Ajout d'information <i>a priori</i>	180
10.1.3.3 Le temps comme modalité	180
10.1.3.4 Classe supplémentaire nécessaire : une classe de rejet	181
10.2 Validité des résultats	181
10.2.1 « Approche de la réalité par l'expérimentation » versus « Réalité »	181
10.2.2 Généralisation des modèles à la personne âgée ?	182
11 Conclusions et perspectives	185
11.1 Conclusions	187
11.2 Perspectives	188
11.2.1 Correction des données	188
11.2.2 Réalisation d'activité et réalisation effective de celle-ci	188
11.2.3 Transmission de ce type de données au médecin ou à la famille	188
11.2.4 HIS et modèle économique	188
Bibliographie	199
Table des figures	203
Liste des tableaux	205
Annexes	209
A Grilles AGGIR pour l'admission en établissement de soins de longue durée	209
B Schéma électrique de la carte ACTIM6D	211
C Plan de l'appartement avec les principales cotes	213
D Expérimentations : protocoles et lettres d'acceptation	215
E Ensemble des corpus de paroles : anodins, détresse et téléphone	223
F Publications	231

Abstract

In the near 2050, about one third of the French population will be over 65. The works of the team AFIRM of the TIMC-IMAG focus on the monitoring of elderly people at home, to detect, as earlier as possible, a loss of autonomy. This thesis work aims at objectivising criterions as ADL or AGGIR grids by automatically classifying the different Activities of Daily Living performed by the subject during the day. A Health Smart Home is used to do this.

Our Smart Home includes, in a real flat, Presence Infra-Red Sensors, (for localisation), door contacts (for the use of some conveniences), temperature and hygrometry sensors in the bathroom and microphones (sound and speech recognition with the GETALP team of the LIG). The subject is also equipped with a kinematic sensor that delivers on the postural transition (by pattern recognition with wavelets) and walk periods (frequency analysis).

This manuscript is compound of two major parts. The first one introduces the realisation of the kinematic sensor, its algorithms and its first evaluation ; but also the set-up, algorithms and the validation of the other sensors inside the flat and finally the Support Vector Machines algorithms to classify the activities of Daily Living (hygiene, toilets, preparing and having a meal, resting, sleeping, communication and dressing/undressing). The second part deals with the experimental protocol to validate these algorithms and the results of these validations on young and healthy subjects. It introduces the results and a discussion about their validity.

Keywords : Classification, Activities of Daily Living (ADL), Support Vector Machines (SVM), Health Smart Home, accelerometers, magnetometers, wavelet transform, design of wavelets, quaternion.

Résumé

À l'horizon 2050, environ un tiers de notre population sera âgée de soixante-cinq ans et plus. Les travaux de l'équipe AFIRM du TIMC-IMAG visent à surveiller les personnes âgées à domicile pour détecter une perte d'autonomie le plus précocement possible. Pour ce faire, les travaux de cette thèse tentent d'objectiver les critères ADL ou les grilles de type AGGIR, en classifiant de manière automatique les différentes activités de la vie quotidienne d'une personne par l'intégration de capteurs, créant un Habitat Intelligent pour la Santé (HIS).

L'appartement HIS possède des détecteurs de présence infrarouges (localisation), des contacteurs de porte (utilisation de certaines commodités), un capteur de température et d'hygrométrie dans la salle de bains et des microphones (classification des sons/ reconnaissance de la parole avec l'équipe GETALP du LIG). Un capteur cinématique embarqué détecte les transferts posturaux (reconnaissance de formes avec la transformée en ondelettes) et les périodes de marche (analyse fréquentielle).

La première partie de ce manuscrit présente la réalisation du capteur cinématique et les algorithmes associés puis une première validation sur des sujets jeunes suivi de la mise en place et de la validation des autres capteurs de l'appartement HIS et enfin l'algorithme de classification des sept activités de la vie quotidienne reconnues (hygiène, élimination, préparation et prise de repas, repos, habillage/déshabillage, détente et communication), par l'intermédiaire des séparateurs à vaste marge. La seconde partie décrit le protocole expérimental pour valider cette classification et discute de la généralisation des premiers résultats présentés.

Mots clés : Classification, Activités de la Vie Quotidienne, Séparateurs à Vaste Marge (SVM), Habitat Intelligent pour la Santé (HIS), accéléromètres, magnétomètres, transformée en ondelettes, création de bases d'ondelettes, quaternions.

Remerciements

La thèse est un travail collectif. C'est le travail d'un doctorant, mais aussi d'un laboratoire et d'une équipe de recherche (ou d'une entreprise). Ce travail sera d'autant meilleur que la collaboration entre chacun des parties sera bonne. En premier lieu, j'aimerai donc remercier le plus chaleureusement possible mes deux directeurs, Norbert Noury et Michel Vacher, pour leur soutien, leur aide, la réponse à mes diverses questions, mais avant tout pour m'avoir proposé ce sujet très intéressant il y a trois ans maintenant. J'espère qu'en regardant cette thèse après coup, j'aurai rempli toutes leurs attentes. Je remercie également le Professeur Jacques Demongeot pour son accueil au sein du laboratoire.

J'aimerai également remercier l'ensemble des personnes participant au jury de cette thèse et qui me donneront, j'en suis persuadé, des retours intéressants et une vision peut-être différente de mes travaux et de leurs implications :

- Le Dr. François Charpillet et le Pr. Olivier Basset pour avoir accepté de rapporter mes travaux,
- Les professeurs Michèle Rombaut, Jacques Duchêne et Kamiar Aminian pour avoir participé à mon jury en tant qu'examineurs,
- Les Docteurs Vincent Rialle et Nicolas Vuillerme pour avoir accepté d'y assister en tant qu'invités et avoir été présents pour mes questions et problèmes durant la thèse.

Le laboratoire et l'équipe de recherche sont également des points déterminants dans la réussite d'une thèse, et la bonne ambiance au sein de l'équipe AFIRM a contribué à ce travail. Je remercie donc d'abord Nicolas Vuillerme, qui m'a souvent donné de très bons conseils expérimentaux, Christophe, ici depuis le début de ma thèse et qui a partagé avec Pierre mon bureau au départ, mettant une très bonne ambiance!! Celle-ci continue avec l'aide de Nicolas et Remy. Damien Flammarion était également présent au départ et m'a apporté son aide et son soutien. Audrey, souvent au Canada pour sa co-tutelle au début de ma thèse, nous a quand même permis de passer de bons moments et nous a fortement inspiré avec ses complots. Je remercie également Mona pour sa présence et lui souhaite bonne chance pour sa rédaction. De même, je souhaite bonne chance à Tarek qui attaque une thèse. Je remercie tous les stagiaires qui sont passés dans nos bureaux (Elodie, Jeremy – qui a travaillé sur ACTIM6D avec Stanley également, Ali, maintenant en thèse dans une autre ville, Carole). Enfin je tiens particulièrement à remercier Julien et Ronald qui partagent mon bureau depuis quelques temps et qui m'ont souvent aidé, et beaucoup supporté au cours de la rédaction et de la fin de la thèse. J'aimerai également remercier les différentes personnes ayant participé à ce travail au LIG, Jean-François Sérignat qui m'a aidé et a relu une partie de la thèse, Hubert Glasson qui a beaucoup travaillé sur le logiciel AuditHIS et Noé Guirand. S'ajoutent aussi les membres des autres équipes de TIMC, que je croise également très fréquemment dans les couloirs, prêts à venir en aide quand ils en ont l'occasion ou qui partagent volontier une tasse de thé : Olivier H., Françoise (que l'on embête avec nos missions en retard!), Yannick,

Nadine, Jennifer, Adeline, Elise, Catalina, Myriam, Vivien, Estelle, Andry, Olivier C...

La thèse peut être aussi l'occasion d'enseigner. Cette partie a été pour moi une très bonne expérience, à l'IUP de chimie comme à Polytech'Grenoble en E2I, donc je remercie chaleureusement les encadrants et les élèves qui ont pu me croiser. J'espère leur avoir appris des choses en C ou en automatisme au moins ! Je remercie plus particulièrement Frédéric Rousseau pour son accueil, ses très bons conseils et pour m'avoir donné cette formidable occasion d'enseigner avec lui en E2I pendant deux ans.

La rédaction de ce manuscrit a touché à sa fin le 30 août. Je regrette de ne pas avoir pu le finir avant. Mon grand-père, qui m'a initié à l'informatique et m'a montré et donné son intérêt pour ces ordinateurs qu'il a appris à connaître très tard alors que c'était si loin de son métier, n'aura pas pu lire ou voir ces travaux, ni me voir les terminer. Il nous a quitté le 28 juin mais j'espère qu'il est encore en train de nous observer pour me voir écrire ces quelques lignes et lui dédicacer l'ensemble des pages suivantes. Ma grand-mère n'aura également pas l'occasion de lire ces lignes ou de raconter à tout le monde, comme elle le faisait avec Pascal et Philippe les ingénieurs, ce que son petit-fils a écrit. Cette maladie qui dévore son cerveau ne lui en laissera pas l'occasion... Je tiens à remercier du fond du cœur l'ensemble de ma famille et de mes amis qui se reconnaîtront. Ces remerciements s'adresse principalement à ma mère, sans qui je n'aurais pas pu mener à bien mes projets d'études, ainsi qu'à mon frère et Gérard. Je tiens à remercier également Séverine qui partage tout avec moi y compris les galères de cette rédaction, la course pour finir à temps... Un remerciement spécial est également adressé à mes grands parents qui m'appellent toujours régulièrement pour prendre des nouvelles malgré la distance ainsi qu'à mon père.

La dernière phrase est destinée à Yohan Payan, reprenant l'activité de l'équipe et sa direction. Je lui souhaite bonne chance dans sa nouvelle mission. Je suis persuadé qu'il pourra amener à l'activité un nouveau regard et qu'il contribuera à son développement.

Première partie

Introduction

Introduction

Mourir de vieillesse, c'est une mort rare, singulière et extraordinaire, et d'autant moins naturelle que les autres ; c'est la dernière et extrême sorte de mourir : plus elle est éloignée de nous, d'autant elle est moins espérable.

Essais de Montaigne, Charles Louandre

Sommaire

1.1 La population en France et dans le monde : Intérêt des solutions pour personnes âgées	19
1.1.1 La démographie française : évolution	19
1.1.1.1 Les chiffres connus	19
1.1.1.2 Projections de population	20
1.1.2 Comparaison avec la démographie mondiale	21
1.1.3 Les solutions orientées personnes âgées : une obligation ?	22
1.1.3.1 Problèmes posés par l'évolution de la population	22
1.1.3.2 Solutions proposées	24
1.2 La Télémédecine et les Géro-technologies : les enjeux et les nécessités . . .	26
1.2.1 Le concept de gérontechnologie	26
1.2.2 La télémédecine	26
1.2.3 Les technologies pour la santé : une nécessité ?	27

DANS ses dernières études parues en mai 2008, l'INSEE¹ annonce un vieillissement rapide de la population et rapporte une estimation montrant cette évolution : en 2050, un tiers de la population française aura plus de soixante ans. Ce vieillissement de la population a de nombreuses implications sur la manière de vivre de nos seniors (qui se trouvent affaiblis et peuvent parfois se retrouver seuls, éloignés de leur famille). C'est la raison pour laquelle des travaux de recherche se tournent vers la détection de la perte d'autonomie des personnes âgées et leur maintien à domicile. L'augmentation de la longévité de la population est un bienfait, à condition que celui-ci s'accompagne d'une bonne qualité de vie. Cela passe d'abord par une bonne santé, un certain confort matériel, mais aussi par la capacité de vivre ces années supplémentaires en autonomie chez soi. Ce constat, évident, a initié de nombreux travaux de recherche sur la détection de la perte d'autonomie et sur la mise en place de solutions technologiques visant à pallier les défaillances et ainsi permettre un maintien à domicile en toute sécurité.

Dans cette introduction, nous allons tout d'abord prendre connaissance de données de démographie française et mondiale, nous permettant ainsi de comprendre les enjeux et les nécessités de tels travaux.

1.1 La population en France et dans le monde : Intérêt des solutions pour personnes âgées

1.1.1 La démographie française : évolution

1.1.1.1 Les chiffres connus

Ces dernières années, les techniques médicales et chirurgicales se sont nettement améliorées. De ce fait, l'espérance de vie de la population augmente. Cette augmentation produit également un accroissement de l'âge moyen de la population et du « rapport de soutien démographique »². La figure 1.1 nous donne une représentation de cette évolution de la population au cours des vingt dernières années. Couplée à la figure 1.2, qui nous donne l'évolution du nombre des naissances par années, nous avons une vue d'ensemble de la population française actuelle et de son type d'accroissement.

Le nombre croissant de personnes âgées peut s'expliquer par ces deux figures et par les évolutions du domaine médical. En effet, comme le montre le second graphique, l'après-guerre fut une époque de « baby-boom »³, marqué un très fort accroissement des naissances après 1945. Ce nombre de naissances ne diminuera pas fortement avant les années 1970. Ces personnes nées lors de l'après-guerre atteignent désormais l'âge de 60 ans et s'approchent donc fortement de la retraite. La France fait donc face à un taux de natalité relativement constant et un taux de mortalité en forte diminution. Par conséquent, sa population vieillit.

¹Institut National de la Statistique et des Études Économiques : <http://www.insee.fr>

²Rapport du nombre de personnes âgées de 15 à 64 ans sur le nombre de personnes de plus de 65 ans. Ce rapport donne des indications sur les problèmes de financement des retraites

³Explosion de la natalité à une période donnée.

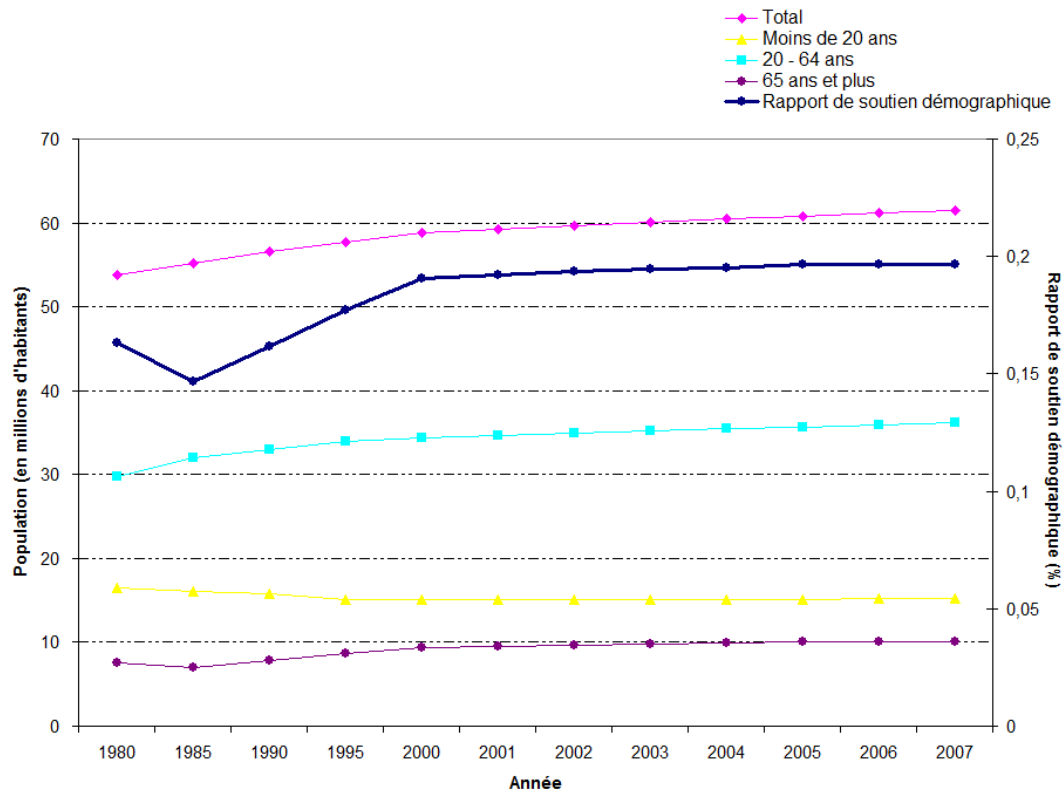


FIG. 1.1 – Évolution de la population française au cours des vingt dernières années. Le trait plus épais représente le rapport de soutien démographique, en pourcent, dont l'échelle se trouve sur la droite

Cette population née du « baby-boom » (1945 – 1970) et qui part donc à la retraite entre 2000 et 2020, fait entrer la France dans une période de « papy-boom »⁴.

L'amélioration des techniques médicales a cependant un effet secondaire car survivre plus longtemps à certaines maladies chroniques implique aussi de vivre plus longtemps avec elles. Certaines pathologies liées au vieillissement telles que la maladie d'Alzheimer et les maladies neurodégénératives apparentées sont en forte évolution et n'ont pour le moment aucun traitement (mis à part les traitements tentant de diminuer la vitesse de l'évolution de la maladie). En effet, le nombre de personnes âgées augmentant, et même si le pourcentage de la population atteint de ce type de maladie n'évoluait pas, la France serait toujours confrontée à une forte augmentation des personnes atteintes de troubles cognitifs et devrait ainsi faire face à une demande accrue d'entrée en institution. Le nombre de places disponibles en institutions équipées pour accueillir des personnes âgées (saines ou atteintes de troubles cognitifs) devra ainsi être fortement augmenté pour suivre ce « papy-boom ».

1.1.1.2 Projections de population

En mai 2008, l'INSEE a donné de nouvelles projections sur l'évolution de la population française d'ici à 2050 [Robert-Bobée, 2008, Coudin, 2008, Léon, 2008]. Les dernières prévi-

⁴Période qui voit une explosion du nombre de personnes qui prennent leur retraite, liée à un baby-boom passé.

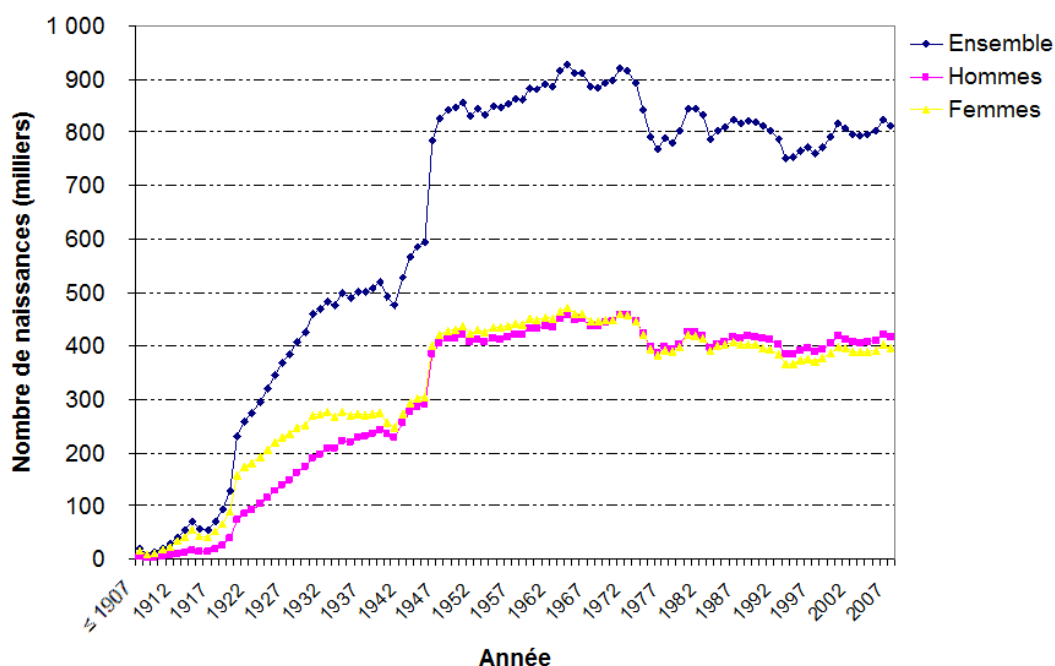


FIG. 1.2 – Répartition de la population française par âge et par sexe au 1er janvier 2008

sions (2002, basées sur le recensement de 1999) ont été revues à la hausse et ces nouveaux chiffres prévoient un vieillissement plus important de la population que celui annoncé dans les précédentes estimations.

Ces chiffres nous obligent à constater que la tendance actuelle au vieillissement de la population n'est pas prête à s'inverser et que des dispositions doivent être prises très vite pour prévenir les changements engendrés.

Le résultat de ces prévisions (en considérant le scénario central⁵) est donné par la figure 1.3. Ce graphique nous montre une augmentation notable de la population (qui serait portée à 70 millions d'habitants aux horizons 2050), avec une stagnation de son pourcentage dans la tranche d'âge 0 – 64 ans et une forte augmentation des seniors de plus de 65 ans. En effet, ce scénario central prévoit un taux de natalité constant (1,9 enfants par femmes) ce qui aura pour conséquence un vieillissement de la population.

1.1.2 Comparaison avec la démographie mondiale

Au sein de l'organisation des Nations Unies, un secteur consacré à la population (« United Nations, Population Division »⁶) centralise les données démographiques des différents pays et émet également des perspectives sur l'évolution de la population mondiale. Les dernières données sont basées sur les statistiques de 2006 et prévoient l'évolution jusqu'à 2050.

⁵Lors de chaque prévision, l'INSEE crée plusieurs scénarios, en prenant en compte une évolution différentes de certains facteurs tels que l'immigration. Le scénario le plus plausible est noté scénario central.

⁶<http://esa.un.org/unpp/index.asp>

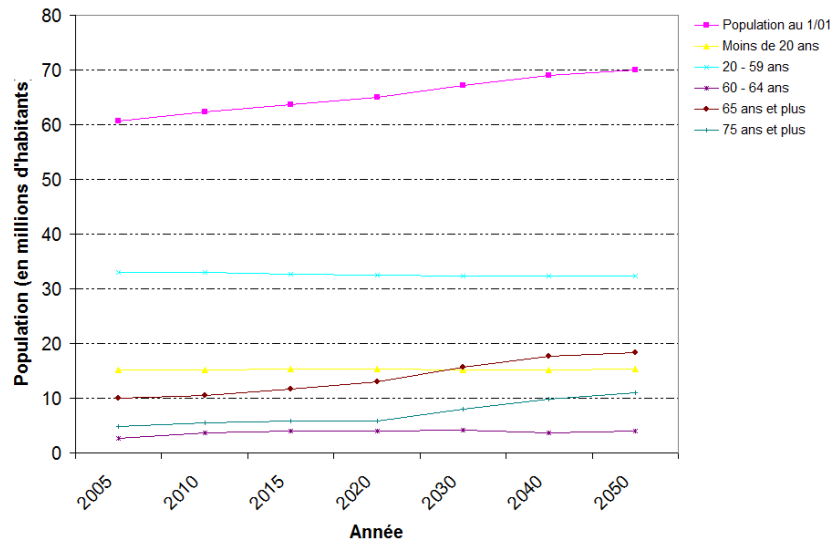


FIG. 1.3 – Répartition de la population française (prévisions 2005–2050)

À partir de ces données, nous pouvons tracer un premier graphique (figure 1.4) qui représente l'évolution de la population entre 1950 et 2050, ainsi que la répartition de la population par continent. Il nous montre un accroissement constant et rapide de la population mondiale. Cette évolution provient surtout de la forte augmentation des populations Européennes et Asiatiques (pays développés et très bien équipés sur le plan médical pour l'Europe, pays émergent et en forte croissance pour l'Asie). La population mondiale passerait, d'après ces statistiques, de 6 à 9,2 milliards d'habitants.

Le second graphique (figure 1.5), représente le pourcentage de la population âgée de 65 ans et plus dans le monde et par continents, entre 1950 et 2050 également. On constate bien que la France n'est pas le seul pays à devoir affronter un vieillissement de sa population. En effet, nous pouvons constater sur ce graphique que ce pourcentage de la population va doubler sur tous les continents. L'âge moyen va ainsi croître. La prochaine stabilisation de la population prévue par les statisticiens, et donc la prochaine évolution notable de notre population, ne se ferait pas avant l'an 2200 (avec une population mondiale qui serait alors d'environ 10 milliards d'habitants).

Ces chiffres nous montrent que, de par le monde, un nouveau problème va surgir : la dépendance.

1.1.3 Les solutions orientées personnes âgées : une obligation ?

1.1.3.1 Problèmes posés par l'évolution de la population

La population mondiale est donc en constant vieillissement. De nos jours, dans tous les pays industrialisés et développés, l'accès aux meilleurs soins et aux médicaments adéquats est aisé et se fait à un prix abordable. Cependant, il n'existe pas de traitement chimique pour

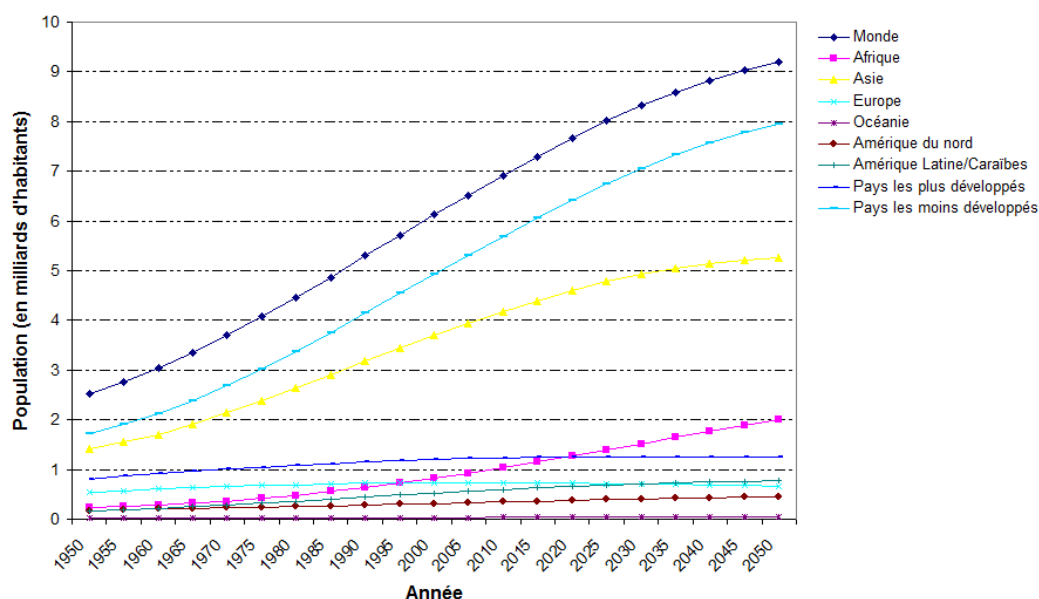


FIG. 1.4 – Répartition de la population mondiale (1950–2050)

certaines pathologies qui entraînent la dépendance de la personne. Pour une personne vivant seule, la famille doit ainsi faire face à ce problème et se rendre disponible pour prendre en charge son aîné.

Une estimation de l'ampleur de ce problème, donnée par les spécialistes, est qu'une personne sur quatre sera, en France, touchée par le problème de dépendance aux alentours des années 2030 (en considérant la dépendance liée à l'âge et aux maladies neurodégénératives, ainsi que la dépendance liée aux handicaps physiques ou mentaux).

Ces estimations modifient la vision sur la population et permettent de réagir au plus vite. Ainsi, en 2008, le plan Alzheimer a été lancé par le gouvernement français. Ce plan va ainsi débloquent des fonds sur la période 2008–2012, pour financer entre autres⁷ :

- un meilleur diagnostic, avec un renforcement des consultations pour une prise en charge plus précoce de la maladie, puis le renforcement du dispositif d'annonce et d'accompagnement des familles.
- un meilleur accompagnement du malade par l'intermédiaire du renforcement du soutien à domicile (avec du personnel formé), la création au sein des EHPAD⁸ d'unités adaptées à l'accueil de personnes souffrant de troubles comportementaux ainsi que la labellisation sur tout le territoire de « portes d'entrée uniques » : les « Maisons pour l'Autonomie et l'Intégration des malades Alzheimer » (MAIA).
- une meilleure aide, par l'intermédiaire du développement et de la diversification des structures de répit ainsi que la mise en place d'un numéro unique et d'un site Internet

⁷Le plan Alzheimer contient et détaille 44 mesures. Il est disponible à l'adresse suivante : http://www.premier-ministre.gouv.fr/IMG/pdf/Plan_Alzheimer_2008-2012.pdf

⁸Établissement d'Hébergement pour Personnes Âgées Dépendantes

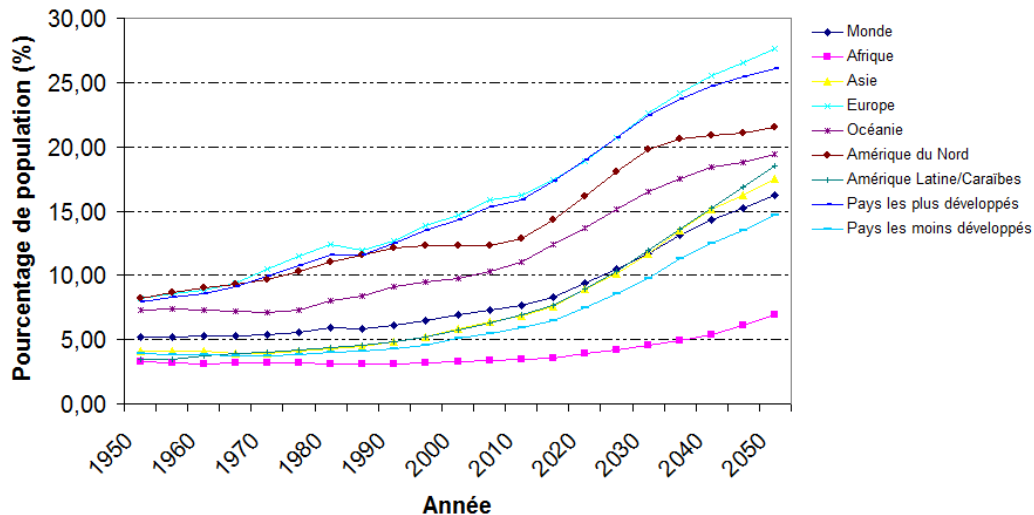


FIG. 1.5 – Pourcentage de la population âgé de plus de 65 ans (1950–2050)

d'informations et d'orientation locale

- une recherche plus rapide avec la création d'une fondation de coopération scientifique pour stimuler et coordonner les équipes autour de la maladie.

Enfin, à partir de l'automne 2008, une commission sera mise en place dans le but de créer une nouvelle branche de la sécurité sociale (la « 5^{ème} branche ») et ce afin que « la solidarité nationale prenne en charge la dépendance, liée au vieillissement comme au handicap »⁹.

1.1.3.2 Solutions proposées

La détection de l'entrée en dépendance se base sur les capacités d'une personne à exécuter seule les ADL (Activities of Daily Living, en français AVQ, Activités de la Vie Quotidienne), dont nous verrons les détails dans la suite de ces travaux.

À l'heure actuelle, une fois l'entrée en dépendance détectée, la prise en charge des personnes âgées peut se faire de plusieurs manières. La première est un soutien aux familles, qui peuvent s'occuper de la personne âgée, et pour qui des allocations existent, afin de leur permettre de rémunérer des aides soignantes ou des auxiliaires de vie qui se chargent de certaines tâches ménagères ainsi que d'effectuer les tâches d'hygiène de la personne. Ces aides financières dépendent des revenus de la personne et de sa famille. En fonction de ce que la famille proche est susceptible de pouvoir payer, l'aide sera donnée en conséquence afin de permettre une vie décente de la personne à son domicile.

Une solution d'accueil familial existe également pour permettre à la personne âgée de ne pas vivre seule à longueur de journée. Cet accueil est fait par des « familles d'accueil » rémunérées pour cela, qui offrent une partie de leur maison à une personne âgée. Ces familles

⁹Extrait d'un discours du ministre du travail, des relations sociales, de la famille et de la Solidarité, Xavier Bertrand

sont en général en contact avec des médecins, des infirmiers et des kinésithérapeutes passant régulièrement pour soigner les personnes accueillies. Ces familles sont agréées par le conseil général. Cette solution est une solution plus humaine et chaleureuse que l'entrée en institution pour les personnes dont la famille est éloignée ou bien pour celles dont la famille n'a pas le temps pour prodiguer l'attention et les soins nécessaires (par exemple une personne atteinte de la maladie d'Alzheimer demande à être constamment surveillée, il n'est donc pas possible pour une famille qui travaille de s'occuper de son aîné malade).

Des entreprises associatives se sont créées, constituées de pharmaciens, kinésithérapeutes, d'infirmiers et de médecins, pour organiser ce maintien à domicile, médicaliser la chambre du patient et s'occuper des soins que la personne doit subir.

Cependant, lorsque la personne ne peut plus être maintenue à domicile, que ce soit à cause d'une maladie neurodégénérative, d'un handicap sévère ou bien tout simplement parce qu'elle ne se sent plus en mesure de réaliser les activités de la vie quotidienne et qu'elle n'a plus personne autour d'elle pour l'épauler, il faut bien alors se tourner vers l'entrée en institution. Cette entrée se justifie par l'accueil et la présence dans l'institution de personnels formés et connaissant les pathologies qui peuvent être développées par la personne. Cependant, l'entrée en institution reste difficile du fait du nombre de places très limité actuellement disponible. En effet, la situation début 2007 révélait un manque estimé de 30000 à 40000 places. Le plan « Vieillesse et Solidarités » a permis de créer 15000 places dans la période 2004–2006 et 5000 en 2007, et le plan « Grand Âge » devrait permettre la création de 25000 places entre 2007 et 2012. Cependant, les personnes âgées fragilisées sont actuellement au nombre de 1,2 millions environ, chiffre qui devrait doubler dans les dix ans.

Parmi les problèmes qu'elles rencontrent, les institutions doivent également faire face à un manque de personnel. En effet, en France, il y a quatre personnes soignantes pour dix personnes âgées. C'est moitié moins que dans des pays tels que l'Allemagne et la Hollande, et cela ne permet pas d'apporter des soins optimaux pour les personnes. Par exemple, la toilette des personnes devrait prendre vingt minutes dans des conditions favorables mais, du fait du manque de personnel, elle doit être effectuée en dix minutes environ.

Enfin, le dernier problème des institutions de personnes âgées est leur coût de séjour. À cet égard, il convient de rapprocher le montant moyen des pensions de retraite, 1100 euros pour les hommes et 800 euros pour les femmes, du coût moyen d'un placement en institution : 1600 euros.

Cela montre que les solutions actuellement proposées doivent être repensées pour faire face à la croissance de la demande dans ce domaine. Comme nous le verrons dans le chapitre suivant, de nombreuses solutions sont en projet dans différents laboratoires afin de permettre un maintien à domicile dans les meilleures conditions. Les solutions pour le moment prisées par les équipes de recherche et certaines entreprises se basent sur la télé médecine et la téléassistance, entrant dans le cadre des gérontechnologies.

1.2 La Télémédecine et les Gêrontechnologies : les enjeux et les nécessités

1.2.1 Le concept de gêrontechnologie

Les gêrontechnologies, un concept apparu dans le milieu des années 1990 dans les pays nordiques, se trouvent au carrefour de deux disciplines, l'une médicale, la gériatrie, et l'autre, la gêrontologie, qui s'intéresse plus largement aux technologies pour la santé des personnes âgées. Le but de la recherche et du développement dans le domaine des gêrontechnologies est de mettre au point des techniques efficaces d'assistance à domicile et de conception de l'habitat pour permettre aux personnes âgées d'y vivre de manière autonome, et ce en bonne santé, dans le confort et la sécurité. À la rencontre des deux domaines qui la composent, le rôle de cette discipline est de mettre en concordance les technologies existantes avec les besoins croissants des personnes âgées.

Les applications proposées par ce domaine sont diverses. Elles concernent plusieurs points, avec, entre autres :

- la conception de bâtiments adaptés aux personnes âgées,
- la communication, au moyen de technologies permettant de rapprocher les personnes de leur entourage qui n'est pas toujours présent à leurs côtés,
- le travail de la mémoire,
- l'assistance à domicile (avec des détections de situations anormales dans le comportement de la personne),
- le suivi de paramètres physiques ou physiologiques.

Les gêrontechnologies se sont énormément développées du fait de la demande croissante des pays développés. Ainsi, ISG (International Society for Gerontechnology)¹⁰ édite une revue depuis 1998, et organise une conférence chaque année sur ce thème, qui permet de mettre en commun, en matière de conception de bâtiments et de technologies, les différents projets.

La France est également très bien représentée dans ce domaine depuis la création de la SFTAG (Société Française des Technologies pour l'Autonomie et la Gêrontologie). À noter une initiative originale du Centre de Gériatrie Sud (CGS) à Grenoble. Sous la direction du professeur Alain Franco, assisté de Vincent Rialle, une consultation en gêrontechnologies doit être mise en place. Elle aura pour but de cerner les besoins des patients en matière de technologie, et peut-être à terme de les « prescrire », au patient et à sa famille, comme c'est le cas pour les médicaments.

1.2.2 La télémédecine

La télémédecine est un domaine d'utilisation des télécommunications pour effectuer un acte médical à distance, pour suivre à domicile des patients venant de subir une intervention

¹⁰<http://www.gerontechnology.org/>

médicale ou encore pour assurer le maintien à domicile de personnes handicapées ou de personnes âgées.

C'est donc un domaine d'application des gérontechnologies dont l'objectif est alors de surveiller à distance des paramètres vitaux selon le type de pathologie. Les données recueillies sont traitées soit sur place, soit dans un centre de traitement, et les anomalies sont détectées au plus vite pour intervenir rapidement auprès du patient.

Un centre de traitement doit être disponible pour relever les alertes et prendre les décisions appropriées. Tout l'apport des technologies et du traitement du signal est de réduire considérablement les traitements effectués manuellement et de réduire au minimum le nombre des fausses détections ou des détections manquées. En effet, manquer une alerte pourrait avoir une incidence grave sur le patient, mais au contraire alerter abusivement a également un coût financier important.

Les travaux de cette thèse s'inscrivent dans cette optique. L'objectif des travaux de l'équipe AFIRM est de détecter, à distance, un changement dans le comportement de la personne, qui pourrait indiquer ainsi une perte ou une baisse de son autonomie.

1.2.3 Les technologies pour la santé : une nécessité ?

Les technologies pour la santé se développent très rapidement face à ces nouveaux marchés, fortement prometteurs du fait des prévisions de population décrites précédemment, qui attirent ainsi fortement les investisseurs et amènent la création d'entreprises dédiées. Si certaines des « technologies » devaient être prises en charge par les collectivités, ce marché se développerait encore plus rapidement.

Mais dans un monde avec une telle démographie, prendre en charge correctement l'ensemble des personnes âgées dans le besoin va se révéler être une tâche très difficile. Ainsi, les technologies d'assistance et de maintien à domicile vont devenir une aide considérable. Elles ne remplaceront bien sûr jamais un médecin spécialisé ni une équipe soignante formée pour prendre en charge un patient. Cependant, elles pourront, à terme, augmenter l'efficacité des médecins en leur permettant de focaliser leurs efforts sur les points importants, en les aidant à poser un diagnostic à partir de données quantitatives sur la situation actuelle du patient.

Aucune technologie n'est une nécessité, cependant elles deviennent, au fur et à mesure de leur développement et de leur perfection, une aide considérable qui augmente l'efficacité de la prise en charge du patient. Les gérontechnologies vont donc très certainement prendre une place importante dans la vie dans les années à venir. Ces technologies proches de celles employées par « Big Brother » peuvent effrayer les personnes âgées actuelles qui n'ont été confrontées que récemment avec l'apparition des technologies de l'information et de la communication. Notre génération, qui a grandi avec ces technologies, sera probablement beaucoup moins craintive envers l'utilisation des aides au maintien à domicile.

État de l'art

Sommaire

2.1 L'actimétrie embarquée	31
2.1.1 Les types de capteurs	31
2.1.2 Les centrales équipées de capteurs cinématiques	32
2.1.3 Applications	33
2.1.3.1 Analyse du mouvement et dépense énergétique	33
2.1.3.2 La détection de la chute	34
2.2 La reconnaissance sonore	35
2.2.1 La reconnaissance de la parole pour l'assistance aux personnes	35
2.2.2 Application de la reconnaissance des sons et de la parole à l'habitat et à la domotique	36
2.3 La domotique au service de la santé et de l'assistance à domicile	37
2.3.1 Les dispositifs de secours et de téléassistance	37
2.3.2 La localisation de la personne	38
2.3.2.1 Suivi à l'intérieur du domicile	38
2.3.2.2 Les risques de fugue	39
2.3.3 La détection de situations anormales	40
2.3.4 Les habitats intelligents	41
2.3.4.1 Surveillance et suivi de personnes	41
2.3.4.2 L'assistance à domicile	44
2.3.4.3 Le suivi d'activités	45
2.3.4.4 La télémédecine	46
2.4 Éthique et habitats intelligents	47
2.5 Conclusion sur cet état de l'art	48

CES dernières années ont vu l'essor de nombreux projets dans les domaines de la télémédecine et des technologies appliquées aux personnes âgées ou handicapées. Émanant de tous les continents, ils visent à proposer divers types de solutions qui pourraient être, dans les années à venir, applicables à grande échelle. Ce chapitre va tenter de faire un tour d'horizon des différents types de solutions proposées et un état de l'art des technologies et algorithmes associés.

2.1 L'actimétrie embarquée

2.1.1 Les types de capteurs

Ces dernières années, la miniaturisation des composants ainsi que les améliorations en matière de consommation ont poussé de nombreuses équipes de recherche à se tourner vers des solutions embarquées pour mesurer l'activité de la personne. Ainsi, les accéléromètres tri-axes se sont démocratisés et sont maintenant d'un prix très abordable et avec une consommation très faible (de l'ordre de $10\mu A$ en veille pour un temps de réveil très court de l'ordre de la milliseconde). De tels capteurs sont fabriqués par exemple par Freescale¹ (MMA7260Q), par Analog Device² (ADXL330), par ST Microelectronics³ (LIS3LV02DQ), et aussi par une société Finlandaise nommée VTI⁴ (SCA3100). Ces accéléromètres sont également de plus en plus utilisés dans les outils de tous les jours. Ainsi, la console Wii a démocratisé la détection de mouvements de la main pour les jeux vidéos, et certains PDA ou appareils photos intègrent de tels accéléromètres pour l'orientation de l'écran et adapter son affichage (en portrait ou paysage) en conséquence.

L'actimétrie embarquée peut également utiliser des magnétomètres (mono-axes, bi-axes ou tri-axes). Ceux-ci mesurent le champ magnétique perçu (donc par exemple le champ magnétique terrestre). Ces magnétomètres consomment encore beaucoup trop par rapport aux accéléromètres, et le matériau sensible doit parfois être « réinitialisé » à la suite de l'exposition à un champ fort, et ce en utilisant en général une impulsion de courant élevée pendant un temps très faible. Ce type de capteur est fabriqué par Honeywell (HMC1053 pour les tri-axes) ou bien par NXP (KMZ51/52 pour respectivement un mono-axe et un bi-axe). Leur prix reste encore relativement élevé par rapport aux accéléromètres mais leur apport est non négligeable. Ils tendent à se démocratiser du fait de leur utilisation dans certains modèles de GPS par exemple (à l'aide d'un magnétomètre mono-axe pour connaître l'orientation par rapport au nord magnétique).

Un troisième dispositif utilisable dans ce contexte est le gyroscope. Ce dernier délivre une mesure de la vitesse angulaire instantanée autour d'un axe. ST Microelectronics propose des références de produits (LISY300AL) ainsi qu'Analog Devices (ADXRS610). Cependant les gyroscopes à l'heure actuelle présentent trois inconvénients majeurs :

¹<http://www.freescale.com/webapp/sps/site/homepage.jsp?nodeId=011269>

²<http://www.analog.com/en/mems-and-sensors/products/index.html>

³<http://www.st.com/stonline/products/families/sensors/sensors.htm>

⁴<http://www.vti.fi/en/products/accelerometers/instruments/>



FIG. 2.1 – Centrales de mesures. De gauche à droite : Dynaport de Mc Roberts, Physilog du LMAM, Trident du CEA-LETI

- une très forte consommation : pour un gyroscope mono-axe, la consommation s'élève à environ $5mA$, ce qui est beaucoup trop important pour un système embarqué raisonnable.
- une dérive importante qui doit être régulièrement compensée, en général par d'autres capteurs tels que les magnétomètres ou même les accéléromètres.
- une intégration moins aisée que les autres capteurs. Pour le moment les produits proposés sont mono-axes. Certains fabricants proposent des boîtiers complets avec trois gyroscopes et un accéléromètre tri-axe, mais il n'existe pas, à l'heure actuelle, de gyroscope tri-axe intégré dans un circuit à monter en surface. Cela oblige à concevoir un boîtier.

Enfin, d'autres capteurs peuvent être envisagés lors de la mesure d'activités, tels que des capteurs de pressions (sous les semelles), des capteurs physiologiques (fréquence cardiaque, mesure de la température, résistivité de la peau...). Ils sont en général associés aux capteurs cinématiques pour améliorer la qualité de l'information délivrée par ces derniers.

2.1.2 Les centrales équipées de capteurs cinématiques

De nombreux projets de développement de centrales d'activités basés sur des capteurs cinématiques ont vu le jour depuis le milieu des années 1990, et ce dans des domaines d'applications divers allant de la rééducation fonctionnelle à l'analyse de la marche et des transferts et à la surveillance d'activités de la vie quotidienne.

La société McRoberts aux Pays-Bas développe des centrales appelées Dynaport équipées au choix d'un accéléromètre tri-axe, de trois gyroscopes, d'un accéléromètre et d'un gyroscope ou bien de deux accéléromètres et d'un gyroscope. Ces centrales, placées sur le sacrum, permettent l'étude de la marche et de la posture du patient. À partir des signaux des accéléromètres, le logiciel de traitement en déduit des variables donnant des informations sur la vitesse de la marche, l'asymétrie, les irrégularités, et la stabilité de la posture. Cette centrale a été validée pour la mesure d'activités sur des personnes saines [Bussmann *et al.*, 1998]. En utilisant ensuite des sujets adultes et sains comme référence, le système a donné des résultats sur des populations différentes tels que les sujets très jeunes [Brandes *et al.*, 2006] ou sur des sujets atteints d'arthrite [Brandes *et al.*, 2008], ou bien encore les sujets amputés

[Stam *et al.*, 1995, Bussmann *et al.*, 2004].

Le laboratoire LMAM de l'EPFL⁵ développe Physilog, un système portable de suivi à long terme de mouvements. Physilog est un dispositif d'enregistrement de signaux permettant la mesure du mouvement à partir d'un ou plusieurs capteurs cinématiques. Les capteurs cinématiques sont reliés à l'enregistreur et fixés sur le patient soit par l'intermédiaire d'une ceinture soit directement collés sur la peau. Les données sont ensuite enregistrées sur une carte mémoire. Les capteurs cinématiques utilisés sont des accéléromètres tri-axes et un ensemble de trois gyroscopes [Favre *et al.*, 2006a]. Cet enregistreur de signaux peut également être connecté à des capteurs de pressions plantaires, un signal cardiaque. . . Les applications proposées sont diverses. Le dispositif est utilisé pour caractériser les transferts posturaux [Najafi *et al.*, 2002], pour la caractérisation de mouvements d'articulations humaines complexes tels que le genou [Favre *et al.*, 2006b], l'analyse de la marche suite à la pose d'une prothèse de hanche [Jolles *et al.*, 2004] ou bien encore l'analyse des signaux de personnes atteintes de la maladie de Parkinson [Salarian *et al.*, 2007].

En France, le CEA-LETI⁶ et sa startup Movea⁷, développent une centrale d'attitude appelée TRIDENT contenant un magnétomètre tri-axe et un accéléromètre tri-axe. Movea développe également une montre (Motion Pod) permettant de surveiller les mouvements d'une articulation. Cette centrale était dans un premier temps utilisée dans le projet ACTIDOM (ACTImétrie à DOMicile), visant à surveiller l'activité des personnes âgées et fragiles à leur domicile [Bonnet *et al.*, 2004, Noury *et al.*, 2004]. Elle est également utilisée pour l'étude du mouvement humain (e.g. les séquences de marche, les transferts. . .) [Bonnet et Heliot, 2007] et la rééducation fonctionnelle de la marche [Héliot, 2007]. En fusionnant les données captées par plusieurs de ces centrales, les travaux permettent d'obtenir la caractérisation des mouvements de certaines articulations complexes [Salmeron-Quiraz, 2007].

2.1.3 Applications

2.1.3.1 Analyse du mouvement et dépense énergétique

Ainsi qu'il a été énuméré précédemment en présentant les travaux de différents laboratoires, l'analyse des mouvements humains est une partie importante de l'utilisation de ces capteurs cinématiques. En effet, de nombreux travaux portent sur l'analyse de mouvements de certaines parties du corps tels que les travaux de Mc Ghee [McGhee *et al.*, 2001, McGhee *et al.*, 2000]. Pour analyser ces mouvements et l'orientation des segments, des méthodes à base de filtres de Kalman ont été développées pour corriger les erreurs de mesure des différents capteurs et avoir une mesure de l'orientation aussi précise que possible [Marins *et al.*, 2001, Luinge et Veltink, 2004].

Les accéléromètres sont utilisés pour détecter différentes activités de la vie quotidienne : la marche ou les différents transferts posturaux (assis, debout et couché) [Bouten *et al.*, 1993, Veltink *et al.*, 1993, Barralon *et al.*, 2005, Lyons *et al.*, 2005, Scanail et al., 2006], chacun des

⁵Laboratory of Movement Analysis and Measurement, École Polytechnique Fédérale de Lausanne

⁶Commissariat à l'Énergie Atomique : <http://www-leti.cea.fr>

⁷<http://www.movea-tech.com/>



FIG. 2.2 – Détecteurs de chute, de gauche à droite : Le détecteur de chute du TIMC-IMAG, le téléphone d'Attendance, la montre Vivago.

auteurs se différenciant par l'utilisation de sa propre méthode de segmentation et de classification des données. Sont également classifiées par ce type de capteurs des activités plus complexes telles que le pédalage sur un vélo d'appartement, la descente et la montée d'escaliers [Muscillo *et al.*, 2007, Tapia *et al.*, 2007].

Lors de changements de posture, notre rythme cardiaque va se modifier. À partir de ce constat, [Chan *et al.*, 2006] analyse les périodes transitoires du signal cardiaque avec une transformée en ondelettes et fusionne ceci avec les données accélérométriques afin d'apporter plus de précisions aux changements de postures données par la cinématique.

Un autre type de calcul, effectué à partir des données cinématiques, est la dépense énergétique engendrée par les mouvements. Ce type de calculs est utilisé pour suivre la dépense énergétique de personnes saines, jeunes ou âgées, mais également de personnes souffrant de maladies pulmonaires chroniques [Steele *et al.*, 2003, Steele *et al.*, 2000].

2.1.3.2 La détection de la chute

La chute est un problème de santé publique important et pour le moment, aucun détecteur n'est fiable à 100% pour détecter tous les types de chutes. De nombreux détecteurs se basent sur des informations vidéos seules [Anderson *et al.*, 2008], sur une fusion informations audio/vidéo [Toreyin *et al.*, 2006] ou même utilisent la détection de vibrations du sol [Alwan *et al.*, 2006].

Cependant, les capteurs cinématiques peuvent être, et sont très largement, utilisés pour détecter la chute. Ainsi, de nombreux projets ont vu le jour. Un projet Européen, Profane⁸, rassemble de nombreux scientifiques et équipes de recherche autour de ce thème, avec pour objectif la détection et la prévention précoce des chutes.

De nombreux systèmes sont réalisés pour cette détection, à partir de capteurs cinématiques. Le TIMC-IMAG développe un capteur de chute basé sur deux accéléromètres bi-axes montés perpendiculairement [Noury, 2002]. Le capteur en lui-même a déjà subi de très nombreux tests et est en passe d'être industrialisé par l'entreprise Vigilio. De même, la société Attendance commercialise un téléphone portable (ATI-2500) détectant la perte de verticalité et alertant les secours ou la famille proche en cas de nécessité. Tout aussi intégrée, la

⁸<http://www.profane.eu.org/>

montre Vivago permet de détecter la chute ou les situations graves (perte de connaissance, baisse de température ambiante) et d'alerter les secours en cas de nécessité. La société NOC-Watch propose également FallSaver, un détecteur sous forme de patch⁹, capable de prévenir la chute et permettant de transmettre une alarme. Celui-ci a pour avantage de pouvoir être porté jour et nuit, sous la douche ou dans le bain. Il doit être remplacé tous les mois. Un récepteur d'alarme avec également un bouton permettant d'appeler un centre de téléassistance est fourni avec le patch.

Pour ce qui concerne les capteurs utilisés, les détecteurs de chute se basent sur des accéléromètres [Noury *et al.*, 2007, Bourke *et al.*, 2007b, Quagliarella *et al.*, 2008] ou bien sur des gyroscopes [Bourke et Lyons, 2008], et utilisent en général des seuils fixés par l'expérimentation pour isoler la chute à l'intérieur des différentes activités de la vie quotidienne (refaire son lacet, ramasser un objet, s'allonger) [Bourke *et al.*, 2007a].

Les causes de chute sont également activement recherchées par analyse de la marche et de la posture chez la personne âgée, afin de prévenir la chute en plus de la détecter, et de surveiller de plus près des personnes dites « à risque ». Dans ce cadre, un projet national PARACHUTE équipe le logement du patient afin de mesurer et de détecter une évolution dans la posture et la marche de la personne. Le but du projet est de mettre en place un protocole simple, s'adaptant à la personne et à son environnement, et permettant d'objectiver le risque de chute.

2.2 La reconnaissance sonore

2.2.1 La reconnaissance de la parole pour l'assistance aux personnes

La reconnaissance automatique de la parole est utilisée dans de nombreux domaines. Ainsi, des logiciels tels que Dragon Naturally Speaking font de la transcription de document sur informatique à partir de paroles énoncées à vitesse et voix normale dans un microphone. Cette technique nécessite cependant un apprentissage pour s'adapter à la personne, ce qui est fait à partir d'un corpus bien défini pour être représentatif du locuteur. Plus aisé, de nombreux serveurs vocaux (tels que ceux des répondants de certains opérateurs mobiles) reconnaissent uniquement certaines commandes, et cette fois-ci sans aucun apprentissage ; un apprentissage dans de tels cas serait trop compliqué à mettre en place et trop coûteux. Chacune de ces deux applications a son défi à surmonter. Dans le cas de la première, il est de reconnaître un vocabulaire extrêmement important (toute la langue française ou anglaise par exemple) avec des phrases de longueur variable. Pour la seconde, la reconnaissance se fait sur un nombre de mot très limité. Cependant, les conditions d'enregistrement du son à traiter peuvent être fortement perturbées. Par exemple, le serveur du répondant de mobile doit reconnaître le mot (qui désigne sa commande) alors que la personne utilise les transports en commun et est environnée par des bruits ambiants très forts. De même dans le cas de commandes vocales pour une voiture, la reconnaissance doit se faire aussi bien moteur

⁹<http://www.fallsaver.net/>

éteint qu'allumé.

Basée sur la reconnaissance d'un faible nombre de mots, France Telecom R&D a mis au point une écharpe communicante. Cette écharpe est utile en cas de problème pour les personnes âgées ou dépendantes (handicap physique par exemple). Cette écharpe permet d'appeler aisément (avec des commandes vocales ou tactiles) un correspondant de la famille ou les secours en cas de problèmes.

Concernant les personnes handicapées et/ou âgées, des travaux ont également démontré la faisabilité du contrôle d'une chaise roulante par l'intermédiaire d'un certain nombre de commandes vocales [Fezari et Bousbia-Salah, 2007]. Ce type de commandes utilise des moteurs de reconnaissance de parole existants qui sont adaptés pour cette utilisation. De même, l'École Nationale Supérieure des Télécommunications de Paris a travaillé sur un système possédant peu de commandes et s'adaptant au fur et à mesure à la voix de la personne [Renouard *et al.*, 2003]. Ce système possède une mémoire permettant de créer une classe de rejet.

Enfin, cette reconnaissance de parole peut aussi être utilisée pour faciliter l'accès aux nouvelles technologies aux personnes âgées. Ainsi, un projet vise à assister les personnes âgées peu habituées à la saisie au clavier, dans l'utilisation de l'ordinateur à l'aide de commandes vocales [Kumiko *et al.*, 2004]. [Anderson *et al.*, 1999] propose, quant à lui, la reconnaissance de la parole de la personne âgée et une application à la recherche de documents informatiques.

2.2.2 Application de la reconnaissance des sons et de la parole à l'habitat et à la domotique

Cette reconnaissance de la voix et des sons peut très bien s'intégrer dans l'habitat afin d'y faciliter la vie à l'intérieur. À l'intérieur de l'habitat, ce type de reconnaissance peut avoir deux utilisations différentes :

- La domotique.
- La reconnaissance d'une situation de détresse.

La première utilisation a été implémentée par [Wang *et al.*, 2008]. Ce système se base sur une classification des sons et une reconnaissance de la parole. Les deux permettent ensuite d'apporter une aide ou d'automatiser des tâches dans l'appartement de la personne. Ce système se base sur un ensemble de microphones intégrés au plafond de l'appartement. La classification se fait à base de Séparateurs à Vaste Marge à partir des coefficients MFCC pour les sons.

La seconde, reconnaissance des situations de détresses, est explorée par le laboratoire CLIPS¹⁰ [Istrate *et al.*, 2006a, Istrate *et al.*, 2006b, Vacher *et al.*, 2006]. Ce système de reconnaissance de sons de la vie courante et de paroles est destiné à permettre la reconnaissance des activités de la vie quotidienne ainsi que la détection des mots clés de détresse dans la parole. Les microphones intégrés dans l'habitat vont donner des informations sur les différents

¹⁰Communication Langagière et Interaction Personne Système devenu le Laboratoire d'Informatique de Grenoble, LIG

sons qui ont été produits ainsi qu'une transcription de l'ensemble des paroles énoncées. À partir de la détection de mots clés dans ces transcriptions, le système peut détecter des situations de détresse avec le retard introduit par le système de reconnaissance automatique de la parole. Ce système au complet étant utilisé dans ces travaux de thèses, il sera décrit de manière complète dans le chapitre 5.

L'ESIGETEL, équipe ANASON¹¹, se penche aussi sur ce problème de reconnaissance sonore dans un habitat, et ce pour des applications embarquées et à coût réduit (en utilisant des cartes audio du marché) [Medjahed *et al.*, 2008, Virone et Istrate, 2007]. L'application principale est de nouveau la télévigilance médicale des personnes âgées ou fragiles au domicile.

La reconnaissance automatique de la parole nécessite des modèles acoustiques (permettant de retrouver les phonèmes prononcés) adaptés à la situation. Les personnes âgées sont un public dont la voix a été modifiée avec le temps et leur manière de parler est différente de la nôtre. Ainsi, des modèles acoustiques ont été construits pour cette population cible et testés afin de montrer qu'il est utile de construire de tels modèles [Baba *et al.*, 2004].

Enfin, en complément, signalons que des recherches ont également été menées sur le dialogue pour les personnes âgées au sein d'un habitat [Takahashi *et al.*, 2003]. Ce type de systèmes s'occupe d'une part de la synthèse de la parole mais aussi de la détection de la parole et de la création d'un dialogue cohérent. Ce type de recherche trouve ses applications dans les robots de compagnie des personnes âgées.

2.3 La domotique au service de la santé et de l'assistance à domicile

2.3.1 Les dispositifs de secours et de téléassistance

À la suite du décès du conjoint, il est fréquent que des personnes âgées se retrouvent seules et isolées. En effet, les demandes croissantes de mobilité dans notre société actuelle amènent les personnes à s'éloigner de leur famille. Que faire en cas de problème grave, si la personne âgée n'est pas en état de se rendre jusqu'au téléphone, et si la famille ne viendra pas la visiter avant une voire deux semaines ? Pour répondre à ce type de problèmes, des systèmes de déclenchement d'alarmes reliés à des entreprises de téléassurances sont disponibles sur le marché. En Europe la société Intervox¹² et aux États-Unis la société AlertOne¹³ proposent ce type de services.

Ce système fonctionne de manière simple. Une base, simple d'utilisation et composée d'un nombre limité de boutons, est placée dans l'appartement (voir Figure 2.3). Cette base est reliée par radio (sur les fréquences dédiées aux applications médicales, bandes d'alarme sociale, en Europe 869,2 MHz), à divers médaillons permettant à la personne de déclencher

¹¹http://new-www.esigetel.fr/index.php?art_id=34

¹²<http://www.intervox.fr/docs/gammes/personnes/>

¹³<http://www.alert-1.com/>



FIG. 2.3 – Exemple de la base Quiatil+ d'Intervox, avec ses trois boutons simples d'utilisation : Urgence, appel d'un proche, appel des services à domicile

une alarme. Ces médaillons peuvent prendre plusieurs formes. La plus courante est un gros bouton poussoir à garder autour du cou (pendentif) ou du poignet (montre), mais il existe également des déclencheurs qui peuvent s'intégrer à l'environnement comme une poire ou une tirette de douche.

Ce type de produit se connecte également à des capteurs environnementaux. Ainsi, ils peuvent être reliés à des détecteurs de gaz, des détecteurs de fumée, ou des détecteurs de présence à infrarouges. Ces derniers sont utilisés principalement afin de détecter des situations de sous-activités ou sur-activités de la personne.

Lorsqu'une alarme est déclenchée (par l'intermédiaire de l'un des bouton ou par un capteur environnemental), la personne entre en contact avec le centre d'appel du service. Celui-ci est appelé directement par la base qui s'y connecte dans les plus brefs délais. En cas de problèmes graves, ce centre se charge de l'appel et de l'envoi des secours au domicile de la personne.

Ces boîtiers de communication sont une première étape rassurante pour l'entourage et pour la personne âgée ; tous deux étant rassurés par la réponse rapide du centre de téléassistance. Certaines initiatives tendent à augmenter le nombre de personnes utilisant ce type de services. Ainsi, le conseil général de l'Essonne finance, depuis 1999, des centres d'appel et de téléassistance (via la société GTS¹⁴) et en fait ainsi bénéficier 5000 personnes (chiffre de 2006).

2.3.2 La localisation de la personne

2.3.2.1 Suivi à l'intérieur du domicile

Afin d'aller un peu plus loin dans l'analyse des activités de la vie de la personne, des équipes de recherche se sont concentrées sur l'utilisation de capteurs infrarouges, vidéos ou ultrasonores pour la localisation et le suivi de la personne à l'intérieur de l'habitat.

Ainsi, [Helal *et al.*, 2003] utilise, dans « la maison de Matilda » (du nom de la personne âgée simulée à partir d'un robot ayant servi aux expérimentations), des capteurs ultrasonores pour un positionnement précis de la personne à l'intérieur du domicile. Celle-ci est

¹⁴<http://www.gts-teleassistance.com/fr/aboutus/homepage.htm>

équipée de deux récepteurs ultrasons sur l'épaule gauche et l'épaule droite, et des enceintes diffusant des ultrasons sont placées dans les quatre coins de l'appartement. Ce système permet une localisation très précise de la personne (de l'ordre de vingt centimètres).

Des solutions à base de GPS existent également [Samama *et al.*, 2005] et permettent une géolocalisation de la personne même en environnement intérieur. Ces solutions se font à base de répéteurs GPS placés à l'intérieur de la structure et permettant une synchronisation qui ne serait pas possible sur les satellites classiques. Cependant, de tels systèmes donnent toujours une précision insuffisante (environ 10 mètres).

Autre solution proposée, les détecteurs infrarouges. C'est le cas du laboratoire TIMC-IMAG de Grenoble avec son HIS dans lequel nous utilisons ces capteurs pour surveiller les endroits stratégiques de l'appartement (associés à des activités). Des statistiques ont été effectuées afin de montrer la pertinence de l'analyse de ces données sur le long terme [Noury *et al.*, 2006, LeBellego *et al.*, 2006]. Cependant, contrairement aux ultrasons, les capteurs utilisés ici ne permettent pas une localisation précise de la personne dans l'appartement. [Ricquebourg *et al.*, 2006], dans l'optique de résoudre ce problème, utilise des capteurs infrarouges pyroélectriques, émettant onze faisceaux (pour chaque capteur). En installant de tels capteurs à plusieurs endroits de l'appartement (donc en créant un effet de recouvrement des zones de détections des différents faisceaux émanant des différents capteurs), les chercheurs proposent un algorithme à base de théorie de Dempster-Shafer (pour tenir compte des incertitudes et des imprécisions des capteurs) qui permet un suivi précis des personnes dans l'appartement.

Enfin, des solutions à base de reconnaissance d'images vidéos sont proposées par Microsoft dans son environnement EasyLiving [Krumm *et al.*, 2000] et par le projet GERHOME [Zouba, 2006]. Ces solutions utilisent un ensemble de caméras disposées dans l'environnement à surveiller et permettent un suivi de personnes. Cette méthode est, par contre, beaucoup plus invasive que les précédentes, l'image étant quelque chose d'important pour les personnes concernées.

2.3.2.2 Les risques de fugue

L'errance et la fugue sont des problèmes récurrents chez les personnes âgées atteintes de la maladie d'Alzheimer. Lorsque la personne réussit à échapper au contrôle de son entourage, elle peut sortir et se perdre, ne sachant plus revenir chez elle.

Pour répondre à ce problème de fugue, la Locate Box permet de positionner une personne avec une grande précision, en utilisant un récepteur GPS de dernière génération. Cet appareil répond aux SMS envoyés par un membre de la famille par exemple, en donnant la position exacte (sous forme de coordonnées GPS). La balise est également équipée d'un bouton d'urgence pour appeler un proche en cas de besoin, en lui envoyant en même temps les coordonnées.

De même, Medical Mobile¹⁵ propose le bracelet Columba. Se présentant sous la forme d'une montre, ce bracelet est équipé d'un GSM et d'un GPS de dernière génération pour

¹⁵<http://fr.eu.medicalmobile.com/iiix/home/>



FIG. 2.4 – Outils de géolocalisation pour les personnes âgées. De gauche à droite : la Locate Box, le bracelet Columba et la balise Aloize

localiser avec précision la personne. Lors d'une fugue ou d'une errance, le porteur peut-être appelé en plus d'être localisé. Un bouton d'urgence est également disponible en cas de problème de la personne porteuse.

La société IGL-France¹⁶ a dessiné un petit boîtier contenant un GPS également. Celui-ci est utilisé par la famille pour localiser une personne errante en se connectant sur un site Internet ou en appelant un centre de télé-assistance disponible 24h/24-7j/7.

Ces quelques exemples nous montrent différentes applications permettant de surveiller à distance les personnes âgées hors de leur domicile. Ces applications se doivent d'être simples d'utilisation afin de convenir au plus grand nombre.

2.3.3 La détection de situations anormales

Toujours dans une optique d'assistance, de nombreux capteurs peuvent déceler des indications sur des changements de rythme de vie de la personne ou des changements dans ses habitudes. Ces capteurs sont souvent très simples mais proposent en général des solutions efficaces et rassurantes pour les personnes de l'entourage.

Ainsi, la société Japonaise Zojirushi Corporation¹⁷ propose une bouilloire intelligente, permettant de préparer et de garder au chaud l'eau du thé ou du café. Cette bouilloire garde en mémoire l'ensemble des informations sur son utilisation et les statistiques sur ces informations peuvent être consultées directement sur le site par les membres de la famille. De plus, une non utilisation de cette bouilloire pendant une longue période est une situation préoccupante (le thé étant une boisson très appréciée et « signe de vie » au Japon) qui est détectée et signalée à la famille.

Les rythmes de vie et la déviation par rapport aux rythmes habituels d'une personne peuvent être mis en évidence au moyen de capteurs très variés. Ainsi, au Japon également [Nakajima *et al.*, 2008] a travaillé sur les rythmes de l'activité « regarder la télévision ». En mesurant sur une certaine période cette activité chez des personnes âgées, il retrouve des « habitudes de vie » qui peuvent ensuite être surveillées et dont les écarts peuvent être signalés à

¹⁶<http://www.igl-france.com>

¹⁷<http://www.mimamori.net>

la famille. Pour cette dernière, les données sont également transmises avec des graphiques accessibles sur une page Internet.

En France, Edelia (filiale d'EDF) propose un service nommé « Tranquil'assistance »¹⁸. Ce service se base sur deux dispositifs. Le premier est un simple médaillon, qui permet d'alerter les secours en cas de danger (tels que les services présentés précédemment) et le second est un système de secours dans le cas de l'impossibilité d'utiliser le médaillon, qui se base sur les habitudes de consommation d'eau de la personne. Le compteur d'eau est relié à un boîtier électronique permettant d'envoyer les informations sur l'utilisation d'eau en permanence par ADSL. En cas d'inutilisation prolongée de l'eau sur des périodes de la journée correspondant aux habitudes de vie de la personne, le centre de téléassistance tente d'entrer en contact avec la personne et alerte les secours en cas de non succès.

En France également, utilisé dans des hôpitaux et des maisons de retraite du Maine-Et-Loire, « Life Assistance »¹⁹ est un coussin équipé de 120 capteurs de pression et d'un temporisateur le tout relié au système d'appel des patients dans les chambres d'hôpital. Ce coussin fonctionne très simplement. Il suffit de régler la temporisation voulue (de dix secondes à quarante minutes) pour que le personnel soignant soit appelé dès que la personne n'est plus en contact avec son matelas. Ce type de capteur est adapté aux personnes alitées et qui ne se lèvent que pour aller aux toilettes ou pour de très courtes sorties (d'où le réglage de la temporisation).

Ces quelques exemples montrent des utilisations de petits objets de l'habitat pour prévenir en cas de problème rencontré par une personne. Plus qu'un équipement, il est aussi possible d'équiper l'habitat en entier afin d'obtenir plus d'information et de produire des alarmes plus pertinentes.

2.3.4 Les habitats intelligents

Les projets d'habitats intelligents, c'est à dire des habitats équipés de capteurs pour le suivi d'activité ou bien l'assistance à la personne sont nombreux en France et dans le monde. La suite de cette section va présenter quelques exemples révélateurs de la philosophie des habitats intelligents et de leur utilisation.

2.3.4.1 Surveillance et suivi de personnes

Le projet House_n du MIT (Cambridge, Massachussets) [Intille, 2002, Intille, 2004] se base sur un appartement équipé de plusieurs centaines de capteurs. Ses domaines d'applications au niveau recherche sont le développement d'interfaces hommes-machines pour l'aide aux activités dans le domicile et le contrôle de l'environnement (domotique), l'aide à l'économie d'énergie, mais aussi des applications plus médicales telles que l'aide pour rester physiquement et mentalement actif et le maintien en bonne santé. La flexibilité de l'environnement permet également de lancer des recherches sur l'impact de nouveaux systèmes ou de nouveaux choix architecturaux sur les activités de la vie quotidienne. Pour réaliser des tests

¹⁸<http://www.tranquilassistance.fr/presentation-service.html>

¹⁹<http://www.ngv1.com/produits/bteu0101.pdf>

sur de longues périodes, des sujets volontaires vivent dans l'environnement pour une durée choisie, sans avoir, pendant ce temps, de contact avec les chercheurs. Les données sont collectées à distance. Les systèmes de décision de l'habitat peuvent également intervenir au près du sujet. En effet, celui-ci est équipé d'un assistant personnel (PDA) qui lui permet de recevoir des messages sur des actions à réaliser. Pour aider les chercheurs, des systèmes d'annotation semi-automatiques ont également été mis au point. Cet habitat compte plusieurs centaines de capteurs. Ceux-ci vont des capteurs environnementaux (température, hygrométrie... de chacune des pièces et de l'extérieur), aux contacteurs de porte, capteur d'utilisation d'eau sur les différents éléments, signatures électriques, capteurs accélérométriques...

Aux États-Unis toujours, mais cette fois-ci à l'Université de Géorgie (Georgia Institute of Technology), le projet « The Aware Home Research Initiative » [Abowd *et al.*, 2002] se base sur une maison de deux étages ayant pour objectif des recherches sur l'habitat pour toutes les tranches d'âge. Au rez-de-chaussée, la maison est réservée à la personne âgée et à son maintien à domicile, alors que le premier étage est réservé, quant à lui, à la famille avec les parents qui travaillent et un ou des enfants possédant des troubles de comportement ou mentaux. La maison est équipée de capteurs de mouvements, de capteurs environnementaux (température, ensoleillement), de caméras vidéos pour une détection plus précise de certains mouvements et de la chute mais aussi pour la reconnaissance d'activités [Moore et Essa, 2002] et l'aide à la mémoire à court terme [Tran et Mynatt, 2003] ainsi que des tags RFID²⁰ sur certains éléments, pour identifier l'utilisation de certains éléments. Des tags Bluetooth ont également été mis en place pour permettre de retrouver les objets perdus dans le cas de personnes malvoyantes et ce en utilisant leur téléphone portable. Il comporte également un système aidant la personne dans la prise de médicaments et un glucomètre enregistrant l'ensemble des mesures qui ont été prises pour un suivi longitudinal de la personne diabétique. Afin d'aider la personne dans son quotidien, un contrôle gestuel de certains appareils nécessitant par exemple une télécommande en temps normal a également été mis en place. Tout ceci permet d'explorer les méthodes d'aide au maintien à domicile de la personne. Pour maintenir la communication avec sa famille, un écran relie également les deux étages permettant une communication par l'intermédiaire de photographies souvenir ou autre. Enfin, cette maison permet également, au premier étage, d'élaborer des recherches sur l'autisme et l'aide qui peut être apportée tant à la famille qu'aux enseignants pour s'occuper des enfants atteints de cette pathologie [Kientz *et al.*, 2007].

En Corée, le projet Intelligent Sweet Home [Bien *et al.*, 2002] est un appartement équipé pour accueillir des personnes âgées ou handicapées. Ce projet vise à les aider dans leur vie quotidienne à l'aide d'un lit automatisé et robotisé. Le lit est automatique et associé à un bras de robot et à un ensemble de caméras pour la détection de mouvements. Cette détection est utile pour un système nommé « SoftRemocon » qui permet de commander à distance le téléviseur ou la lumière par exemple. Ce système se destine plutôt aux personnes alitées et qui ne peuvent utiliser les éléments de la vie quotidienne de manière normale (qu'elles

²⁰Radio Frequency Identification

soient âgées ou handicapées).

Au Japon, le projet « Ubiquitous Home » [Yamazaki, 2007] rassemble, dans un même appartement, un ensemble de caméras et de microphones, des écrans LCD au mur, des capteurs de pression au sol, des détecteurs infrarouges, des accéléromètres et des capteurs de vibration posés sur les quatre coins du lit (pour détecter plus précisément les mouvements de la personne lorsqu'elle est allongée) ainsi qu'un petit robot capable de reconnaître les visages enregistrés à partir de sa caméra embarquée et équipé d'un microphone et d'un haut parleur pour la communication. Deux systèmes à base de RFID, l'un actif dans le plafond de chaque pièce et l'autre passif à l'intérieur des murs, permettent de détecter l'entrée dans une pièce d'un tag RFID. Le projet vise à implémenter des services à la personne qui s'aident du contexte (donné par les différents capteurs) pour prendre une décision et apporter une aide. Pour l'heure, les services implémentés sont les suivants : recommander un programme TV à la personne qui désire la regarder, en se basant sur ses goûts télévisuels et sur les programmes précédemment regardés ainsi que sur le programme TV actuel ; sélectionner et afficher la recette de cuisine lorsque le résident énonce un mot de cuisine ; une assistance lorsque le résident sort de l'appartement pour une activité précise, qu'il doit emmener des affaires spécifiques préalablement listées, le système se base sur les tags RFID présents dans l'appartement pour rappeler au résident ce qu'il doit apporter avec lui.

En France, le premier grand projet sur le sujet a été le projet TISSAD²¹ lancé en 1999 [Thomesse *et al.*, 2001]. Ce projet consistait en la modélisation d'un habitat intelligent capable d'accueillir les personnes âgées, les insuffisants rénaux et les insuffisants cardiaques.

Le projet GARDIEN²² [Steenkeste *et al.*, 2001] de l'INSERM à Toulouse a ensuite été mis en place pour surveiller les patients dans des institutions gériatriques. Le but de ce projet était la surveillance des déplacements des personnes âgées dans une chambre d'une institution équipée de détecteurs infrarouges. Ce système se déclenche automatiquement de 22h à 8h.

Le projet AILISA²³ a également équipé le logement de plusieurs personnes (en chambre d'hôpital et au domicile) et ce pendant une très longue période avec des capteurs infrarouges afin de suivre l'évolution des rythmes d'activité de la personne. Un système de mail est utilisé pour envoyer des situations d'urgence détectées ou des rapports sur l'activité de la personne durant la dernière période [Noury *et al.*, 2006].

Les projets PROSAFE²⁴ et ERGDOM²⁵ sont tous deux développés au LAAS²⁶ à Toulouse. Le premier a pour objectif le maintien de personnes âgées à leur domicile, avec détection des situations d'urgence et envoi d'alarme automatique. Le système, à base de détecteurs infrarouges, calcule les informations sur le niveau d'activité de la personne et repère des situations de détresse par des changements dans ces niveaux d'activités [Bonhomme *et al.*, 2008].

²¹ Technologies de l'Information Intégrées aux Services des Soins À Domicile

²² Gériatrie Assistée par la Recherche et le Diagnostic des Incidents et des Errances Nocturnes

²³ Plateformes d'évaluations pour des technologies de télésurveillance médicale et d'assistance en gériatrie

²⁴ Système d'aide au suivi des personnes âgées

²⁵ ERGonomie du DOMicile : <http://www.laas.fr/MIS/ERGDOM/Webergdom.htm>

²⁶ Laboratoire d'Analyse d'Architecture des Systèmes

Le système ERGDOM, quant à lui, est un gestionnaire d'ambiance et de confort dans la maison. Il se base sur les données de détecteurs infrarouges pour déterminer les habitudes de vie de la personne et adapte la température de chaque pièce en conséquence, dans un but de confort et d'économie d'énergie. En fonction des réglages sur le souhait de la personne ainsi que sur le modèle thermique de l'habitat, le système s'adapte et prévoit à l'avance le chauffage à mettre en place.

Le projet GERHOME²⁷, de l'Université de Nice Sophia-Antipolis, vise également à développer un système de suivi de la personne à domicile et de détection des activités de la vie quotidienne, pour soumettre des rapports réguliers à des équipes médicales. Le système est un appartement de deux pièces équipé de plusieurs types de capteurs : des capteurs de présence volumétriques, des capteurs d'ambiance (température, humidité, luminosité), des capteurs d'occupation sur certaines parties de l'habitat (lit, chaise, canapé...), des capteurs de consommation d'eau et de consommation électrique, des contacteurs de porte sur les fenêtres, les meubles, et certaines portes de communication et enfin des caméras vidéos pour de l'analyse d'image et de la détection d'activité.

2.3.4.2 L'assistance à domicile

Le Swedish Handicap Institute a mis en place le projet d'appartement intelligent SmartBo [Elger et Furugren, 1998], d'une surface de $78m^2$, adapté aux handicapés moteurs, visuels et auditifs ainsi qu'aux personnes démentes. Basé sur un système de reconnaissance de parole, un ensemble d'écrans tactiles, un synthétiseur de parole, ainsi qu'un ensemble de capteurs environnementaux capables de détecter des situations d'urgence. L'habitat a également été complètement dessiné et équipé afin qu'un ensemble d'automatisme soit capable d'aider les personnes handicapées ou âgées. Ainsi, des commandes à distance ou des commandes automatiques permettent l'allumage de lumière, la fermeture et l'ouverture de volets... En cas de situation d'urgence détectée, des signaux compréhensibles par tout type de population (sourde et aveugle entre autre) sont envoyés à l'habitant afin qu'il remédie au problème.

Au Québec, le laboratoire Domus de l'Université de Sherbrooke a également équipé un appartement complet dans le but de l'assistance à domicile des personnes à troubles cognitifs. Le laboratoire a créé un habitat intelligent, contenant des capteurs environnementaux, des caméras vidéo, des tags RFID et des contacteurs. Avec des retours sur des dispositifs mobiles ou dans l'appartement, le laboratoire vise à aider les personnes ayant des problèmes cognitifs à réaliser les activités de la vie quotidienne. Pour atteindre les objectifs définis, les travaux ont mené à effectuer une modélisation des processus cognitifs durant les activités [Dion et Pigot, 2007, Serna *et al.*, 2007] et développé des systèmes d'assistance dans la tâche [Bouchard *et al.*, 2006].

Le projet ENABLE [Hagen *et al.*, 2001, Hagen *et al.*, 2003] s'est, quant à lui, spécialement intéressé aux malades d'Alzheimer (et aux personnes démentes) et s'est concentré sur l'intérêt des technologies dans ce cas. Ils se sont tournés vers des solutions de type calendrier électronique permettant de rappeler des choses aux patients, des solutions pour retrouver

²⁷<http://gerhome.cstb.fr/>

des objets perdus, des téléphones affichant les photos des personnes en plus de leur numéro de téléphone pour en faciliter l'usage, ou bien des systèmes de gestion de l'allumage et de l'extinction automatique des lumières.

2.3.4.3 Le suivi d'activités

Le projet MavHome [Cook *et al.*, 2003, Das *et al.*, 2002, Youngblood *et al.*, 2005] de l'Université d'État de Washington et de l'Université du Texas se base sur des modèles de Markov des tâches afin de reconnaître les activités de la vie quotidienne et d'agir comme un agent intelligent. Son architecture en plusieurs couches prend en compte des capteurs de signature électrique et tente de prédire les activités qui vont être réalisées par l'habitant afin de lui venir en aide et réduire les « coûts d'entretien » de la maison.

À Portland, le projet Elite Care [Adami *et al.*, 2003, Adami *et al.*, 2005] s'occupe de surveiller à distance, à l'aide de capteurs de pression, les heures de coucher et de lever de personnes atteintes de la maladie d'Alzheimer. Ceci permet d'adapter les soins et de voir des changements dans le comportement de la personne.

En se basant sur les données des infrarouges du HIS du TIMC, le GIPSA-Lab a mis au point des algorithmes à base de chaînes de Markov pour la reconnaissance des activités de la vie quotidienne et l'évolution dans leur réalisation [Le *et al.*, 2007].

Aux Pays-Bas, le projet CARE [Kröse *et al.*, 2008] utilise un habitat intelligent équipé de contacteurs, capteurs d'immersion, capteurs de température et enfin capteurs d'humidité, placés à différents endroits de l'habitat. Il reconnaît, après un apprentissage à base de Modèles de Markov Cachés, les activités de la vie quotidienne d'une personne, afin de permettre le maintien à domicile des personnes. Pour vérifier les résultats, il a été demandé aux sujets d'annoter leurs actions. Les résultats présentés concernent pour l'instant la sortie de l'appartement et aller aux toilettes.

Au Royaume-Uni, [Hong *et al.*, 2008] équipe l'appartement de divers capteurs, principalement sur les objets (réfrigérateur, porte de placard, ...) mais aussi sur les condiments (lait, jus d'orange, ...) et se base sur l'utilisation de ces divers objets pour créer un modèle de deux activités : « se préparer une boisson chaude ou une boisson froide » et se préparer (se laver, se brosser les dents, se coiffer). La fusion des données suivant la théorie de l'évidence de Dempster-Shafer permet de créer les modèles d'activité en fonction de l'utilisation des différents éléments. Aux États-Unis, [Philipose *et al.*, 2004] utilisait des tags RFID sur l'ensemble des objets pour effectuer le même type de reconnaissance.

Enfin, au Japon, [Tsukamoto *et al.*, 2008] se base sur les signatures électriques (en mettant des capteurs sur la télévision, un chauffage et un four à micro-ondes) pour déterminer ces activités de la vie quotidienne à l'intérieur de la maison. Cette étude est préliminaire et n'inclue que très peu d'appareils électriques. En effectuant cette étude sur plusieurs semaines, des modèles statistiques d'utilisation et une vérification des bonnes détections des signatures ont pu être réalisés. Orange a travaillé également sur ce sujet, avec un système plus complet de détection des signatures électriques. Cette fois, toutes les lampes électriques étaient surveillées, ainsi que les chauffages et des équipements électriques de la cuisine. Les

résultats détaillés de cette étude sont présentés dans [Berenguer *et al.*, 2008].

2.3.4.4 La télémedecine

Les habitats et leurs équipements peuvent également être utilisés dans le cadre d'une hospitalisation à domicile, et ce pour diminuer le temps passé par le malade à l'hôpital. Ceci a deux conséquences. La première est une diminution des coûts et la seconde est une meilleure remise en forme du patient, étant donné que chez lui, celui-ci peut adopter son propre rythme et n'est donc pas obligé de se soumettre à celui de l'institution dans laquelle il est accueilli.

Ainsi, entre 1999 et 2002, s'est déroulé le projet DIATELIC. Il avait pour but de développer un système permettant la surveillance de patients dialysés traités à leur domicile par une dialyse péritonéale continue ambulatoire. Dans ce projet, une architecture client-serveur a été développée, avec un serveur dédié à l'analyse et au stockage des données des différents patients et des clients chez chacun des patients et chez les différents médecins participant à l'étude. Chaque jour, chaque patient entrait ses données dans une fiche informatisée qui remplace la fiche papier habituelle. Ces deux fiches ont une présentation similaire, de façon à permettre une adaptation aisée des patients et des médecins au système. Un module de diagnostic a également été conçu, à base de Chaînes de Markov Cachées, prenant en compte les différentes données du patient. Ce système évalue ensuite la condition du patient par des calculs de probabilité d'appartenance à une classe de patient. À la suite des bons résultats de cette expérimentation, des brevets ont été déposés et une société créée²⁸ afin de proposer les services de ces travaux de recherche.

Un peu plus tard, entre 2005 et 2007, Medical Care Continuity (MCC)²⁹ est un projet lancé en Europe et en expérimentation dans quatre pays différents, afin de permettre à des patients soignés pour des cancers d'être médicalisés à domicile. Ce projet permet au patient d'être suivi à son domicile grâce à une station réunissant une caméra télé-pilotable et un écran, établissant le lien avec son centre d'assistance ou la structure hospitalière. Le patient peut activer ce système à l'aide d'une télécommande à bouton unique. Il peut alors appeler ou accepter une communication et entrer ainsi en contact vocal puis visuel. En cas de problème, l'appel peut être transféré vers le médecin du centre d'assistance, puis vers le centre hospitalier si nécessaire via Internet (ADSL). Un médecin peut ainsi s'entretenir avec le patient pour suivre l'évolution de son état de santé, visionner son environnement en télé-pilotant la caméra. Afin de voir des détails et de faire un diagnostic à distance, le médecin a la possibilité d'agrandir l'image sur une zone précise du malade ou sur les équipements médicaux à proximité, et peut également prendre des clichés et les imprimer.

Le projet CareLab, de Philips (Pays-Bas)³⁰ est un programme de recherche qui vise à amener les éléments permettant la télémedecine pour les personnes âgées et les personnes malades chroniques dans l'appartement. Par exemple, il a été développé le lit smartBed, qui per-

²⁸<http://www.diatelic.fr/>

²⁹http://www.eten-mcc.org/images/CP_MCC_VDEF_VA.pdf

³⁰<http://www.research.philips.com/newscenter/archive/2006/061025-experiencelab.html>

met de mesurer différents paramètres physiologiques de la personne (fréquence cardiaque, respiratoire, ...) pendant le sommeil, pour s'apercevoir de possible troubles chez la personne et ainsi la soigner et augmenter sa qualité de vie.

Un dernier exemple qui peut être cité, et né d'un projet Européen, est le projet Mobil-Health³¹. Ce projet a permis de créer des dispositifs de mesure embarqués de plusieurs paramètres physiologiques (pression artérielle, rythme cardiaque, ...). Ensuite, ce système transfère par liaison sans fil (à l'aide d'un GSM) les paramètres mesurés au prescripteur des mesures (médecin, hôpital). Ceci permet aux personnes nécessitant seulement une surveillance de certains paramètres physiologiques à court ou long terme de faire cette mesure à leur domicile et d'y vivre normalement. D'autres projets en France (gilet VisuResp de RBI, VTAMN de TAM-Telesanté) sont des gilets ou des tee-shirts possédant des capteurs intégrés et permettant de suivre les paramètres physiologiques, et si besoin de prévenir le médecin en cas de besoin (par GSM).

2.4 Éthique et habitats intelligents

Ces nombreux projets posent des questions d'éthique. En effet, de nombreuses données sensibles sortent de l'appartement, et des images de la personne peuvent aussi transiter par différents réseaux. Certaines études, par l'intermédiaire de questionnaires sur la population cible, se sont intéressées à ces problèmes éthiques.

[Beaudin *et al.*, 2006] a réalisé des questionnaires à destination des personnes âgées et des personnels de soins, à propos du suivi à court et long termes de paramètres et d'activités dans un habitat intelligent. Il ressort de cette étude une grande variabilité en fonction des gens sur les données à recueillir à l'intérieur de l'appartement ou en embarqué. De plus, ces paramètres changent en fonction de l'âge mais aussi des pathologies des différentes personnes questionnées. Ainsi, les conclusions de cette étude préconisent des systèmes de suivi médical longitudinal personnalisables et qui s'adaptent aux demandes des personnes, qui doivent requérir le moins d'interaction possible entre le système et la personne suivie.

[Rialle, 2007] présente les résultats d'une étude faite avec l'association départementale France Alzheimer de Seine-Saint-Denis, et envoyée à de nombreuses familles d'aidants. Cette étude les questionne sur leurs souhaits en technologie et sur les problèmes que posent cette maladie et ces recherches pour la famille et les aidants. Il ressort de cette étude que pour le moment les aidants, touchés par cette maladie qui affaiblit leur proche, ont des opinions très tranchées et font ainsi apparaître deux populations, dites technophiles et technophobes, très distinctes. La première apprécie certaines technologies telles que les systèmes de géolocalisation, très utiles pour prévenir les fugues, mais aussi par exemple la visiophonie de lien social et la visiophonie de vigilance mobile privée. Au contraire, certaines technologies sont fortement rejetées par les personnes répondant aux questionnaires. Ainsi, une machine donnant des conseils oraux pour réaliser des recettes de cuisine est la plus rejetée suivie par le capteur de chute inséré sous la peau. Une dernière technologie proposée, le robot de com-

³¹<http://www.mobihealth.org/>

pagnie, est quant à lui rejeté par les aidants plus âgés et accepté par les aidants plus jeunes. Outre cet exemple, cette étude fait surtout ressortir le fait que la technologie doit vraiment s'adapter à la pathologie pour être acceptée, et que les technologies proposées pour l'instant pour la maladie d'Alzheimer sont difficiles à faire accepter.

[Rumeau *et al.*, 2006] investigue, également par l'intermédiaire d'un questionnaire, l'acceptabilité de systèmes à base de capteurs de présence infrarouges tel que celui proposé dans le HIS de Grenoble. Les résultats de cette étude font ressortir un risque modéré sur l'intimité pour un besoin perçu important et un usage potentiel pour un proche comme pour soi-même quasi-unanime. Cependant, plus de résultats présentés aux personnes sur la pertinence des données obtenues permettraient de les convaincre d'autant plus.

2.5 Conclusion sur cet état de l'art

Ce rapide état de l'art sur les technologies d'information et de la communication appliquées aux habitats des personnes âgées ou dépendantes montre que le sujet de recherche est en plein développement depuis quelques années. Ces technologies s'intéressent autant au vieillissement et au maintien à domicile des personnes qu'à l'hospitalisation à domicile, pour permettre aux malades de réintégrer plus rapidement leur appartement et ainsi se retrouver dans un environnement à l'intérieur duquel ils se sentent bien. Ces habitats font intervenir des connaissances technologiques bien sûr, pour l'installation des capteurs et le traitement des données, mais aussi des connaissances médicales, pour l'interprétation des données et des connaissances sur les modèles cognitifs pour comprendre la manière dont les personnes ayant des troubles peuvent réaliser des activités.

Nous allons maintenant présenter les objectifs de ce travail de thèse et le plan de ce manuscrit.

Objectifs des travaux de thèse

Sommaire

3.1	Historique des travaux du laboratoire	51
3.2	Nécessité d'une approche multi-paramétrique	52
3.3	Choix des capteurs	52
3.4	Les Activités de la vie Quotidienne (critères ADL)	53
3.4.1	Les ADLs de Katz	53
3.4.2	Les Activités Instrumentales de la vie quotidiennes (IADLs)	53
3.4.3	La grille AGGIR	55
3.4.4	Objectivation des critères	57
3.5	Organisation du manuscrit	57

3.1 Historique des travaux du laboratoire

Ces travaux s'inscrivent dans la continuité de thèses précédemment réalisées au laboratoire par Gilles Virone [Virone, 2003], Florence Duchêne [Duchêne, 2004] et enfin par Pierre Barralon [Barralon, 2005].

La première thèse, après avoir décrit le HIS (Habitat Intelligent pour la Santé) grenoblois, porte sur l'apprentissage et la surveillance multi-paramétrique des habitudes de vie d'un patient suivi à domicile, à travers le calcul des Rythmes Circadiens d'Activités, pour détecter une évolution anormale dans le comportement de la personne suivie. Lors de ces travaux, un simulateur a été développé pour mettre en place et vérifier la cohérence de cet indicateur.

Les seconds travaux concernent la conception d'un assistant intelligent pour analyser les données hétérogènes issues des capteurs du domicile pour détecter et prévenir toute situation à risque. À l'aide d'un système d'apprentissage des habitudes de vie de la personne, une mesure de l'écart à ce modèle est calculée pour classer une situation comme critique si la mesure est trop élevée. Ces travaux portent sur la mise en place d'un simulateur de données, puis des méthodes génériques d'extraction non supervisées de motifs temporels, pour expérimenter ensuite sur l'identification de comportements récurrents d'une personne suivie.

Enfin, les derniers travaux portent sur l'actimétrie embarquée de la personne et sur la fusion de données. À partir des données des accéléromètres délivrées par le capteur de chute, des algorithmes ont été développés et testés, sur des personnes jeunes et âgées, pour détecter les séquences de marche ainsi que les transferts « Assis-Debout » et « Debout-Assis ». Ensuite, ces données, ainsi que celles des détecteurs infrarouges et enfin le temps, ont été utilisés dans un modèle de type bayésien pour déterminer quelle activité de la vie quotidienne est en cours de réalisation.

Les travaux de thèse présentés dans ce manuscrit se situent dans le prolongement des travaux présentés ci-dessus. Les objectifs de cette thèse sont les suivants :

- Développer un système d'acquisition constitué d'un magnétomètre tri-axe et d'un accéléromètre tri-axe « ACTIM6D » ainsi que les algorithmes associés, pour détecter les différentes postures (assis, debout, couché) de la personne et avoir une idée de son niveau d'activité à chaque instant. Cette partie du travail sera décrite dans le chapitre 4.
- Tester et valider le système de reconnaissance de sons et de paroles développé par Hubert Glasson [Glasson, 2008] qui fait suite au système de détection et reconnaissance de sons de la vie courante étudié par Dan Istrate [Istrate, 2003]. Le chapitre 5 revient sur cette partie du travail.
- Développer un algorithme de classification supervisé pour les activités de la vie quotidienne d'une personne. Les chapitres 6 à 8 présentent les différentes parties, allant du choix de structuration des données au test en situation expérimentale de l'algorithme retenu.

Pour précéder cet ensemble, un état de l'art des Habitats Intelligents (ainsi que de l'uti-

lisation du son dans ces domaines), de l'actimétrie embarquée a déjà été présenté dans le chapitre 2. Ce manuscrit se terminera par une discussion générale rappelant, synthétisant et débattant des résultats obtenus et enfin permettra d'ouvrir sur la suite de ces travaux.

3.2 Nécessité d'une approche multi-paramétrique

Notre but est de classifier chaque fenêtre temporelle de la journée comme partie d'une activité de la vie quotidienne. Or, il n'existe aucun capteur permettant de détecter une activité de la vie quotidienne actuellement effectuée. Un accéléromètre, qui nous donnera l'activité et la posture de la personne, ne sera pas suffisant. De même un capteur infrarouge nous donnera la pièce dans laquelle elle se trouve mais il n'existe pas de bijection entre les pièces et les activités.

Ainsi, le laboratoire se base depuis quelques années sur une approche multimodale permettant de résoudre ce problème. Comme l'indique la définition de la fusion de données, un ensemble de capteurs différents nous donneront une classification plus précise qu'un seul capteur isolé, pour peu que nous utilisions les capteurs adaptés d'une part, et que d'autre part les données soient bien interprétées avant d'être apprises.

Ce travail s'est donc inscrit dans cette logique, à savoir que nous allons utiliser un ensemble de capteurs non-invasifs placés dans l'appartement pour tenter d'améliorer la décision qui pourrait être prise à l'aide d'un seul capteur et ainsi avoir une définition que nous espérons plus précise des différentes activités qui sont faites dans la journée par la personne.

3.3 Choix des capteurs

La classification à partir de capteurs hétérogènes permet d'obtenir de bons résultats dans le cas où un ensemble de trois conditions sont respectées :

- Un nombre suffisant de capteurs sont utilisés et le choix est fait de capteurs appropriés aux classes recherchées (décrivant suffisamment les classes).
- Un formatage correct des données (si celles-ci ne sont pas utilisables brutes) est effectué précédemment à l'apprentissage.
- Choisir un algorithme de classification et une règle de décision adaptés au problème.

Les deux derniers points seront développés plus précisément dans les chapitres suivants. En ce qui concerne le premier point, nous pouvons déjà, dans cette introduction, justifier quelques-uns de nos choix.

Notre but était de faire un système non-invasif, très peu encombrant, à prix limité et également le moins contraignant possible pour la personne. Nous nous sommes donc tournés vers un certain de type de capteurs :

- Un capteur embarqué sur le sujet, qui s'intègre à son tee-shirt de par sa taille réduite et qui est totalement autonome (fonctionnement à base de piles et de carte mémoire). Ce capteur va nous permettre de connaître la posture actuelle de la personne ainsi que d'avoir une idée de son taux d'activité.

- Un ensemble de microphones reliés à un système de classification des sons de la vie quotidienne et de reconnaissance de la parole. Ce système permet ainsi d’avoir des précisions dans le cas de certaines activités, par exemple de savoir qu’en un moment précis il a été détecté des sons de vaisselle dans la cuisine.
- Des capteurs de présence infrarouge donnant la pièce dans laquelle se trouve actuellement la personne. Cela réduit le nombre d’activités possibles à un moment donné.
- Des contacteurs de porte sur certains placards de la cuisine et de la chambre, ainsi qu’un contacteur sur la porte de l’entrée pour savoir quand la personne sort ou entre dans l’appartement.
- Une station météorologique donnant température et hygrométrie dans la salle de bain. Lors de la prise d’une douche par exemple, ces données vont fortement être modifiées.

Des caméras (webcams grand angles avec une ouverture d’environ 85°) ont également été ajoutées dans l’appartement. Cependant, ces caméras ne sont utilisées que pour indexer les activités réalisées. Les images ne sont pas traitées en tant que telles, du fait du caractère invasif de ce type de modalité.

Nous reviendrons plus en détails, dans le chapitre 6 sur l’installation de l’appartement et sur son organisation.

3.4 Les Activités de la vie Quotidienne (critères ADL)

3.4.1 Les ADLs de Katz

Les Activités de la Vie Quotidienne (AVQ, de l’anglais Activities of Daily Living – ADL) sont un ensemble d’activités basiques que l’on effectue « sans même y penser » dans le cadre de notre vie de tous les jours, et que nous effectuons dans le but d’avoir une qualité de vie correcte et de survivre. Ces activités incluent le fait de se laver, de se nourrir, d’aller aux toilettes...

Une échelle d’évaluation est utilisée par les professionnels de santé (gériatres entre autre) pour décrire l’aptitude ou l’inaptitude à s’assumer soi-même. Elle peut être utilisée dans la surveillance des personnes âgées, de la démence (maladies de type Alzheimer), dans le cadre de personnes souffrant d’une maladie chronique ou bien d’un handicap, et ce dans le but d’évaluer le type et la quantité de service médicaux individuels nécessités par une personne.

Les critères ADL présentés par Katz [Katz et Akpom, 1976] ont permis la création de plusieurs échelles toutes plus ou moins équivalentes et conformes aux définitions ci-dessus. Le tableau 3.1 présente l’une d’entre elle ainsi que les scores associés. Dans le cas de cette échelle, un score supérieur à 6 pour un patient montre une dépendance.

3.4.2 Les Activités Instrumentales de la vie quotidiennes (IADLs)

Une autre échelle, ayant les mêmes champs d’application, prend quant à elle en compte les instruments de la vie quotidienne et la capacité de la personne à les utiliser. Elle se base sur les IADLs définis par Lawton [Lawton et Brody, 1969]. Cette échelle est présentée dans le

Activité	Description	Score
Hygiène corporelle	Autonome	0
	Aide partielle pour une partie du corps	1
	Aide pour plusieurs parties du corps ou toilette impossible	2
Habillage	Autonome pour le choix et l'habillage	0
	S'habille mais besoin d'aide pour se chausser	1
	Besoin d'aide pour choisir ses vêtements ou s'habiller reste partiellement ou complètement déshabillé	2
Aller aux toilettes	Autonome	0
	Doit être accompagné, besoin d'aide	1
	Ne va pas aux WC, n'utilise pas le bassin, l'urinoir	2
Locomotion	Autonome	0
	Besoin d'aide	1
	Grabataire	2
Continence	Continent	0
	Incontinence occasionnelle	1
	Incontinence permanente	2
Repas	Autonome	0
	Aide pour couper la viande ou peler les fruits	1
	Aide complète ou alimentation artificielle	2

TAB. 3.1 – Grille ADL de Katz avec les scores associés à chaque comportement

tableau 3.2. Celui-ci présente l'échelle complète telle que définie par Lawton. Elle se réduit parfois à la première colonne du tableau. La notation est appréciée en fonction du nombre d'items notés 1 (donc dépendance plus ou moins forte du patient), mais aussi de l'évolution de ce nombre d'items.

3.4.3 La grille AGGIR

La grille **Autonomie Gérontologique - Groupes Iso-Ressources**, abrégée AGGIR, est utilisée en France dans le but d'évaluer l'autonomie physique et psychique grâce à l'observation des activités effectuées seules par la personne âgée, puis de définir, en fonction de la perte d'autonomie démontrée par cette grille, un groupe iso-ressources (une notation du degré de dépendance). Ce groupe est utilisé en France pour entre autre l'attribution d'aides telles que l'APA (Allocation Personnalisée d'Autonomie) ou bien l'évaluation du besoin d'intégration de la personne dans un environnement médicalisé (entrée en unité d'hospitalisation longue durée).

L'évaluation du degré d'autonomie se fait suivant un total de dix-sept variables, dont dix seulement entrent en compte dans le calcul (ces variables sont dites discriminantes). Les variables définies sont les suivantes :

- les variables discriminantes : cohérence, orientation, toilette, habillage, alimentation, élimination, transferts (se lever, se coucher, s'asseoir), déplacement à l'intérieur, déplacement à l'extérieur et enfin communication à distance.
- Les variables illustratrices : gestion personnelle de son budget et de ses biens, cuisine, ménage, transports, achats, suivi du traitement et enfin activités de temps libre.

Chacune de ces dix sept variables est notée par l'un des trois états suivants :

- A : n'a besoin d'aucune aide pour réaliser cette activité.
- B : réalise cette activité avec une aide partielle.
- C : une aide totale est requise pour réaliser l'activité.

À partir de ces critères, six groupes iso-ressources ont été définis, évaluant le degré d'autonomie et les besoins en aides et soins requis par le patient :

- GIR 1 : comprend les personnes âgées confinées au lit ou au fauteuil, dont les fonctions mentales sont gravement altérées et qui nécessitent une présence indispensable et continue d'intervenants.
- GIR 2 : concerne les personnes âgées confinées au lit ou au fauteuil, dont les fonctions intellectuelles ne sont pas totalement altérées et dont l'état exige une prise en charge pour la plupart des activités de la vie courante. Ce groupe s'adresse aussi aux personnes âgées dont les fonctions mentales sont altérées, mais qui ont conservé leurs capacités de se déplacer.
- GIR 3 : réunit les personnes âgées ayant conservé leur autonomie mentale, partiellement leur autonomie locomotrice, mais qui ont besoin quotidiennement et plusieurs fois par jour d'être aidées pour leur autonomie corporelle.
- GIR 4 : intègre les personnes âgées n'assumant pas seules leurs transferts mais qui, une fois levées, peuvent se déplacer à l'intérieur de leur logement. Elles doivent parfois

Téléphone		Approvisionnement	
0	Je me sers du téléphone, cherche et compose les numéros	0	Prend soin aux besoins de manière indépendante
1	Je compose un petit nombre de numéros bien connus	1	Réalise les commissions seul(e) pour de petits achats
1	Je réponds au téléphone mais n'appelle pas	1	Nécessite un accompagnement pour tout approvisionnement
1	Je suis incapable d'utiliser le téléphone	1	Incapable de sortir réaliser des achats
Moyens de transport		Entretien du logement	
0	Je peux voyager seul(e) et de façon indépendante	0	Maintient le logement en état seul(e) ou avec une assistance occasionnelle
1	Je peux me déplacer seul(e) en taxi, pas en autobus	1	Effectue des tâches journalières légères (laver la vaisselle, faire le lit, ...)
1	Je peux prendre les transports en commun si je suis accompagné(e)	1	Effectue des tâches journalières légères mais n'est pas capable de maintenir un niveau d'hygiène suffisant
1	Transport limité au taxi ou à la voiture en étant accompagné(e)	1	Nécessite une aide pour toute tâche de maintenance du logement
1	Je ne me déplace pas du tout	1	Ne participe à aucune tâche d'entretien du logement
Gérer son budget		Lessive	
0	Je suis totalement autonome	0	Effectue la lessive personnelle complète seul(e)
1	Je me débrouille pour les dépenses au jour le jour mais j'ai besoin d'aide pour gérer mon budget à long terme	1	Nettoie un petit nombre d'items (rince et stock...)
1	Je suis incapable de gérer l'argent nécessaire à payer les dépenses au jour le jour	1	Toute ma lessive doit être faite par quelqu'un d'autre
Prise de médicaments		Nourriture	
0	Je m'occupe moi même de la prise : dosage et horaire	0	Prévoit, prépare et sert un repas adapté de manière autonome
1	Je peux les prendre moi même si ils sont préparés et dosés à l'avance	1	Prépare des repas adaptés si la liste des ingrédients est fournie
1	Je suis incapable de les prendre moi-même	1	Fait chauffer, sert et prépare des repas ou bien prépare des repas mais n'utilise pas les aliments appropriés
		1	Nécessite que les repas soient préparés et servis

TAB. 3.2 – Échelle des Activités Instrumentales de la Vie Quotidienne de Lawton

être aidées pour la toilette et l'habillage. Ce groupe s'adresse également aux personnes âgées n'ayant pas de problèmes locomoteurs mais devant être aidées pour les activités corporelles et pour les repas.

- GIR 5 : comporte des personnes âgées ayant seulement besoin d'une aide ponctuelle pour la toilette, la préparation des repas et le ménage.
- GIR 6 : réunit les personnes âgées n'ayant pas perdu leur autonomie pour les actes essentiels de la vie courante.

Un classement dans les GIR 1 à 4 ouvre droit à des aides financières, dont l'APA, ... Un classement en GIR 5 et 6 ouvre néanmoins des accès à des prestations d'aides ménagères ou d'aides légères aux soins (par le régime de retraite ou par l'aide sociale départementale).

Un exemple concret de bilan d'autonomie, contenant cette grille ainsi que d'autres informations nécessaires pour une admission en service de soin de longue durée est donné dans l'annexe A.

3.4.4 Objectivation des critères

Le but de ces travaux de thèse est de détecter et classer les activités de la vie quotidienne. Celles-ci sont au cœur du calcul des différents indices présentés dans les trois sections précédentes. Cette classification des activités a pour but d'objectiver le remplissage de ces différentes grilles et d'effectuer un suivi longitudinal de la personne afin de détecter des changements dans son comportement. Ce travail n'est que la base d'un tel objectif.

Une telle objectivation peut être utile pour plusieurs raisons :

- Sachant que la personne est suivie à domicile, son entrée en institution est d'autant plus retardée, tant qu'aucun signe détecté ne la nécessite. Vivre à son propre domicile apporte en plus une meilleure qualité de vie à la personne âgée.
- Le suivi longitudinal permet de détecter des modifications du comportement révélatrices d'un certain état, et ce de manière peut-être plus précoce qu'en se basant sur des critères cliniques. Cette voie reste ouverte pour des recherches ultérieures.

3.5 Organisation du manuscrit

Ce manuscrit va s'articuler en trois grandes parties qui suivront cette introduction. La première sera consacrée à la mise en place des capteurs et des algorithmes alors que la seconde sera consacrée à l'expérimentation, à ses résultats et à une discussion sur ceux-ci. La troisième est composée d'une discussion générale et d'une conclusion.

La première partie contient quatre chapitres qui sont les suivants :

- Le chapitre 4 décrit l'implémentation et les algorithmes associés au circuit ACTIM6D, capteur cinématique mis au point au cours de cette thèse. Il décrira également une expérimentation sur ce capteur et les résultats associés.
- Le chapitre 5 décrit, quant à lui, la mise en place, la configuration et les algorithmes du système de reconnaissance de sons et de paroles installés dans le HIS par le laboratoire LIG. De même, une expérimentation a été réalisée pour en valider l'implémentation et

en connaître la sensibilité et la spécificité. Celle-ci ainsi que ses résultats seront décrits à la fin de ce chapitre.

- Le chapitre 6 aborde le problème du formatage des données en vue de leur classification. Il décrit d’abord le HIS et sa structure, avant de décrire les différents choix effectués pour chacun des capteurs.
- Le chapitre 7 rappelle tout d’abord les principes de la méthode de classification choisie et explique ce choix, avant d’aborder les implémentations qui en ont été faites pour répondre à notre problème.

La seconde partie contient deux chapitres :

- Le chapitre 8 décrit le protocole expérimental et rappelle les détails sur la classification que nous voulons effectuer à partir des données du HIS.
- Le chapitre 9 synthétise les résultats obtenus lors de cette expérimentation puis introduit une première discussion sur la classification des données dans un habitat intelligent.

La troisième partie clôt ce manuscrit par une discussion générale puis une conclusion et une ouverture sur l’avenir.

Deuxième partie

**La classification des activités dans le HIS :
de la mise en place des capteurs et des
algorithmes...**

Actimétrie embarquée : Conception, algorithmes et tests de l'ACTIM 6D

Depuis des temps immémoriaux, disons depuis dix mille ans, il y a peu de doute que l'évènement le plus marquant sera pour un temps la découverte (au *XIX^{ème}* siècle) de Maxwell sur les lois de l'électromagnétisme.

Richard Phillips Feynman

Sommaire

4.1 Conception de la carte ACTIM6D	63
4.1.1 Description	63
4.1.2 Positionnement du capteur	64
4.1.3 Acquisition et enregistrement des données	65
4.1.4 Récupération des données enregistrées sur carte mémoire	66
4.1.5 Perturbations extérieures	68
4.2 Algorithmes	70
4.2.1 Les données accélérométriques	70
4.2.1.1 Description des données brutes	70
4.2.1.2 Pré-traitement des données accélérométriques	71
4.2.1.3 Identification des motifs temporels	71
4.2.2 Les données magnétométriques	72
4.2.2.1 Description des données brutes	72
4.2.2.2 Identification des motifs temporels	74
4.2.2.2.1 Motifs et dépendance par rapport à l'orientation du porteur	74
4.2.2.2.2 Passage dans le domaine des quaternions	75
4.2.3 Segmentation du signal	77
4.2.3.1 Critère choisi pour la segmentation	77
4.2.3.2 Motifs détectés	77
4.2.4 Classification	78
4.2.4.1 La transformée en ondelettes continue pour la classification	78
4.2.4.2 Réalisation des motifs de détection	81
4.2.4.2.1 Séquence Debout-Assis	81
4.2.4.2.2 Séquence Assis-Debout	83
4.2.4.2.3 Séquence Debout-Allongé	84
4.2.4.2.4 Séquence Allongé-Debout	85

4.2.5	Périodes de marche	85
4.2.5.1	Détection des périodes de marche	85
4.2.5.2	Périodes de marche et segmentation	87
4.3	Résultats sur quelques scénarios	88
4.4	Fichier de sortie et résultats finaux	90
4.5	Expérimentations et résultats	91
4.5.1	But de l'expérimentation et population	91
4.5.2	Protocole expérimental	92
4.5.3	Résultats	92
4.6	Conclusion et discussion sur l'implémentation de la carte ACTIM6D	94
4.6.1	Limitations matérielles	94
4.6.1.1	Bruit de mesure	94
4.6.1.2	Positionnement du capteur	94
4.6.2	Logiciel	95
4.6.2.1	Intérêt des ondelettes	95
4.6.2.2	Discussion sur la segmentation et la classification	95
4.6.2.3	Analyse temps-fréquences de la marche	96
4.6.2.4	Perspectives	96
4.6.2.4.1	Correction des incohérences des données	96
4.6.2.4.2	Système embarqué?	97

4.1 Conception de la carte ACTIM6D

4.1.1 Description

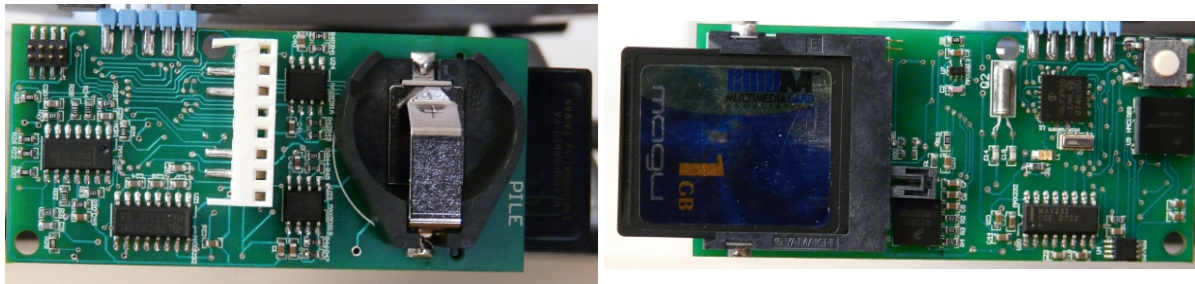


FIG. 4.1 – Prototype du circuit ACTIM6D

La carte ACTIM6D a été conçue dans le but d'obtenir différentes informations sur la personne qui la porte :

- ↪ Les changements de posture du porteur : passages Assis-Debout et Debout-Assis ainsi que Debout-Allongé et Allongé-Debout. Cette première étape nous permet de déduire les phases de la vie quotidienne dans lesquelles la personne est couchée, assise ou debout.
- ↪ Le niveau d'activité de la personne, qui pourrait nous donner des informations sur les phases jour/nuit, ainsi que sur les phases très actives (lorsque le porteur fait la cuisine, le ménage...) ou non-actives, lors de siestes, repos, séquence de détente devant la télévision, etc.
- ↪ Les séquences de marche de la personne.

Pour réaliser ces différentes tâches, nous avons équipé la carte ACTIM6D de plusieurs capteurs reliés aux entrées analogiques d'un microcontrôleur. Le microcontrôleur utilisé est un microchip PIC18LF2580 (architecture 16 bits), alimenté par une tension $V_{ss} = 3.3V$. Ce microcontrôleur est cadencé par un résonateur céramique à 4MHz, et possède en interne un convertisseur analogique-numérique à dix entrées (réparties sur ses ports), d'une précision de 10 bits (1024 niveaux).

Le premier capteur est un accéléromètre tri-axes (Freescale, MMA7260Q), qui fournit l'accélération ressentie sur les trois axes, en g, avec différentes résolutions : $\pm 1.5g$, $\pm 2g$, $\pm 4g$ et $\pm 6g$. Pour notre application, nous avons décidé de nous limiter aux accélérations $\pm 2g$ et $\pm 6g$, réglables par une ligne numérique d'un port du microprocesseur. Cet accéléromètre donne des valeurs analogiques avec un $0g$ à $\frac{V_{ss}}{2}$, et une résolution de $800mV/g$ à $\pm 1.5g$ et $600mV/g$ à $\pm 2g$. La bande passante de la réponse du capteur est à 150 Hz.

Le second capteur est un magnétomètre tri-axes (Honeywell, HMC1053), mesurant le champ magnétique environnant avec une étendue de $\pm 6Gs$, une bande passante de 5MHz et une résolution de $1mV/V/Gs$. Ce magnétomètre est présent pour nous donner les orientations du corps (une base mobile) par rapport à un champ fixe (le champ magnétique terrestre – considéré fixe dans l'intervalle de temps et de position sur Terre induit par l'expérimentation). Sa fonction sera donc la mesure du champ magnétique Terrestre, qui, rappelons-le, a

une amplitude approximative de $47\mu T = 0.47\text{Gs}$ (à Grenoble). Chaque axe devra donc mesurer le champ dans un intervalle de $\pm 0.6\text{Gs}$. Considérant les données du capteur, chaque couple de sorties du magnétomètre (un ensemble de deux sorties différentielles reliées à un pont de Wheastone donnant l'image de la mesure) est connecté au couple d'entrées d'un montage différentiel à amplificateur opérationnel, de gain 1000. La référence du montage à amplificateur opérationnel est positionnée à $\frac{V_{ss}}{2}$ pour ne donner que des mesures positives (redressement rendu obligatoire par l'entrée du convertisseur).

Le constructeur Honeywell fournit également un dispositif intégré pouvant délivrer les signaux magnétométriques à une fréquence de 8Hz (HMR3300), et qui va nous servir de référence afin de connaître la qualité de nos signaux magnétométriques et de vérifier l'implémentation réalisée. Un connecteur pour ce circuit a été implémenté sur la carte, ainsi que les circuits nécessaires à son utilisation. Ce circuit étant alimenté en 5V , ACTIM6D est alimentée par deux piles 3V (CR2025) et possède deux régulateurs (un 3.3V et un 5V). Deux circuits RS-232 de la société Maxim (MAX232CPE et MAX3232CPE) ont également été ajoutés pour la conversion de niveaux entre le 5V TTL du HMR3300 et le 3.3V TTL du microcontrôleur.

Le dernier capteur est un capteur de température (Maxim, MAX6613), donnant la température (sur le die du composant) entre $-55\text{ }^{\circ}\text{C}$ et $+130\text{ }^{\circ}\text{C}$.

Pour des questions de confort, ce circuit doit être autonome et embarqué sur la personne. Pour ce qui est de l'alimentation, un connecteur de piles bouton a été placé et permet l'accueil de deux piles 3V (CR2025). Ces piles ont une capacité moyenne de 250mAh . L'enregistrement des données se fait sur une carte mémoire, placée dans un connecteur de carte SD/MMC. À la vue de la consommation actuelle du capteur, l'autonomie lors de l'alimentation par les piles bouton s'élève à environ 20h. L'autonomie de la carte mémoire est quant à elle largement suffisante étant donné qu'elle atteint 23 jours avec une carte de 1 Go.

Le circuit implémentant le cahier des charges ainsi défini a été réalisé et routé sous Protel™. Le schéma électrique est donné en Annexe B. Pour le confort du porteur, la taille de ce circuit a été réduite au minimum, pour obtenir une taille de circuit de $6.5\text{cm} \times 3.7\text{cm}$. Du fait de la précision attendue, la réalisation du circuit imprimé a été sous-traitée à un industriel.

Le prototype ainsi réalisé est montré sur la figure 4.1.

4.1.2 Positionnement du capteur

Du fait de sa taille réduite, le circuit ACTIM6D se loge dans une poche de $9\text{cm} \times 5\text{cm}$ cousue sous l'aisselle gauche du porteur, directement sur le tee-shirt. Le positionnement sous l'aisselle a été choisi pour être cohérent avec les mouvements à détecter. Ces mouvements sont : s'asseoir sur une chaise, se lever d'une chaise, s'allonger sur un lit, se lever d'un lit et marcher. Lors de ces mouvements, et principalement lorsque la personne s'assoit ou se lève, la partie du corps qui représente le mieux les changements de posture est le buste. Nous avons donc intégré notre capteur au tee-shirt. Nous aurions également pu équiper le sujet d'un ensemble de capteurs identiques placés sur différentes parties du corps (typiquement la jambe). Cependant, pour notre type d'application, posséder un capteur embarqué est déjà une contrainte forte. Nous voulions limiter la contrainte en n'utilisant qu'un seul capteur et

en compensant ce choix par des algorithmes plus complexes. La figure 4.2 présente le positionnement du capteur sur le buste de la personne, ainsi que le placement des axes associés sur le sujet et sur le circuit. Les axes x , y et z ont été choisis pour former un trièdre direct et être identiques aux notions respectives de roulis (autour de x , angle ϕ), de tangage (autour de y , angle θ) et lacet (autour de z , angle ψ).

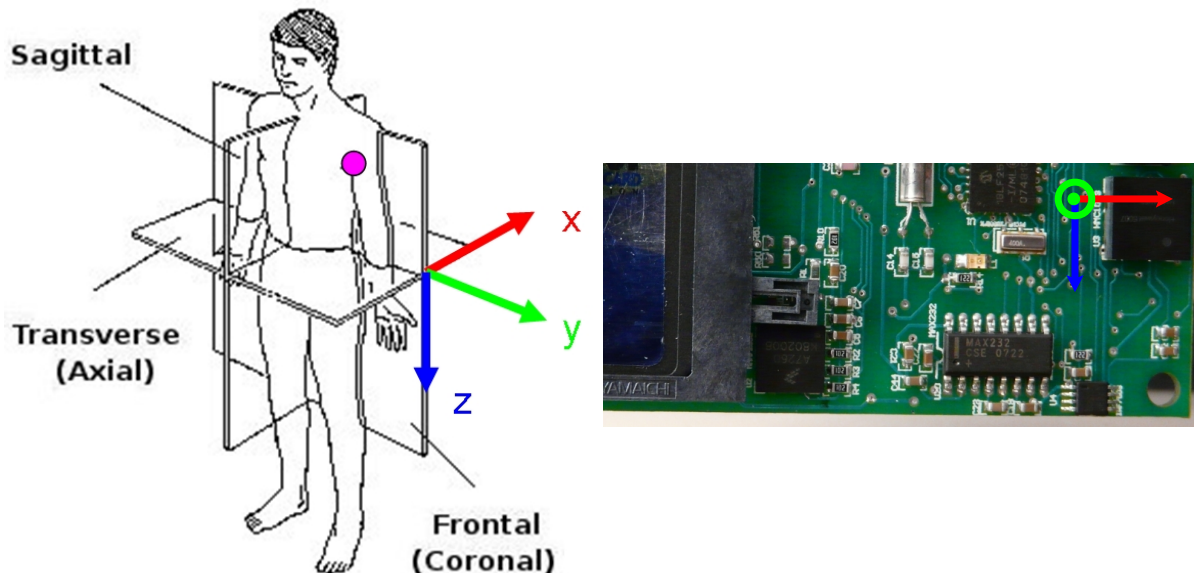


FIG. 4.2 – Positionnement du capteur sur le sujet et trièdre direct créé par la centrale et fixe sur la personne (figure de gauche). Positionnement de ce trièdre sur le capteur embarqué (figure de droite)

4.1.3 Acquisition et enregistrement des données

Le microcontrôleur a été programmé pour gérer l'acquisition des sept données avec une fréquence d'échantillonnage de 100 Hz. Toutes les 10ms, une interruption, lancée par le timer du microcontrôleur, le réveille pour lancer la lecture des valeurs de sept des entrées du convertisseur analogique/numérique. À la fin de la conversion des sept valeurs analogiques, le microcontrôleur est remis en veille à l'exception des moments où l'écriture sur la carte mémoire doit être exécutée. Ce fonctionnement est décrit par les diagrammes de la figure 4.3

L'écriture sur les cartes mémoires SD/MMC, sur le bus SPI, se fait par blocs de 512 octets. Les données acquises par le convertisseur analogique/numérique sont sur 10 bits. Pour un stockage efficace, et peu gourmand en calcul, seuls les 12 bits de poids faible de l'entier 16 bits contenant les valeurs sont stockés sur la carte mémoire pour les données cinématiques, et 8 bits (qui sont suffisant pour coder la plage entière de valeurs) pour le capteur de température. L'ensemble des sept valeurs tient ainsi sur 10 octets. Les valeurs sont alors stockées sur la carte mémoire toutes les 500ms. Sont ajoutés aux 500 octets de données des indicateurs sur le numéro de l'expérience en cours (dépendant du nombre d'allumage du capteur) et le temps depuis le début de l'expérience pour le bloc en cours (si la carte mémoire n'est pas vide, l'écriture reprend à l'endroit où elle avait été interrompue à l'expérience précédente).

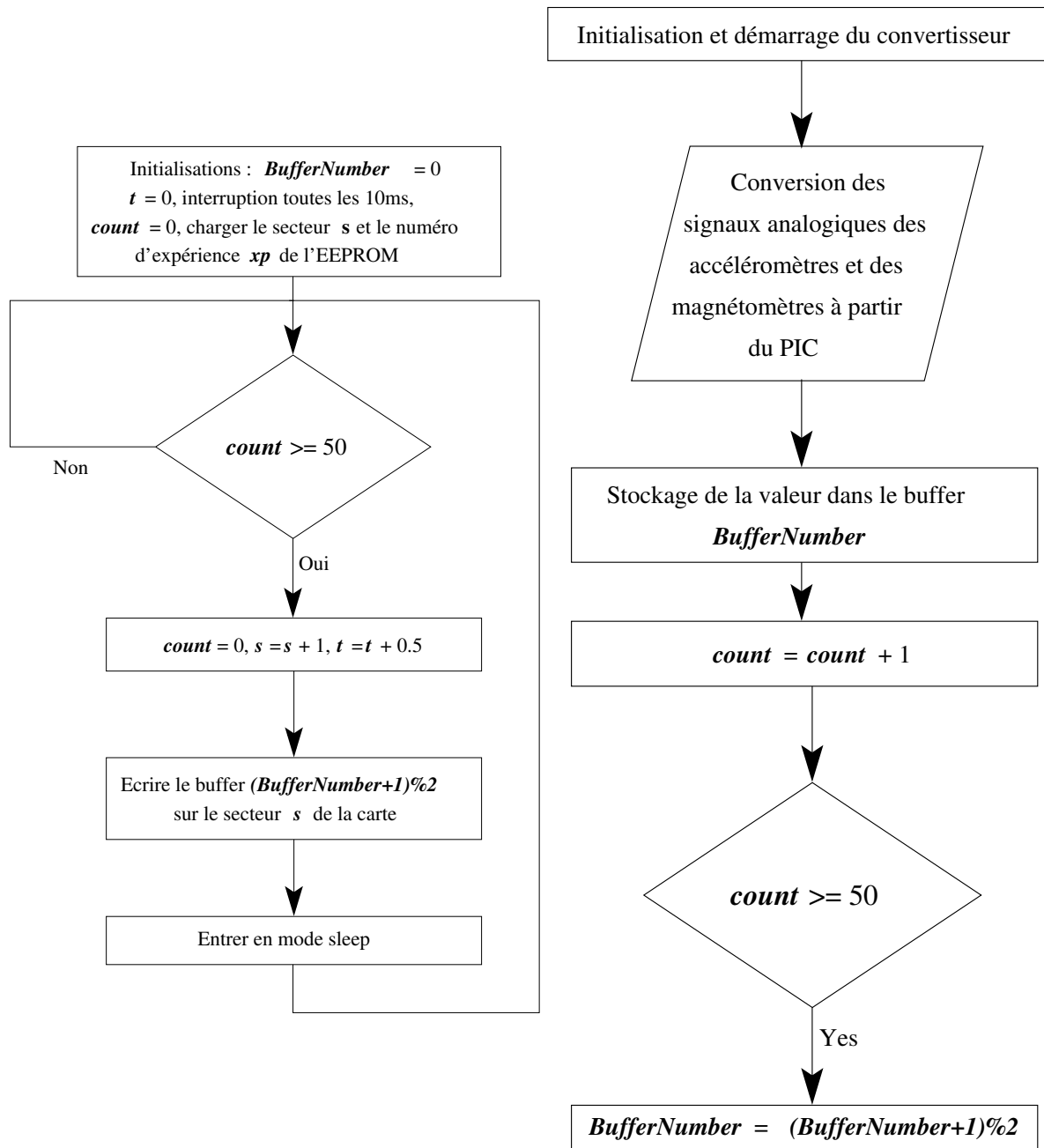


FIG. 4.3 – Diagramme décrivant le fonctionnement d'ACTIM6D. Le diagramme de gauche représente la procédure principale et le diagramme de droite la procédure d'interruption associée au timer.

La carte mémoire peut ultérieurement être lue par un programme permettant de récupérer les données brutes et de les reconstruire (aucun système de fichier n'a été créé pour des raisons de place en mémoire programme et de consommation).

4.1.4 Récupération des données enregistrées sur carte mémoire

Le microcontrôleur que nous utilisons, associé au lecteur de carte mémoire, permet l'écriture secteur par secteur sur les cartes insérées, qu'elles soient de type SD ou MMC. Pour être lisible sur un ordinateur, la carte mémoire devrait être équipée d'un système de fichiers (type FAT16 ou FAT32). Cependant, la présence de ce système de fichiers induit deux contraintes complémentaires. La première est un alourdissement du code source pour gérer ce système

de fichiers et implémenter une gestion partielle des noms de fichiers, des secteurs. . . La seconde est que chaque écriture de données engendrera la lecture et l'écriture de plusieurs secteurs (une, voire deux lectures puis au minimum trois écritures), et augmenter ainsi le temps de traitement pour un secteur. Ces lectures et écritures sont indispensables pour maintenir, à tout instant, un système de fichiers cohérent afin de ne perdre aucune donnée si la pile tombe en panne ou si le porteur la retire par exemple. Ces contraintes vont d'une part augmenter la consommation électrique du circuit et d'autre part réduire la fréquence maximale de capture.

Pour ces raisons, nous avons décidé de ne pas implémenter un tel système de fichiers et de créer un programme C sur PC capable de lire la carte mémoire (en format « brut ») et de récupérer les données. Ainsi, dès qu'un secteur est prêt à être écrit sur la carte, il le sera et ce de manière unique. La fréquence d'échantillonnage choisie est de 100Hz . Ainsi, un secteur de 512 octets sera écrit toutes les 500 *ms*. Chacun des secteurs contient les informations suivantes :

- Deux caractères '#' à la suite, permettant de repérer un secteur utile.
- L'identifiant de l'expérimentation, un chiffre sur 2 octets.
- Le nombre de secondes depuis le début de l'expérimentation, chiffre sur 8 octets.
- Les valeurs des capteurs sous la forme : *AA BC DD EE FG HH II JK LL MM* » mises à la suite les unes des autres, jusqu'à ce qu'un secteur soit plein :
 - AA les deux octets de poids faible de l'accéléromètre sur l'axe **x**.
 - B le dernier octet de l'accéléromètre **x**.
 - C le premier octet de poids faible de l'accéléromètre **y**.
 - DD les deux octets de poids fort de l'accéléromètre **y**.
 - EE les deux octets de poids faible de l'accéléromètre sur l'axe **z**.
 - F le dernier octet de l'accéléromètre **z**.
 - G le premier octet de poids faible du magnétomètre **x**.
 - HH les deux octets de poids fort du magnétomètre **x**.
 - II les deux octets de poids faible du magnétomètre **y**
 - J l'octet de poids fort du magnétomètre **y**
 - K le premier octet de poids faible du magnétomètre **z**
 - LL les deux octets de poids fort du magnétomètre **z**
 - MM les deux octets représentant la valeur du capteur de température.

Le numéro de l'expérimentation en cours ainsi que le numéro du dernier secteur écrit sont maintenus en mémoire flash (intégrée au processeur) afin de pouvoir reprendre l'écriture après le dernier secteur à la mise en route suivante, en incrémentant le numéro de l'expérience. Le fichier de sortie contient le numéro de l'expérimentation, la seconde dans laquelle a été prise l'échantillon, les valeurs des trois accéléromètres, des trois magnétomètres et enfin du capteur de température. Sa forme est la suivante :

0	90	015d	0260	01ac	0146	001c	0295	31
0	90	0158	0264	01ac	0148	001a	0296	31
0	90	0161	026a	01ac	014c	001b	0298	31
0	91	0148	028f	0188	0143	001c	028f	31
0	91	0173	0240	01cf	0149	001c	0287	28
0	91	0156	025b	01bc	0147	001a	029a	32

4.1.5 Perturbations extérieures

Le bruit de mesure peut être un problème lors de l'acquisition de signaux très faibles pouvant se noyer dans du bruit. Les accéléromètres pour leur part peuvent être soumis aux bruits usuels tel que le bruit Johnson, et à la diaphonie (réduite au minimum lors de l'implémentation en appliquant les règles élémentaires de la compatibilité électromagnétique lors de la conception de notre circuit). Il pourrait être également soumis à un bruit mécanique, dans le cas où un objet très massif viendrait perturber le champs gravitationnel à côté de la personne, cependant ce cas est peu probable. Enfin, le dernier bruit de mesure pour ces accéléromètres est le bruit de quantification et les approximations calculatoires faites après-coup sur les calculs. La figure 4.4 présente les signaux acquis avec la carte posée, immobile, sur un bureau. Lors de cette séquence de mesure d'une demi-heure, nous mesurons ce bruit avec la moyenne et l'écart-type de la norme des accéléromètres et des magnétomètres. Pour les accéléromètres, la norme moyenne était de $1,0712 \text{ m/s}^2 \pm 0,0449$. Pour les magnétomètres, la norme était de $435,6 \text{ mGs} \pm 6,7$. À noter que cette figure 4.4 présente les signaux avec un code de couleurs similaire à la figure 4.2 donnant le positionnement des capteurs. Afin de ne pas alourdir inutilement l'ensemble des courbes présentées, les signaux accéléromètres et magnétométriques seront toujours tracés avec ce même code de couleurs qui ne sera plus rappelé dans les graphiques suivants.

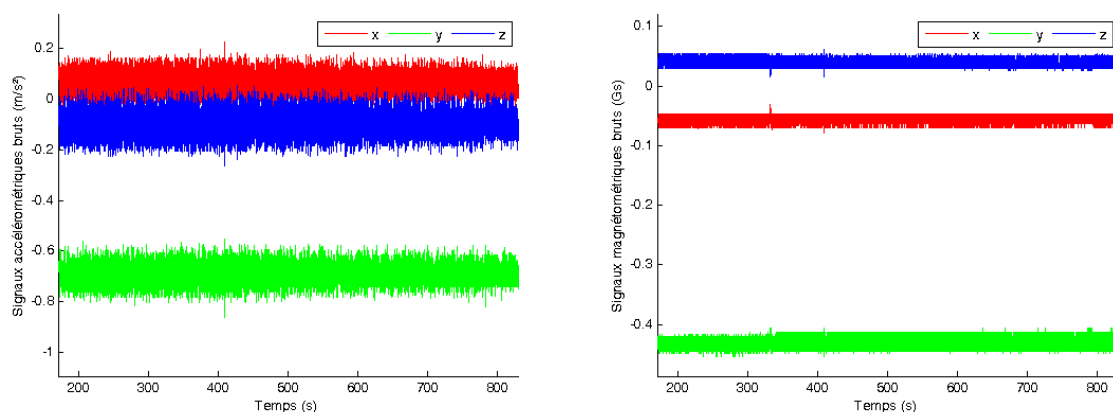


FIG. 4.4 – Signaux accélérométriques (à gauche) et magnétométriques (à droite), période d'immobilité éloignée de fortes perturbations électromagnétiques

Contrairement aux accéléromètres, les magnétomètres mesurent par contre, quant à eux, un champ très faible qui peut être perturbé par certains appareils utilisés dans la vie quotidienne. Pour vérifier ceci, nous avons effectué des mesures dans diverses conditions afin

de mesurer l'influence des différents champs sur la mesure du champ magnétique terrestre. Par exemple, un ascenseur, qui possède un moteur à courant continu (et de grosses parties métalliques), va créer un champ variable beaucoup plus important que celui de la Terre. Ainsi, la figure 4.5 nous montre les variations du champ magnétique perçu, lors d'une séquence d'une demi-heure, pendant laquelle ACTIM6D a été posée fixe à côté de l'ascenseur et celui-ci a été appelé vers différents étages. Ceci nous donne comme statistique un champ gravitationnel de $1,0328 \text{ m/s}^2 \pm 0,0230$ ainsi qu'un champ magnétique de $468,6 \text{ mGs} \pm 75,6$.

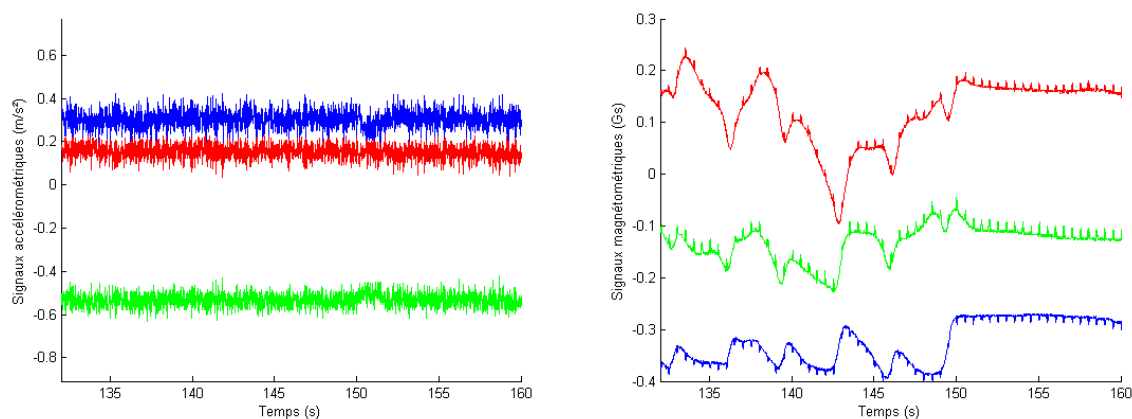


FIG. 4.5 – Signaux accélérométriques (figure de gauche) et magnétométriques (figure de droite), période d'immobilité à proximité (20cm) d'un ascenseur

Cette mesure a été faite à 20cm de l'ascenseur environ. À noter que lorsque nous nous éloignons à 1m, l'influence de l'ascenseur se réduit et les mesures redeviennent correctes. Des tests ont également été réalisés avec différents appareils de l'appartement. Ainsi, pendant une demi-heure également, la carte a été posée à proximité du réfrigérateur, de la télévision, d'un écran d'ordinateur (tube cathodique) ainsi que d'un four à micro-ondes. Des mesures ont été faites en comparant les statistiques du signal à une même distance et dans une même position avec le même appareil éteint puis allumé. L'ensemble de ces mesures a révélé qu'aucun de ces appareils ne modifiait le champ perçu (à distance acceptable, équivalente à la plus courte distance que nous pourrions atteindre lorsqu'une personne est à proximité de l'appareil). Le seul appareil modifiant le champ est donc l'ascenseur. Annuler l'effet d'un champ non constant sans référence est relativement difficile. Pour l'instant, nous considérons donc qu'une mesure faite à proximité de ces appareils est inutilisable et les fenêtres temporelles sont alors annulées (en se basant sur une variation trop importante de la norme du champ).

4.2 Algorithmes

4.2.1 Les données accélérométriques

4.2.1.1 Description des données brutes

Chaque personne et chaque objet sur Terre subissent la force gravitationnelle. Cette force, responsable du fait que nous sommes attirés vers le sol (et donc qu'il est plus difficile sur Terre que sur la Lune d'effectuer un bond d'une hauteur donnée...), est une force d'interaction qui crée une attraction réciproque entre deux corps massiques. Cette force s'exprime par la relation :

$$\mathbf{F}_g = -\frac{G \cdot m_1 \cdot m_2}{R^2} \mathbf{u}_{m_1 m_2}$$

Dans laquelle $G = 6,67428 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2}$ représente la constante de gravitation universelle et $\mathbf{u}_{m_1 m_2}$ est le vecteur unitaire allant de la masse m_1 vers la masse m_2 . En considérant, dans cette équation, la masse de la Terre ainsi que son rayon, nous retombons sur la formule de l'attraction Terrestre (appelée plus communément le « poids ») :

$$\mathbf{P} = m \cdot \mathbf{g}$$

Où $\mathbf{g}$ est le vecteur accélération de la pesanteur sur Terre, de norme approximative $9,81 \text{ m s}^{-2}$ dirigé vers le sol et orthogonal à celui-ci.

En statique, l'accéléromètre va mesurer la projection sur chacun des axes de cette force de pesanteur. Ainsi, par exemple, lorsque la personne est allongée sur le dos et immobile, l'axe selon lequel la pesanteur est dirigée est $\mathbf{x}$, les sorties de l'accéléromètre (en considérant une mesure idéale et non bruitée) seront donc $+1g$ pour l'axe $\mathbf{x}$ et $0g$ pour les deux autres axes. Dans la position debout l'axe $\mathbf{z}$ sera à $+1g$ alors que les deux autres seront à $0g$. La sortie à $+1g$ sur l'axe $\mathbf{y}$ sera atteinte lorsque la personne est allongée sur son côté gauche. À titre d'exemple également, lorsque le buste est positionné à 45° vers l'avant alors que la personne est debout, nous obtenons les projections du vecteur de pesanteur sur notre accéléromètre, la sortie sur l'axe $\mathbf{x}$ sera alors de $-\cos\left(\frac{\pi}{4}\right) = -\frac{\sqrt{2}}{2}$ (l'axe $\mathbf{x}$ étant dirigé vers l'arrière la sortie est négative) alors que la sortie sur $\mathbf{z}$ sera $\sin\left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2}$.

Ces précédents cas représentent le cas statique, lorsque l'accéléromètre ne subit aucune autre accélération supplémentaire que celle de la pesanteur. Or, dans un cas dynamique, c'est à dire lorsque le corps en mouvement ajoute une accélération du fait de la force qu'il met en jeu pour ce mouvement, les valeurs captées comprendront une part statique et une part dynamique,

$$a_{\text{mesurée}} = a_{\text{pesanteur}} + a_{\text{mouvement}}$$

Dans ce cas, la norme du vecteur accélérométrique va évoluer et différer de sa norme de $+1g$. Lors d'un mouvement représentant un changement de posture, ces changements de normes vont donner une indication sur le mouvement actuellement effectué.

4.2.1.2 Pré-traitement des données accélérométriques

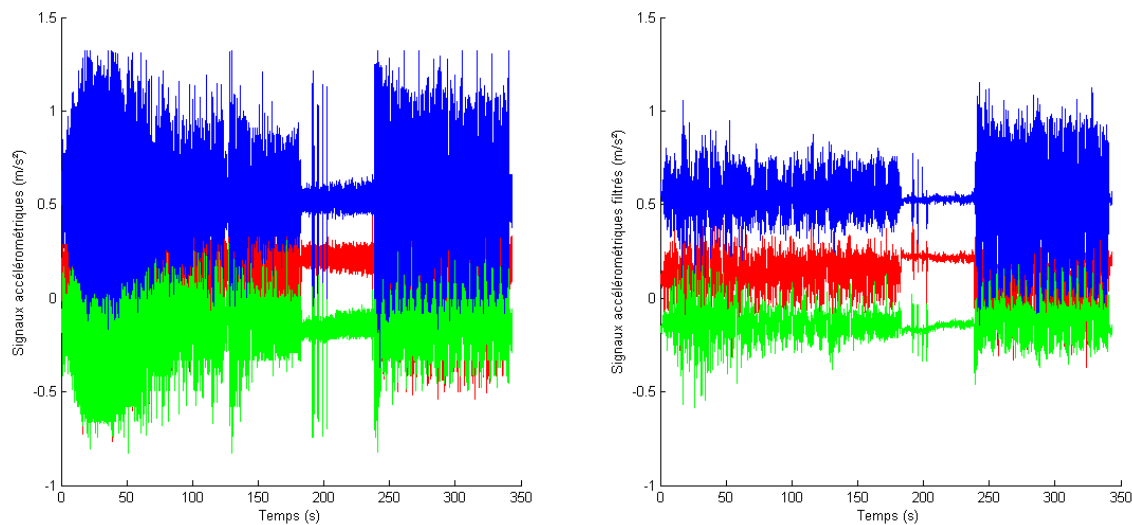


FIG. 4.6 – Signaux des accéléromètres pour une séquence de marche lente puis une séquence de marche rapide. À gauche, les signaux bruts et à droite, les signaux filtrés par le filtre de Butterworth

Les données accélérométriques sont acquises à une fréquence de 100Hz . Pour un capteur positionné sur le buste du sujet, [Bouten *et al.*, 1997] montre cependant que les mouvements de plus haute fréquence du corps humain se situent à 5Hz pour les activités de la vie quotidienne. Ainsi, nous utiliserons un filtre de Butterworth d'ordre 6 (qui lui donne une pente de -120dB/decades). La raison de ce choix est la constance de ce type de filtre dans la bande passante (filtres dits « maximally flat »). Ils engendrent ainsi beaucoup moins de bruits dans la bande passante que les autres filtres numériques du même type (Cauer, Chebyshev, Bessel). La contrepartie de cette absence de rebonds est que la valeur de la pente aux abords de la fréquence de coupure (roll-off steepness) est plus faible. Cependant, notre but, afin de faire une analyse temps-fréquences par la suite, est d'éliminer tout rebond dans la bande analysée. La figure 4.6 présente les résultats de ce filtrage sur une séquence de dix aller-retours lents (15m de marche) puis dix aller-retours rapides.

4.2.1.3 Identification des motifs temporels

La figure 4.7 présente une séquence Debout-Allongé-Debout. La figure de droite nous montre un agrandissement de deux des dix réalisations de gauche. De même, la figure 4.8 montre une séquence Debout-Assis-Debout répétée dix fois à gauche et un agrandissement de deux réalisations à droite.

L'observation de ces deux figures nous montre que la forme des signaux accélérométriques dans ces deux types de mouvement est répétable. En effet, ces mouvements font intervenir une force donnée par la personne pour le réaliser. La forme de la force donnée pour ce mouvement, et donc de l'accélération qui en découle, est la même entre deux réalisations. L'intensité du mouvement, ainsi que le temps pour le réaliser (déformation avec un coefficient sur l'échelle temporelle de la forme), peuvent cependant être modifiés.

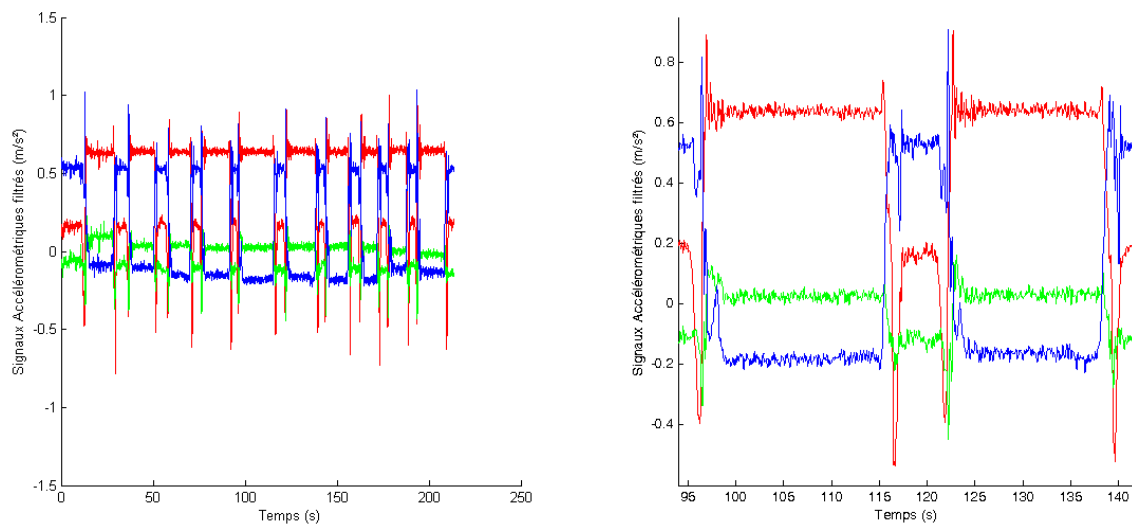


FIG. 4.7 – Signaux filtrés des accéléromètres de dix séquences Debout-Allongé-Debout à gauche. Agrandissement sur deux des séquences à droite

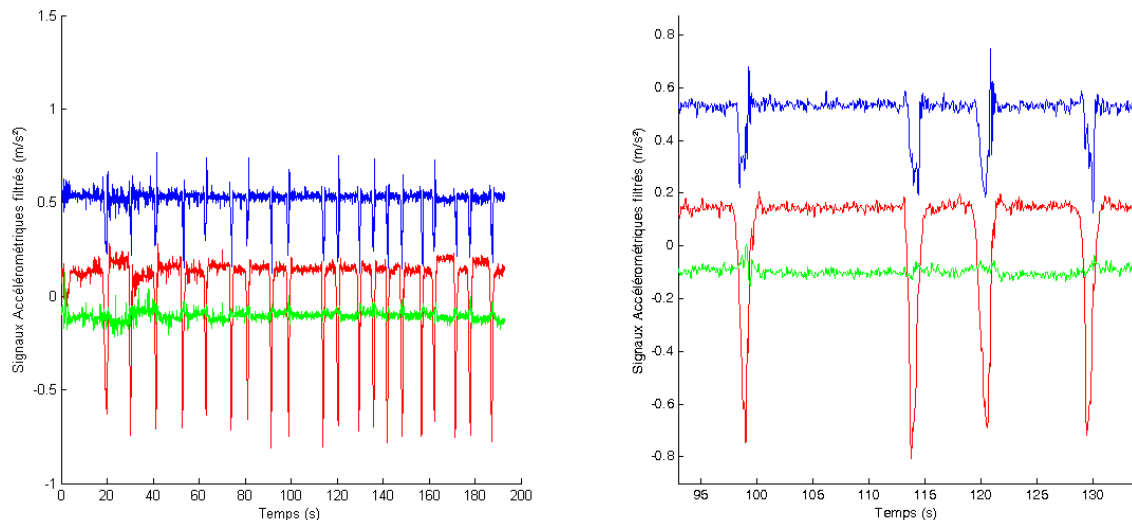


FIG. 4.8 – Signaux filtrés des accéléromètres de dix séquences Debout-Assis-Debout à gauche. Agrandissement sur deux des séquences à droite

En ce qui concerne le signal accélérométrique, les signaux filtrés seront donc pris comme signaux d'entrée pour la détection et la classification des postures.

4.2.2 Les données magnétométriques

4.2.2.1 Description des données brutes

Les magnétomètres mesurent le champ magnétique environnant. Or, la Terre, du fait de son noyau métallique liquide, crée un champ magnétique que nous pouvons observer à l'aide d'une boussole (qui nous indique le nord géographique ou sud magnétique de la Terre). Ce champ est décrit par la figure 4.9. En première approximation, ce champ peut être vu comme un dipôle magnétique, dont l'orientation est décalée de 11° par rapport au dia-

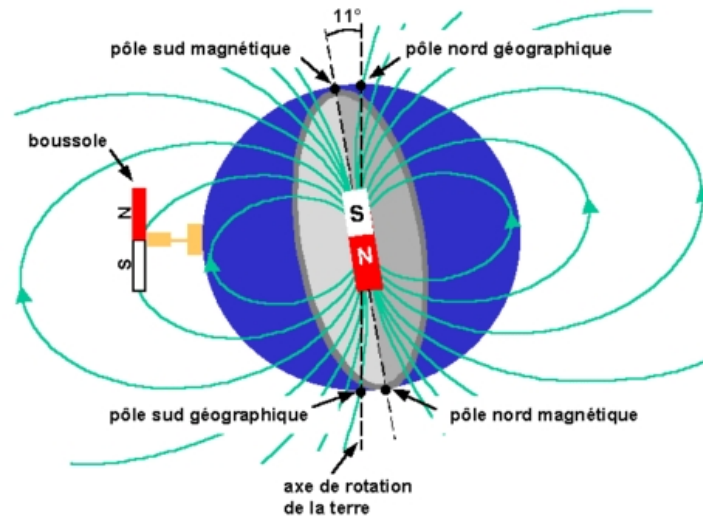


FIG. 4.9 – Positionnement du dipôle magnétique équivalent au champ magnétique terrestre mètre passant par les deux pôles terrestres.

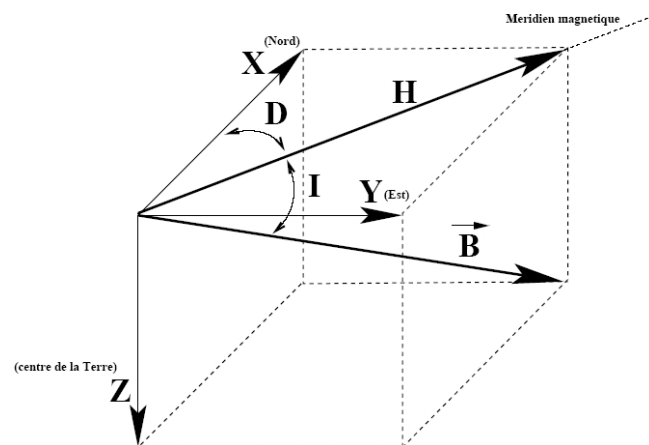


FIG. 4.10 – Orientation du champ magnétique Terrestre ([Chambodut, 2004])

Ce champ magnétique est la donnée mesurée à l'aide des magnétomètres du circuit. Cependant, celui-ci est très faible. Il est de l'ordre de $470mGs$ en France. Il est décrit et analysé plus précisément dans [Chambodut, 2004]. La figure 4.10 montre l'orientation de ce champ. Il est donc orienté à 11° par rapport à la ligne Nord-Sud des pôles (angle D sur la figure) et à environ 60° en direction du centre de la Terre par rapport à la surface (angle I sur la figure). Sur cette figure, le plan (x,y) représente le sol terrestre alors que le plan (x,z) se dirige vers le centre de la Terre et réalise un plan de coupe de notre planète révélant sa structure en couches.

Comme pour le champ gravitationnel, le magnétomètre tri-axes mesure la projection, sur les trois axes de la carte (figure 4.2) du champ magnétique Terrestre. Ainsi, la valeur de $470mGs$ serait mesurée (et donc en sortie du magnétomètre) par son axe x, lorsque la carte est positionnée suivant ce champ. Les modifications des projections de ce champ sont une

donnée représentative, en donnant le mouvement d'un vecteur fixe dans un repère mobile, mathématiquement équivalent à un repère fixe et un vecteur mobile.

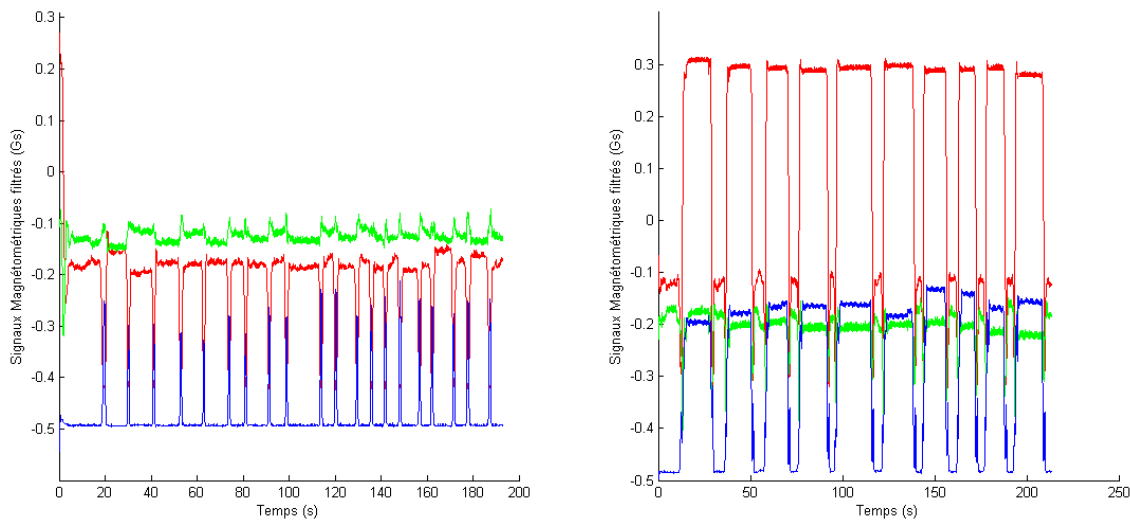


FIG. 4.11 – Signaux filtrés des magnétomètres. À gauche pour une séquence de dix Debout-Assis-Debout et à droite pour une séquence de dix Debout-Allongé-Debout

4.2.2.2 Identification des motifs temporels

4.2.2.2.1 Motifs et dépendance par rapport à l'orientation du porteur

La figure 4.11 donne le tracé des magnétomètres dans deux séquences différentes. À gauche, la figure représente le mouvement Debout-Assis-Debout alors que la figure de droite représente la séquence Debout-Allongé-Debout. Ces deux figures montrent un signal répétable entre chacune des dix séquences de chaque mouvement. Cependant, les signaux magnétométriques sont différents des signaux accélérométriques. Pour les accéléromètres, en plus de l'accélération créée par le corps, la mesure se fait sur le champ gravitationnel, qui est toujours suivant le corps du sujet lors du départ du mouvement et dont la position évolue toujours de la même manière. Le seul moyen de changer cette évolution et cette « orientation de départ » serait de faire le mouvement dans une position qui ne peut être réalisée du fait de la gravitation justement (par exemple s'asseoir sur une chaise dont le siège serait orienté parallèlement à un mur).

Cependant, l'orientation du champ magnétique terrestre rend les choses plus complexes pour ce paramètre. En effet, les valeurs de départ (en statique) ainsi que l'évolution de la forme des projections du champ dépendent de la position du sujet par rapport au nord magnétique au moment du mouvement. Pour se rendre compte de ce phénomène, la figure 4.12 présente la même séquence (dix mouvements Debout-Assis-Debout) réalisée sur deux chaises identiques placées à angle droit l'une de l'autre. Comme nous pouvons le constater, la représentation du signal magnétométrique, contrairement aux accéléromètres, est fortement dépendante de l'orientation du sujet. Pour résoudre ce problème, il est nécessaire de changer de représentation pour les magnétomètres en considérant l'écriture de la transformation entre deux instants dans le domaine des quaternions.

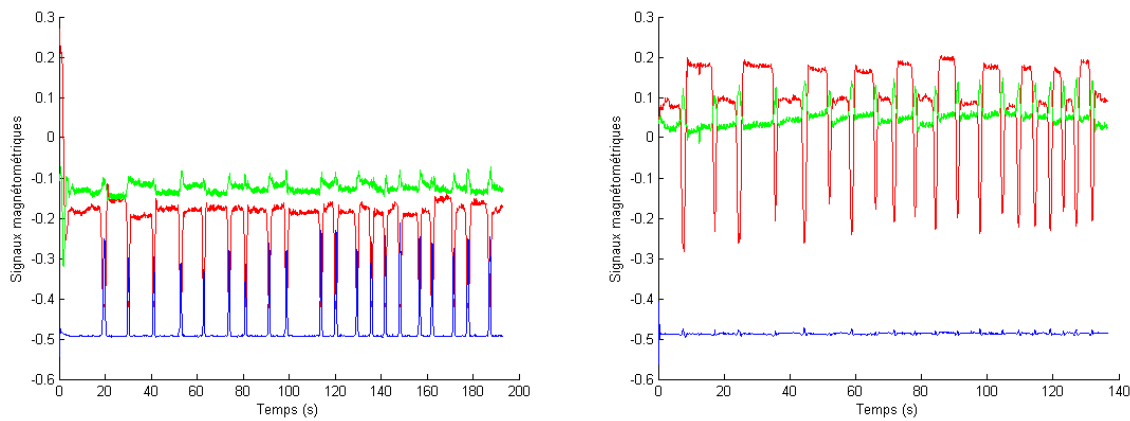


FIG. 4.12 – Signaux filtrés des magnétomètres pour une séquence de dix Debout-Assis-Debout. À gauche sur une première chaise et à droite sur une seconde placée à 90° de la première

4.2.2.2.2 Passage dans le domaine des quaternions

Les rotations dans notre espace en trois dimensions sont des applications linéaires d'un groupe appelé $SO(3)$ (groupe Special Orthogonal). Les rotations appartiennent à ce groupe car chacune d'entre elles peut être représentée par une matrice orthogonale à déterminant unitaire. Les deux propriétés de chaque rotation, dont la matrice peut être notée R , sont donc : $R^T R = I$ (avec I la matrice identité) et $\det R = 1$.

En deux dimensions, les rotations peuvent être représentées par les nombres complexes. En effet, une matrice de rotation d'angle θ dans le plan est donnée par :

$$R_2 = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \quad (4.1)$$

Cette rotation peut également être écrite à l'aide d'un nombre complexe de module 1. En considérant un nombre complexe $z = x + iy = \cos \theta + i \sin \theta$, l'argument de ce nombre complexe est aussi l'angle de la rotation ($\mathbf{Ox}$, $\mathbf{Oz}$). Afin d'étendre cette notion et de représenter les rotations de $\mathbb{R}^3$, Hamilton a introduit un plan hypercomplexe, algèbre en quatre dimensions sur le corps des nombres réels, noté $\mathbb{H}$ [Tait, 1873]. Ce groupe est défini par l'ensemble suivant :

$$\mathbb{H} = \{a + bi + cj + dk \mid [a, b, c, d] \in \mathbb{R}^4\} \quad (4.2)$$

le triplet (i,j,k) de cette relation est défini par :

$$i^2 = j^2 = k^2 = ijk = -1 \quad (4.3)$$

L'algèbre des quaternions n'est pas commutative. Elle est partiellement anticommutative, c'est à dire que $1 \cdot i = i \cdot 1 = i$ mais que $i \cdot j = k$ et $j \cdot i = -k$.

Considérons deux quaternions $Q_1 = \begin{pmatrix} a_1 \\ b_1 i \\ c_1 j \\ d_1 k \end{pmatrix}$ et $Q_2 = \begin{pmatrix} a_2 \\ b_2 i \\ c_2 j \\ d_2 k \end{pmatrix}$

Les principales opérations sur les quaternions sont les suivantes :

– la somme de deux quaternions :

$$Q_3 = Q_1 + Q_2 = \begin{pmatrix} a_1 + a_2 \\ (b_1 + b_2) i \\ (c_1 + c_2) j \\ (d_1 + d_2) k \end{pmatrix} \quad (4.4)$$

– le produit de deux quaternions :

$$Q_3 = Q_1 \cdot Q_2 = \begin{pmatrix} a_1 a_2 - b_1 b_2 - c_1 c_2 - d_1 d_2 \\ (a_1 b_2 + b_1 a_2 + c_1 d_2 - d_1 c_2) i \\ (a_1 c_2 + c_1 a_2 + d_1 b_2 - b_1 d_2) j \\ (a_1 d_2 + d_1 a_2 + b_1 c_2 - c_1 b_2) k \end{pmatrix} \quad (4.5)$$

– le conjugué d'un quaternion :

$$Q_3 = Q_1^* = \begin{pmatrix} -a_1 \\ -b_1 i \\ -c_1 j \\ -d_1 k \end{pmatrix} \quad (4.6)$$

– la norme d'un quaternion :

$$\| Q_1 \| = \sqrt{a^2 + b^2 + c^2 + d^2} \quad (4.7)$$

– inverse d'un quaternion :

$$Q_1^{-1} = \frac{1}{\| Q_1 \|^2} Q_1^* \quad (4.8)$$

Les quaternions ont également un autre type de notation, faisant intervenir un angle θ et un vecteur k . Cette notation repose sur le fait que toute combinaison de rotation dans l'espace en 3 dimensions peut être représentée par une unique rotation autour d'un vecteur $\mathbf{k} = (k_x \ k_y \ k_z)^T$ et dont l'angle de la rotation est θ . Cette représentation revient à noter le quaternion Q de la manière suivante :

$$Q = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) k_x \\ \sin\left(\frac{\theta}{2}\right) k_y \\ \sin\left(\frac{\theta}{2}\right) k_z \end{pmatrix} \quad (4.9)$$

Ainsi, un quaternion unitaire (il est trivial que tout quaternion donné par l'équation 4.9 est unitaire) représentera la rotation entre un vecteur $\mathbf{B}_1$ et un vecteur $\mathbf{B}_2$, sous la forme vecteur principal de la rotation et angle principal de la rotation. La combinaison de deux rotations dans ce nouvel espace est donnée par le produit des deux quaternions (équation 4.5).

tion 4.5). Les équations 4.6 et 4.8 montrent également que pour « inverser l'effet d'une rotation », comme en deux dimensions avec les complexes, on considère la rotation décrite par le conjugué du quaternion. À partir de ces deux vecteurs, nous pouvons trouver une nouvelle base $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ définie par :

$$\begin{pmatrix} \mathbf{i} \\ \mathbf{j} \\ \mathbf{k} \end{pmatrix} = \begin{pmatrix} \frac{\mathbf{B}_1}{\|\mathbf{B}_1\|} \\ \frac{\|\mathbf{B}_1\|^2 \mathbf{B}_2 - (\mathbf{B}_1 \cdot \mathbf{B}_2) \mathbf{B}_1}{\|\|\mathbf{B}_1\|^2 \mathbf{B}_2 - (\mathbf{B}_1 \cdot \mathbf{B}_2) \mathbf{B}_1\|} \\ \frac{\mathbf{B}_1 \wedge \mathbf{B}_2}{\|\mathbf{B}_1 \wedge \mathbf{B}_2\|} \end{pmatrix} \quad (4.10)$$

En considérant les vecteurs $\mathbf{b}_1$ et $\mathbf{b}_2$, projections respectives de $\mathbf{B}_1$ et $\mathbf{B}_2$ dans la nouvelle base. La rotation transformant $\mathbf{b}_1$ en $\mathbf{b}_2$ peut être écrite, dans $(\mathbf{i}, \mathbf{j}, \mathbf{k})$, comme une rotation autour de $\mathbf{k}$ d'angle θ avec :

$$\theta = \arccos\left(\frac{\mathbf{b}_1 \cdot \mathbf{b}_2}{\|\mathbf{b}_1\| \|\mathbf{b}_2\|}\right) \quad (4.11)$$

Nous utiliserons la représentation donnée par l'équation 4.9 afin de trouver des motifs temporels indépendants de la position du porteur du capteur par rapport au nord magnétique. Le premier vecteur ($\mathbf{b}_1$) servant de référence sera pris au début de la fenêtre d'analyse donnée par l'algorithme de segmentation suivant. Les motifs temporels obtenus ainsi que leur invariance sont illustrés dans les sections suivantes.

4.2.3 Segmentation du signal

4.2.3.1 Critère choisi pour la segmentation

Cette segmentation, première étape du traitement, est utilisée pour réduire la quantité de calcul lors de la classification. En effet, lorsqu'aucun mouvement significatif n'a lieu, il n'est pas utile de tenter d'y détecter un changement de posture. Les figures 4.4, 4.5, 4.7 et 4.8 nous montrent que (comme nous pouvions nous y attendre) l'amplitude des variations des accéléromètres est plus importante dans le cas d'une séquence Allongé-Debout ou Debout-Allongé que dans une séquence Assis-Debout ou Debout-Assis.

Ainsi, la décision sera prise sur l'écart-type de la norme des accéléromètres. Cette norme va évoluer au cours du temps du fait que les accéléromètres, contrairement aux magnétomètres, vont capter aussi l'accélération dynamique induite par le corps en mouvement. Les magnétomètres, par contre, ont des variations qui ne sont pas assez significatives et donneraient lieu à trop de faux positifs pour être utilisés pour cette segmentation.

4.2.3.2 Motifs détectés

À partir de plusieurs réalisations par une même personne des deux mouvements Assis-Debout et Allongé-Debout, nous avons calculé les deux écart-types respectifs s_{ass} et s_{all} nous permettant de prendre la décision suivant le critère suivant :

Soit $N_Accelero_i$ la norme des accéléromètres sur la fenêtre numéro i , et s_i l'écart-type

des données à l'intérieur de cette fenêtre :

$$\text{Segmentation}_i = \begin{cases} 1 & \text{si } 0,5 \cdot s_{ass} \leq s_i \leq 1,5 \cdot s_{ass} \\ 2 & \text{si } 0,5 \cdot s_{all} \leq s_i \leq 1,5 \cdot s_{all} \\ 0 & \text{sinon} \end{cases} \quad (4.12)$$

Les seuils de détection à plus ou moins la moitié de la valeur initiale ont été choisis suite à des expérimentations préliminaires.

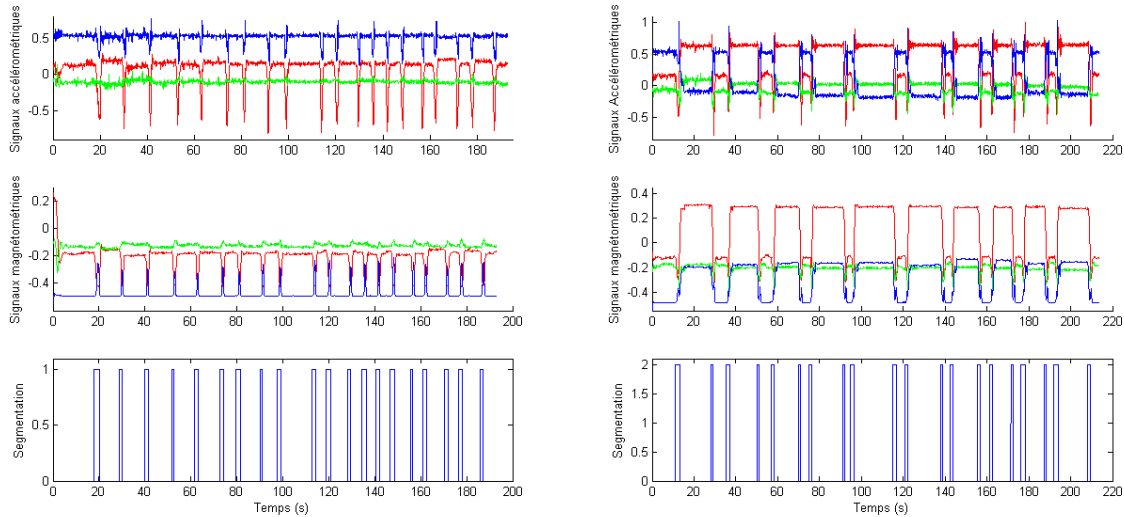


FIG. 4.13 – Résultats de l'algorithme de segmentation des signaux, à gauche la segmentation d'une séquence Debout-Assis-Debout répétée dix fois, à droite la segmentation d'une séquence Debout-Allongé-Debout répétée dix fois

La figure 4.13 montre les résultats de cet algorithme de segmentation. Comme nous pouvons le vérifier, les périodes de non activité sont bien éliminées et la classification peut alors se faire uniquement sur les fenêtres sélectionnées, diminuant ainsi la quantité de calcul à réaliser. Cette segmentation ne suffit cependant pas. En effet, lors d'une période de marche rapide par exemple, de nombreux faux positifs apparaissent sur le transfert Assis-Debout, et ramasser un objet à Terre donne de faux positifs sur le transfert Allongé-Debout. La classification va nous permettre d'éliminer ces faux positifs en trouvant des critères plus sélectifs sur les signaux accélérométriques et magnétométriques, mais également plus gourmands en calcul.

4.2.4 Classification

4.2.4.1 La transformée en ondelettes continue pour la classification

Pendant de très nombreuses années, la transformée de Fourier et l'analyse fréquentielle des signaux ont dominé le traitement du signal. La transformée de Fourier $\mathcal{F}$ cherche à décomposer une fonction en série infinie de cosinus et de sinus à différentes fréquences. Elle se définit par l'application suivante :

$$\mathcal{F}(f) : \omega \mapsto \hat{f}(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-2i\pi\omega t} dt \quad f \in L^2(\mathbb{R}) \quad (4.13)$$

Pour rappel, $L^p(\mathbb{R})$ est la notation standard pour l'ensemble des fonctions dont la puissance p -ième est sommable, c'est à dire $\{f \mid \int_{\mathbb{R}} |f(t)|^p dt < +\infty\}$.

Cette analyse fréquentielle par utilisation de la transformée de Fourier et l'analyse du spectre des signaux a permis de nombreuses avancées sur la connaissance de divers phénomènes et a révolutionné les télécommunications. Cependant, l'analyse fréquentielle par la transformée de Fourier se limite aux fonctions dont le contenu fréquentiel ne varie pas au cours du temps. L'exemple classique pour montrer les résultats dans de tels cas est de construire un signal $f(t) = \cos\left(\frac{1}{3}\pi t\right) \mathbb{1}_{[1,2]} + \cos\left(\frac{2}{3}\pi t\right) \mathbb{1}_{[3,4]}$. Une analyse globale du signal telle que celle effectuée par la transformée de Fourier ne fera pas ressortir deux pics du fait de la non stationnarité du signal de départ.

Pour résoudre ce problème, des analyses temps-fréquences ont été introduites. L'analyse temps-fréquences consiste à analyser le signal avec des fonctions (atomes) localisées en temps et/ou en fréquence. La première analyse temps-fréquences s'est faite à l'aide des atomes de Gabor, famille d'atomes $\{w_{\omega,b}\}_{(\omega,b) \in \mathbb{R}^2}$, définis par :

$$w_{\omega,b}(t) = e^{-\pi(t-b)^2} + e^{+2i\pi\omega t} \quad (4.14)$$

Ces atomes représentent les exponentielles complexes de l'analyse de Fourier classique modulées et atténuées par des enveloppes exponentielles. Cependant, dans cette analyse, lorsque ω augmente pour une localisation précise (fixée par b), le nombre d'oscillation augmente mais la résolution temporelle reste tout de même fixe.

Pour pallier à ce problème, la transformée en ondelettes permet une analyse semblable mais avec une grande gamme de résolutions temporelles. À la fin des années 1970, Jean Morlet a mis au point les bases de la transformée en ondelettes, qui s'appuie sur une ondelette mère, ψ . Celle-ci est une fonction du temps appartenant à $L^1 \cap L^2$. L'analyse en ondelette va s'effectuer en translatant (d'un facteur réel b) et en dilatant (d'un facteur d'échelle a positif) cette ondelette mère :

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad a \in \mathbb{R}^{+*}, b \in \mathbb{R} \quad (4.15)$$

Puis en utilisant ces différents atomes comme base de projection, on calcule les différents coefficients de la transformée :

$$C_f(a,b) = \int_{-\infty}^{+\infty} f(t) \psi_{a,b}(t) dt = \langle f, \psi_{a,b} \rangle \quad (4.16)$$

L'ensemble des coefficients $C_f(a,b)$ représente la transformée en ondelettes du signal f . Les différents coefficients représentent les fluctuations et le contenu fréquentiel de la fonction f à proximité de la position temporelle b (proximité donnée par le support de ψ), et à l'échelle a . Du fait de ces analyse locales des fonctions, on peut alors imaginer effectuer de la reconnaissance de formes avec les ondelettes. Ainsi, [Tang et al., 2000] présente entre autre la reconnaissance de caractères écrits à la mains et imprimés dans diffé-

rents langages à l'aide de B-splines, [Lai *et al.*, 1999] présente une reconnaissance de visage, [Haley et Manjunath, 1999] a mis au point un algorithme de classification de texture avec une propriété d'invariance en rotation alors que [Selin *et al.*, 2007] présente de la classification de bruits d'oiseaux.

Cette reconnaissance de formes peut se faire de deux manières. La première est l'utilisation d'une ondelette classique, et la classification des différents motifs à reconnaître en se basant sur le résultat de la transformée avec cette ondelette mère. La seconde est la construction d'ondelettes adaptées, permettant de mieux détecter mais également de mieux localiser la forme voulue [Misiti *et al.*, 2003]. Cette ondelette mère va permettre d'isoler, à l'intérieur d'un signal bruité et contenant plusieurs motifs, l'ensemble des échelles (a) et des positions temporelles (b) où apparaissent chacun des motifs.

La construction de telles ondelettes mères doit respecter la condition donnée précédemment, que la fonction ψ appartienne à l'ensemble $L^1 \cap L^2$. Afin que la transformée soit inversible, une condition sur la transformée de Fourier de ψ est également nécessaire :

$$\int_{\mathbb{R}^-} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega = \int_{\mathbb{R}^+} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < +\infty \quad (4.17)$$

La condition 4.17 peut se réécrire par les conditions suffisantes suivantes (dans notre cas, c'est à dire le cas d'ondelettes mères réelles) :

$$\int_{\mathbb{R}} \psi(t) dt < +\infty \text{ et } \int_{\mathbb{R}} |t\psi(t)| dt < +\infty \quad (4.18)$$

Ces deux conditions montrent que $\hat{\psi} \in C^1$, avec $\hat{\psi}(0) = 0$. À partir de ces conditions, la construction de l'ondelette mère se fait par approximation de la fonction donnée. Cette approximation est calculée au sens des moindres carrés et se fait en considérant la forme initiale comme une somme de fonction :

$$\psi(t) = \sum_{i=0}^{N-1} \alpha_i \cdot f_i(t) \quad (4.19)$$

L'approximation au sens des moindres carrés donnera l'ensemble des coefficients α_i approximant au mieux la forme. Les fonctions f_i sont choisies afin d'approximer au mieux le signal initial et de respecter la « forme » de celui-ci. Dans les constructions présentées par la suite, nous utilisons trois formes différentes de fonctions d'interpolation (avec N le nombre de fonctions utilisées pour l'approximation) :

- les polynômes, $f_i(t) = t^i$, famille génératrice de l'espace $\mathbb{P}[X]$ des polynômes à une inconnue.
- les polynômes de Lagrange :

$$f_i(t) = \prod_{j=1, j \neq i}^n \frac{X - x_j}{x_i - x_j}$$

avec $x_k, k \in [1; n]$ l'ensemble des abscisses auxquelles est donnée la forme de départ.

- des cosinus,

$$f_i(t) = \cos\left(\frac{2\pi i t}{x_n - x_1}\right)$$

Le choix de l'une ou l'autre des formes ainsi que du nombre de fonction N se fait en considérant la forme à approximer ainsi que le degré maximum rendant compte de la forme du signal sans pour autant ajouter trop d'artefacts d'une part, ni obtenir une matrice singulière dans la résolution du système linéaire. Pour respecter les conditions données précédemment, le signal est centré avant interpolation, afin de donner une moyenne nulle au signal interpolé.

Une fois cette ondelette mère créée, la transformée en ondelette continue d'un signal dans lequel la forme doit être recherchée, est calculée, et l'on recherche les maximums sur la surface des coefficients (en deux dimensions suivant a et b). Ces zones (traçables sous forme de contours) vont nous donner la localisation en temps et en fréquence des maximums détectés pour chacune des formes. En recherchant les maximums locaux au même endroit avec les différentes ondelettes (construits à partir des accéléromètres et des magnétomètres), nous obtenons une décision sur la position et la présence ou non d'une réalisation des signaux recherchés. Nous allons présenter par la suite la construction des ondelettes pour les différentes formes que nous recherchons.

4.2.4.2 Réalisation des motifs de détection

Dans les paragraphes suivants, nous allons illustrer la réalisation des modèles pour chacun des quatre mouvements détectés, et vérifier que le calcul de l'angle du quaternion est bien invariant lors d'un changement de position de la personne par rapport au nord magnétique. Pour ceci, les séquences debout-assis et assis-debout seront effectuées sur deux chaises identiques placées à 90° l'une de l'autre. Pour illustrer la pertinence des modèles, les figures contiendront le modèle créé pour un mouvement donné puis une réalisation de ce modèle tirée au hasard (et différente de celles utilisées pour la création du modèle).

Les modèles sont construits en utilisant la meilleure approximation, trouvée grâce à l'observation de plusieurs réalisations du signal, parmi celles décrites précédemment. Ce choix fait en sorte de conserver les formes importantes du signal. Les transferts Assis-Debout et Debout-Assis se passent principalement dans le plan antéropostérieur ($\mathbf{x}, \mathbf{z}$). L'accéléromètre sera donc étudié uniquement dans ce plan. Pour les deux autres mouvements, une partie de ceux-ci se passe suivant l'axe $\mathbf{y}$ qui sera alors pris en considération. Pour ce qui est des magnétomètres, il vont nous indiquer l'angle principal de la rotation.

4.2.4.2.1 Séquence Debout-Assis

Le premier mouvement illustré est montré par les figures 4.14 et 4.15.

Ces figures illustrent deux points cruciaux :

- Le premier, montré par la figure 4.14 est que le modèle est bien conforme à une autre réalisation du mouvement, tirée au hasard. Ce point est à la base de tout algorithme de reconnaissance de forme.

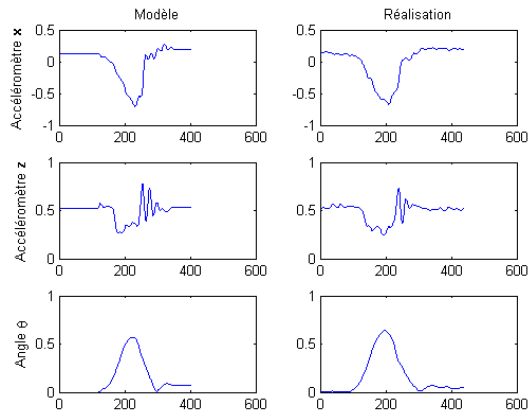


FIG. 4.14 – Séquence Debout-Assis sur une première chaise. À gauche le modèle utilisé pour la classification ultérieure, à droite une réalisation du mouvement. Les abscisses représentent le numéro d'échantillon et l'ordonnée la valeur en g pour les accéléromètres et en radian pour les angles

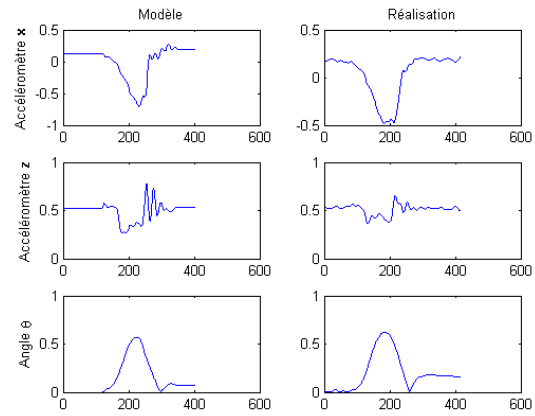


FIG. 4.15 – Séquence Debout-Assis sur une seconde chaise (à 90°). À gauche le modèle utilisé, à droite une réalisation du mouvement

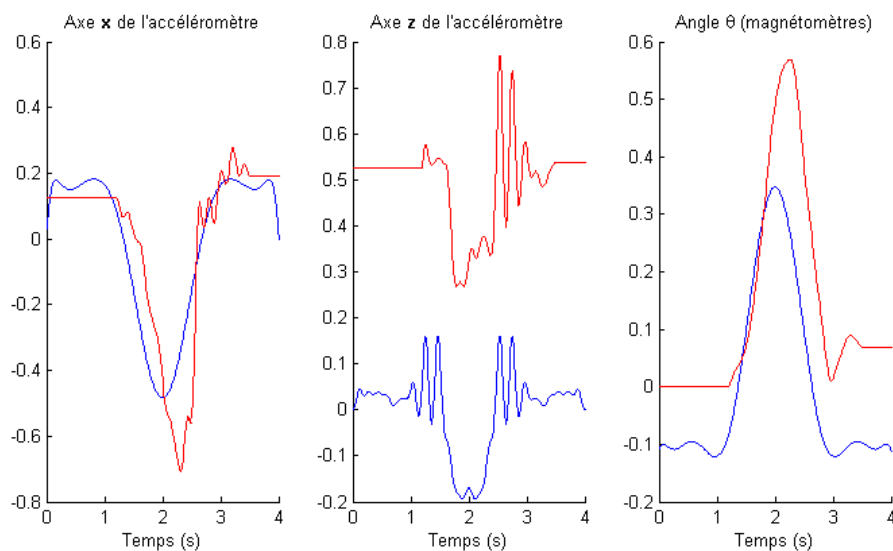


FIG. 4.16 – Construction des ondelettes adaptées pour la séquence Debout-Assis. En rouge le signal d'origine, en bleu l'ondelette correspondante respectant les critères d'admissibilité

- Le second est que, comme attendu, la transformée dans le domaine des quaternions du vecteur donné par les magnétomètres amènent à une forme répétable que l'on retrouve sur deux mouvements effectués avec des orientations différentes par rapport au nord magnétique. Cela nous permet d'utiliser le signal des magnétomètres pour améliorer la détection faite avec celui des accéléromètres. Ce signal des magnétomètres nous donne le mouvement de penché en avant puis retour en arrière que nous effectuons dans un mouvement Debout-Assis. En conjonction avec l'accélération créée sur les accéléromètres, cela nous permet de se distinguer du mouvement « se pencher en avant » (pour récupérer un objet) puis retour à la position verticale.

La construction des ondelettes adaptées est montrée sur la figure 4.16. Comme nous le voyons, les ondelettes adaptées conservent les formes du signal pour les détecter de nouveau par la suite. Elles ne conservent par contre pas l'échelle en ordonnée, ceci s'explique par les conditions à respecter. Cependant, cela n'a aucune incidence sur la détection future.

4.2.4.2.2 Séquence Assis-Debout

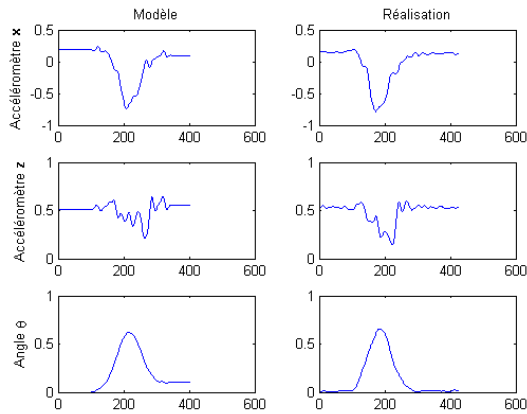


FIG. 4.17 – Séquence Assis-Debout sur la première chaise. À gauche le modèle utilisé, à droite une réalisation du mouvement.

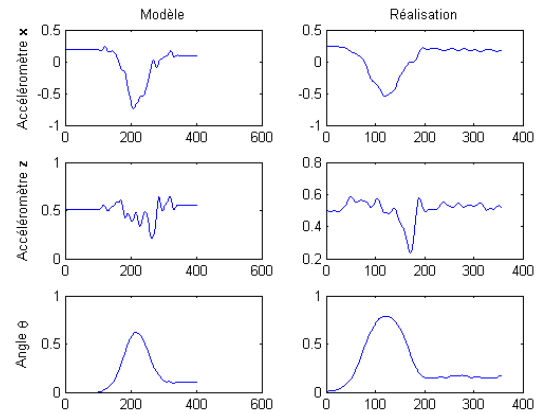


FIG. 4.18 – Séquence Assis-Debout sur la seconde chaise. À gauche le modèle utilisé, à droite une réalisation du mouvement.

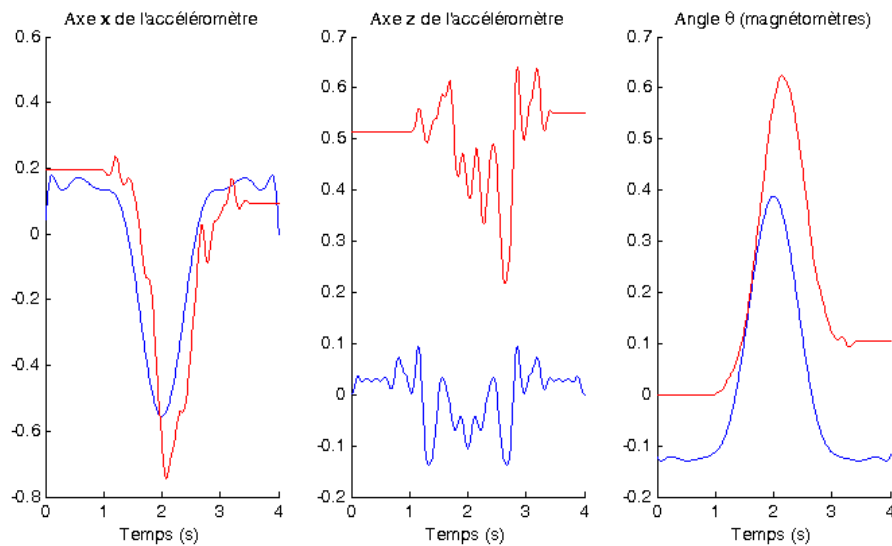


FIG. 4.19 – Construction des ondelettes adaptées pour la séquence Assis-Debout. En rouge le signal d'origine, en bleu l'ondelette correspondante respectant les critères d'admissibilité

Les figures 4.17 et 4.18 illustrent quant à elles le mouvement Assis-Debout.

De même que précédemment, ces figures montrent bien une indépendance en fonction de la position par rapport au Nord magnétique mais également une répétabilité des critères choisis. Les algorithmes de détection pourront se comporter de manière correcte sur de tels signaux.

La figure 4.19 montre le résultat, après approximation, de la construction des ondelettes adaptées pour ce mouvement. Ces trois ondelettes sont admissibles et pourront être utilisées pour calculer une transformée en ondelettes continue sur notre signal.

4.2.4.2.3 Séquence Debout-Allongé

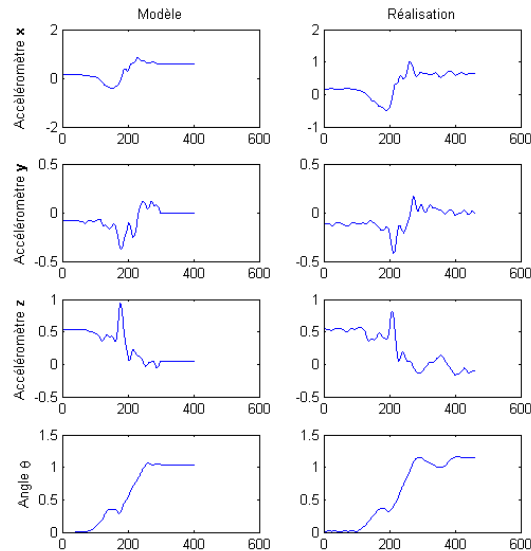


FIG. 4.20 – Séquence Debout-Allongé. À gauche le modèle utilisé, à droite une réalisation du mouvement.

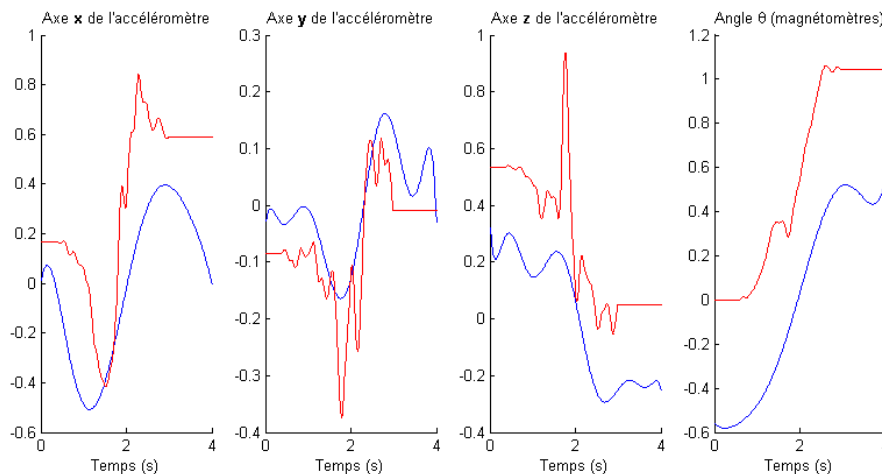


FIG. 4.21 – Construction des ondelettes adaptées pour la séquence Debout-Allongé. En rouge le signal d'origine, en bleu l'ondelette correspondante respectant les critères d'admissibilité

En réalisant le même travail sur la séquence Debout-Allongé, nous obtenons la figure 4.20. Les deux précédentes figures ayant illustré l'invariance en fonction de l'orientation du porteur, cette figure présente une seule séquence Debout-Allongé tirée au hasard et confrontée avec le modèle utilisé. De même que précédemment, nous obtenons bien un modèle

répétable sur les critères choisis et nous pouvons ainsi envisager une classification correcte de cette activité.

La figure 4.21 présente la construction, par approximation, des ondelettes adaptées, admissibles pour la transformée en ondelettes continue, de ce mouvement Debout-Allongé. Les remarques à propos du respect de la forme et du non-respect des échelles sont identiques aux paragraphes précédents, pour les mêmes raisons mathématiques.

4.2.4.2.4 Séquence Allongé-Debout

Enfin, le dernier mouvement à détecter est le mouvement Allongé-Debout. Celui-ci est présenté, par l'intermédiaire de son modèle confronté à une autre réalisation tirée au hasard, sur la figure 4.22. Cette dernière figure nous montre également la répétabilité du mouvement.

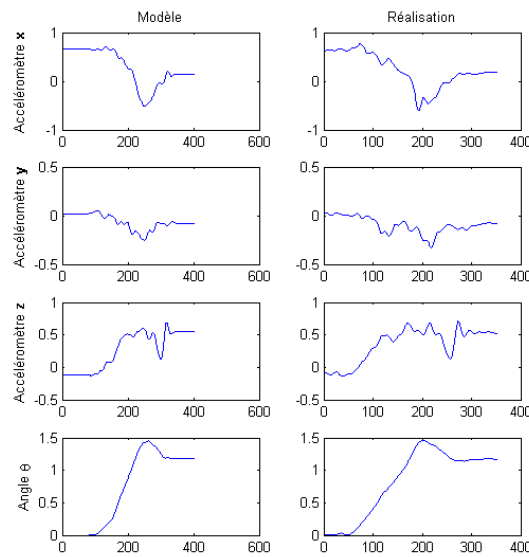


FIG. 4.22 – Séquence Allongé-Debout. À gauche le modèle utilisé, à droite une réalisation du mouvement.

Identiquement aux paragraphes précédents, la figure 4.23 montre l'approximation réalisée pour la construction des ondelettes admissibles.

4.2.5 Périodes de marche

4.2.5.1 Détection des périodes de marche

Les articles de la littérature ont montré que les périodes de marche pouvaient se détecter avec la transformée en ondelettes (discrète et continue) ou bien avec la transformée de Fourier à court terme (analyse fréquentielle). En effet, pour un accéléromètre placé sur le buste de la personne, [Bouten *et al.*, 1997] a montré que la marche était perçue comme un mouvement oscillatoire dans la bande de fréquence $[0.8; 5\text{Hz}]$. La figure 4.24 nous montre le signal

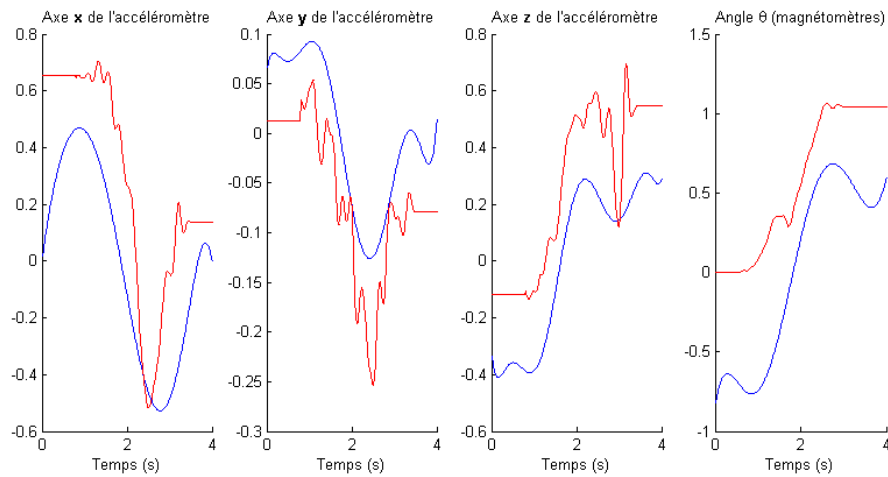


FIG. 4.23 – Construction des ondelettes adaptées pour la séquence Allongé-Debout. En rouge le signal d'origine, en bleu l'ondelette correspondante respectant les critères d'admissibilité

de marche pour une expérimentation qui a consisté à faire dix aller-retours (d'une distance approximative de 15m) à vitesse lente puis dix aller-retours à vitesse rapide.

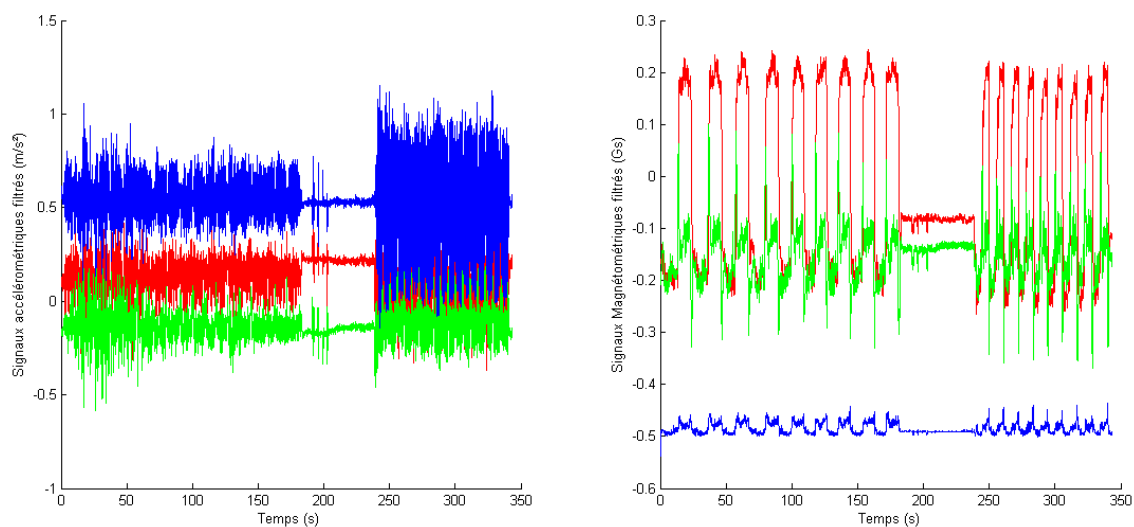


FIG. 4.24 – Signaux filtrés des accéléromètres à gauche et des magnétomètres à droite correspondant à une séquence de dix aller-retours (distance de 15m) à marche lente puis dix à marche rapide

Nous avons choisi de nous baser sur le critère fréquentiel et d'analyser et de détecter les périodes de marche à l'aide de la transformée de Fourier à court terme. Cette transformée sera calculée sur la norme du signal accélérométrique, norme qui représente bien les oscillations puisque le mouvement effectué par le torse va engendrer une accélération se superposant à celle de la pesanteur.

Le calcul de la transformée de Fourier à court terme se fait en calculant la transformée de Fourier classique mais sur une fenêtre temporelle de notre signal. Cette fenêtre temporelle pourrait très bien être réalisée en multipliant notre signal par une porte ($\mathbb{1}_{[a,b]}$, $(a, b) \in \mathbb{R}^2$).

Cependant, la transformée de Fourier d'une multiplication est donnée par la convolution des transformées de Fourier de chacune des fonctions dans le domaine des fréquences. La porte ayant comme transformée de Fourier un sinus cardinal, il entraînerait des rebonds dans le domaine fréquentiel. Pour résoudre ce problème, on utilise en général une fenêtre d'apodisation aux variations plus lentes aux limites. Nous avons utilisé une fenêtre de Hamming, dont l'amplitude du premier rebond est extrêmement faible. Cette fenêtre a pour équation

$$w(t) = 0,54 - 0,46 \cos\left(\frac{2\pi t}{N}\right), \text{ N longueur de la fenêtre} \quad (4.20)$$

Étant donné qu'une partie du signal fenêtré est atténuée, la fenêtre de calcul glissante de la transformée de Fourier inclut une zone de recouvrement pour ne perdre aucune information.

Cette transformée de Fourier est calculée sur des fréquences linéairement espacées dans la gamme [0.8; 5Hz] (M fréquences). La transformée de Fourier est alors donnée par :

$$X(\omega_j, \tau) = \int_{\mathbb{R}} x(t) w(t - \tau) e^{-j\omega_j t} dt, \quad \tau \in \mathbb{R}, \omega_j \in [0.8; 2\text{Hz}] \quad (4.21)$$

La décision est prise sur le dépassement d'un seuil sur ce calcul, noté s_{marche} . Le signal de sortie, sera alors décrit par :

$$S(\tau) = \begin{cases} 1 & \text{si } \sum_{j=1}^M X(\omega_j, \tau) > s_{marche} \\ 0 & \text{sinon} \end{cases} \quad (4.22)$$

C'est à dire si le contenu fréquentiel dans la bande de fréquence considérée est suffisant pour dépasser un seuil donné (et fixé par des expérimentations préalables).

Les résultats de cette détection dans deux scénarios incluant des séquences de marche différentes est donné par la figure 4.25. Nous voyons bien les passages à 1 du signal nous montrant les différentes périodes de marche.

4.2.5.2 Périodes de marche et segmentation

Étant donné l'utilisation visée de ces données actimétriques (analyse sur de longues périodes des activités de la vie quotidienne d'une personne), nous avons choisi de réaliser cette détection des périodes de marche sur des fenêtres temporelles de cinq secondes, afin de réduire les quantités de données à traiter et de ne prendre que les données les plus significatives. Cependant, cette granularité du signal de la marche a un effet sur la détection des mouvements. En effet, la classification n'est pas effectuée si la période considérée est de la marche (car dans ce cas la posture est debout). Cette fenêtre de cinq secondes, si elle est mal positionnée, peut avoir un effet de masquage de certains mouvements. Ainsi, sur les longues périodes de marche comme la figure 4.25 en présente, les mouvements Debout-Assis et Assis-Debout au début des périodes de marche sont « masqués » par la marche dans la même fenêtre temporelle (les expérimentations étant faites à vitesse normale, la marche

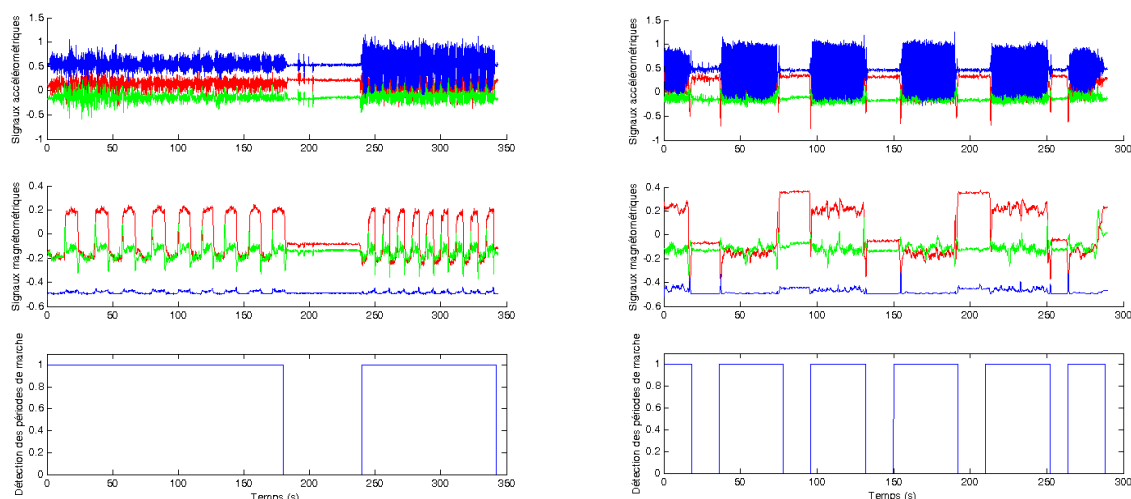


FIG. 4.25 – Résultats de la détection de la marche sur deux séquences. À gauche, une séquence de dix aller-retours d’une distance de 15m environ, d’abord lents, puis rapides. À droite, une séquence de cinq sessions de marche (plus deux sur le début et la fin pour se rendre jusqu’à la chaise), chaque séquence étant composée de 80m de marche suivie d’un transfert Debout-Assis

débute juste après le lever, rapidement effectué).

Pour résoudre ce problème, l’analyse de fenêtres de marche est toujours considérée sur la même période de cinq secondes, mais les deux premières secondes et les deux dernières secondes de chaque période de marche (donc si une période de marche s’étend sur huit fenêtres, on ne considère que la première et la huitième), est analysée plus précisément pour y détecter le début réel de la marche et pour vérifier si un changement de posture y est présent. Le reste des périodes de marche n’est par contre jamais analysé par la classification. Cela permet de réduire les non-détections de changements de posture mais également de ne pas limiter les changements de posture détectés à de « faux » changements qui interviendraient après avoir fait une pause en arrivant devant la chaise par exemple. La figure 4.28 présente ainsi le résultat des algorithmes sur un scénario de Up&Go, susceptible de présenter les défauts corrigés par l’algorithme de détection de la marche.

4.3 Résultats sur quelques scénarios

L’implémentation MatlabTM de ces algorithmes permet de visualiser les postures et la marche aux différents moments du signal. Elle donne également un fichier de sortie, utilisable par la classification, qui sera décrit par la suite. La figure 4.26 montre le résultat sur un scénario de dix séquences Debout-Allongé-Debout. La figure présente les signaux filtrés, le résultat de la détection de la marche et le résultat de la segmentation puis de la classification. La classification donne, à un instant t , la posture de la personne (debout, assis, allongé). On peut y voir la bonne segmentation puis la bonne classification des différentes fenêtres de changement de posture.

La figure 4.27 suit la même présentation pour, cette fois-ci, un scénario Debout-Assis-

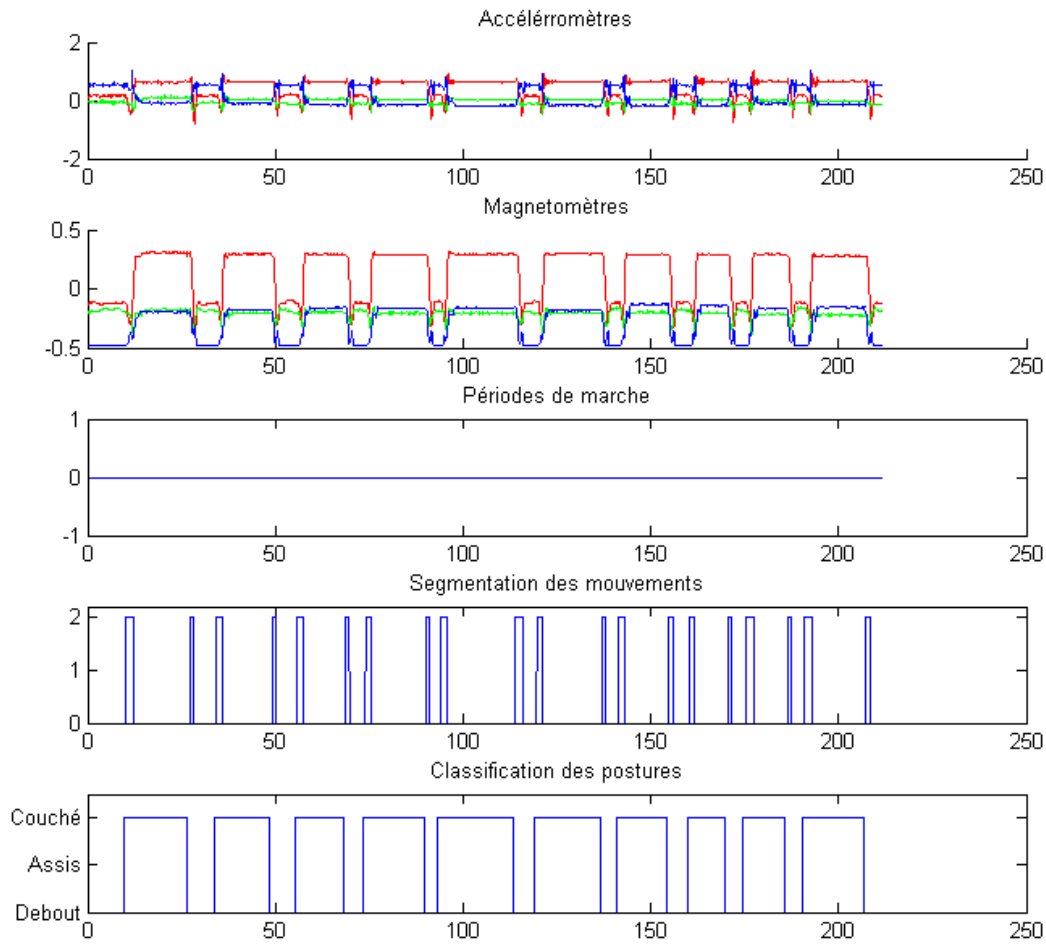


FIG. 4.26 – Résultats de l'algorithme sur un scénario de dix séquences de Debout-Allongé-Debout

Debout répété dix fois. De même, il ressort de ce scénario (réalisé sur la chaise numéro 2, à 90° de celle sur laquelle s'est faite l'apprentissage), une bonne classification des fenêtres des changements de posture pour s'asseoir.

Enfin, la figure 4.28 présente un scénario de Up&Go qui, rappelons-le, est un scénario dans lequel la personne est assise, se lève, marche, fait un demi-tour, marche puis fait demi-tour devant la chaise avant de s'asseoir. Dans ce scénario, les périodes de marche sont importantes (la marche s'est faite sur une distance de quatre-vingt mètres), et le changement de posture est « collé » temporellement au début et à la fin de la marche. Dans ce cas, la correction des fenêtres de marche, pour y détecter le réel début et la réelle fin, est utile. La segmentation fait également dans ce cas ressortir des erreurs (dues par exemple à l'élan donné par la marche qui fait que le passage assis-debout se fait avec une plus grande accélération). Ces erreurs sont, comme nous le voyons, corrigées par la classification.

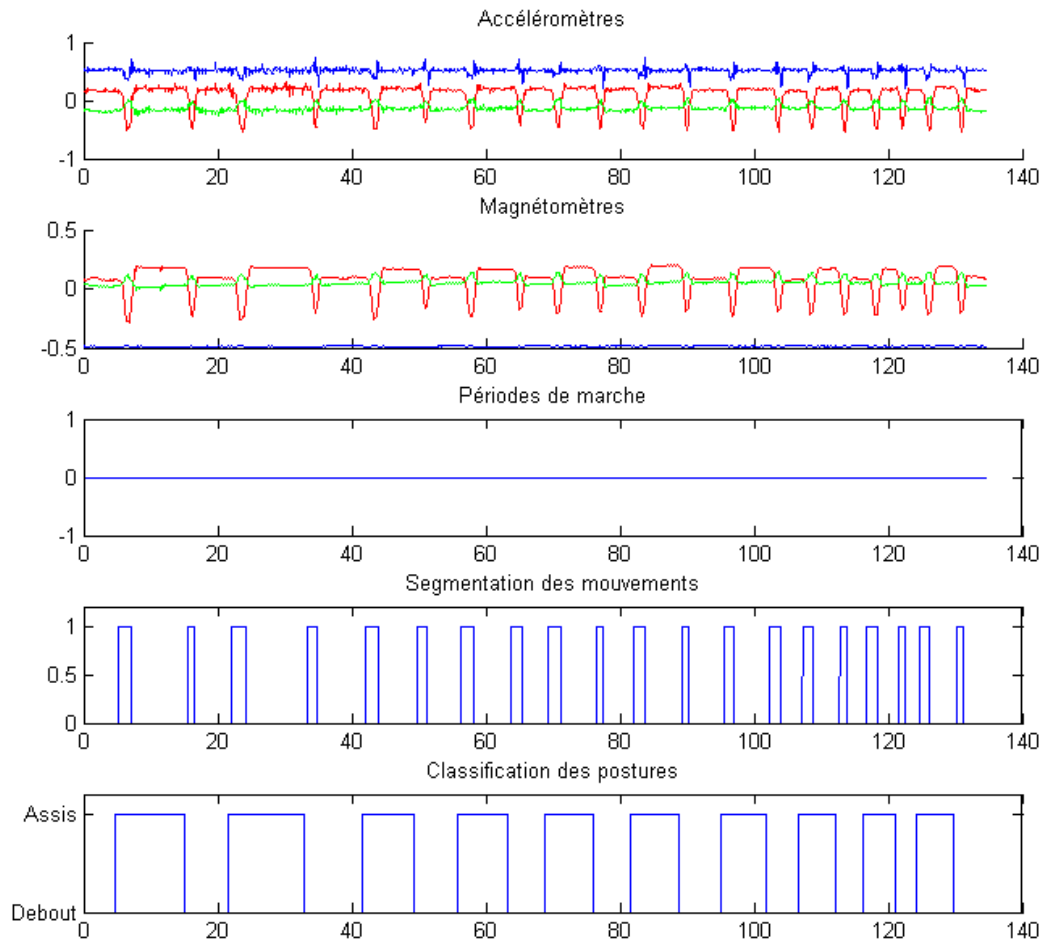


FIG. 4.27 – Résultats de l'algorithme sur un scénario de dix séquences Debout-Assis-Debout

4.4 Fichier de sortie et résultats finaux

Le fichier de données en sortie du système d'analyse de la marche et des postures va contenir les informations sur les actions effectuées par le sujet et détectées par la carte ACTIM6D. Ce fichier va être ensuite utilisé pour la classification des activités. Il contient donc principalement les informations sur les changements de posture et sur les périodes de marche de la personne. Cela permettra au système d'extraire le temps passé à marcher et le temps passé dans chaque position (debout, assis, couché). La première colonne décrit le type de détection qui a eu lieu, P pour Posture et M pour Marche. Dans le cas de la marche, la seconde et la troisième colonne donnent l'échantillon de départ et l'échantillon final à laquelle l'action a été détectée. Pour le cas d'une posture, la seconde colonne donne le début de la détection du changement de posture (en échantillon toujours) puis la posture de départ (D pour Debout, AS pour Assis et AL pour Allongé), et la posture d'arrivée. Il revient ensuite au système de classification de s'occuper de la synchronisation de ces données avec les autres capteurs.

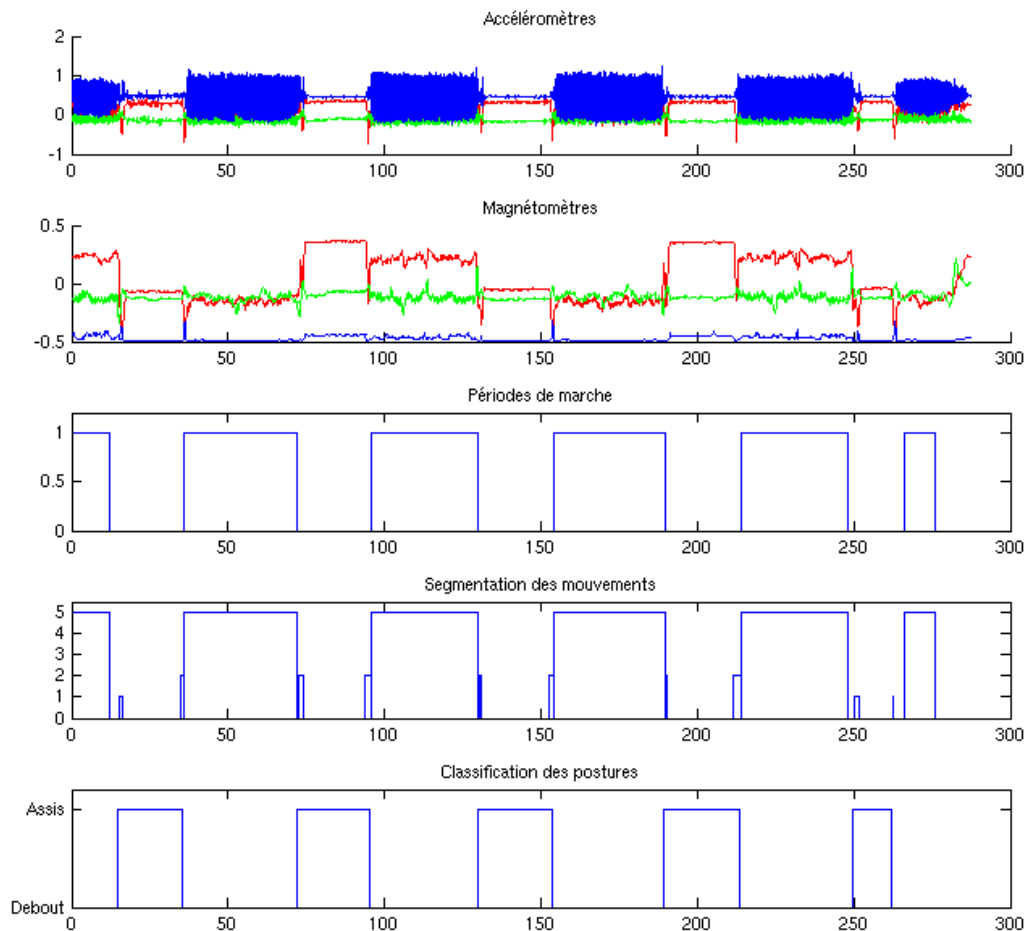


FIG. 4.28 – Résultats de l'algorithme sur un scénario de Up&Go

Le début du fichier, pour la succession de Up&Go, se présentera ainsi (en considérant toujours que l'allumage du capteur se fait debout) :

M	0	1198		
P	1483	D	AS	
P	3549	AS	D	
M	3599	7198		
P	7207	D	AS	

4.5 Expérimentations et résultats

4.5.1 But de l'expérimentation et population

L'expérimentation proposée a pour but de caractériser la sensibilité et la spécificité des algorithmes proposés ci-dessus. Cette expérimentation visait donc à la réalisation de plusieurs itérations des transferts considérés dans ce chapitre, ainsi que de séquences de marche

et de mouvements de la vie quotidienne (ramasser un objet et refaire son lacet). Ces derniers mouvements sont utilisés pour connaître la spécificité des algorithmes. En effet, les mouvements cités vont tous deux provoquer une importante amplitude sur l'accéléromètre, classant la fenêtre considérée comme l'un des transferts posturaux. La transformée en ondelette ne devrait détecter aucune des formes dans ce mouvement et ne pas le classer, le considérant comme un faux positif de la segmentation. Le scénario a été effectué par douze sujets jeunes et sains. Ces douze sujets font partie de ceux ayant fait l'expérimentation qui sera présentée par la suite, pour valider la classification des activités. L'ensemble de ces personnes était âgé en moyenne de $30,4\text{ans} \pm 5,9$ (maximum 43 ans, minimum 24 ans), de taille moyenne $1,76\text{m} \pm 0,08$ (taille maximum : 1,92m, taille minimum 1,62m), pour un poids moyen de $69\text{ kg} \pm 7,42$ (pour un maximum de 80 kg et un minimum de 57 kg). La répartition par sexe était de six femmes et six hommes. Notre population initiale était donc assez diversifiée.

4.5.2 Protocole expérimental

La séance expérimentale est filmée en vue d'une indexation. La séquence commence par la réalisation d'un saut vertical (permettant la synchronisation des signaux avec la vidéo), puis le sujet s'assoit. Une fois assis, il lui sera demandé de se lever, de marcher quatre mètres, de faire demi-tour, de marcher de nouveau quatre mètres, faire demi-tour devant la chaise puis se rasseoir. Cette succession (appelée Up&Go dans la littérature) sera répétée cinq fois.

Le mouvement suivant est le coucher et le lever. Pour tester ces mouvements, le sujet démarre debout, marche trois mètres, s'allonge sur le lit et y reste allongé quelques secondes (il peut remuer), puis se relève. Cette séquence est également répétée cinq fois.

Enfin, la dernière séquence est de refaire son lacet puis de ramasser un objet à terre.

Cet ensemble de mouvements est effectué sans aucune contrainte de temps ou de réalisation. Les mouvements sont effectués à la vitesse de confort de la personne. La chaise est une chaise classique, ferme, sans accoudoir, de hauteur de siège 45cm et d'angle entre le siège et le dossier de 10° . Le lit est également un lit classique, une personne, ferme et de hauteur 46cm .

4.5.3 Résultats

Le tableau 4.1 et 4.2 décrivent les résultats de ces expérimentations.

	Classification Incorrecte	Classification Correcte
Segmentation Incorrecte	8% (10)	10,4% (13)
Segmentation Correcte	6,4% (8)	75% (93)

TAB. 4.1 – Matrice de confusion pour les mouvements Assis-Debout et Debout-Assis

Ces tableaux nous donnent des indications sur le taux d'erreur mais aussi sur le type d'erreur commis par l'algorithme de segmentation et de classification. Les chiffres indiqués

	Classification Incorrecte	Classification Correcte
Segmentation Incorrecte	13,4% (15)	17,8% (20)
Segmentation Correcte	4,4% (5)	64,2% (72)

TAB. 4.2 – Matrice de confusion pour les mouvements Allongé-Debout et Debout-Allongé

entre parenthèses représentent les nombres d'exécution correspondant aux cas respectifs. Les quatre cas présentés ci-dessus dans les deux tableaux correspondent à différents cas :

- Segmentation incorrecte et classification incorrecte. Dans ce cas, un transfert postural a été effectué, et celui-ci a été mal segmenté (donc mauvaise décision sur la segmentation ou non détection du mouvement), et mal classifié, c'est à dire qu'au final aucun transfert postural n'a été détecté.
- Segmentation incorrecte mais classification correcte. Ce terme de classification correcte doit être expliqué. Il signifie que la classification a détecté un transfert postural à l'endroit où il y en avait un dans la réalité, mais que ce transfert postural n'est pas correct car le module de segmentation a proposé la mauvaise décision.
- Segmentation correcte mais classification incorrecte. Dans ce cas, la segmentation est correcte, et a détecté le bon mouvement, par contre la classification a rejeté la présence d'une fenêtre de transfert postural à cet endroit.
- Dans le dernier cas, autant la segmentation que la classification sont correctes, la décision prise correspond à la réalité.

Le taux d'erreur global est acceptable dans les deux types de transferts posturaux, avec un taux de bonne détection de 65% dans le cas du transfert assis-debout et debout-assis et 75% dans le cas des transferts posturaux debout-allongé et allongé-debout. Il est à noter que l'ensemble des modèles et paramètres de l'application n'ont pas été modifiés pour traiter les données de ces sujets et que l'apprentissage s'est fait sur un ensemble de données n'appartenant pas à cette expérimentation.

Les différents taux d'erreurs s'expliquent également par un phénomène que nous avons beaucoup rencontré lors du traitement de cette expérimentation : la mauvaise segmentation. En effet, lors de l'expérimentation, il était demandé à la personne de s'asseoir puis de faire ses 5 Up&Go départ assis. Ce que nous avons noté est que, sans consigne supplémentaires, la personne restait assise un minimum (alors que nous avons demandé plusieurs secondes), et se relevait très vite, non consciente du temps qui passe. Ainsi, nous arrivons à des données qui ne sont pas vraiment représentatives d'activités de la vie quotidienne du fait que parfois, le passage assis-debout se fait une seconde après le passage debout-assis, trompant ainsi l'algorithme de segmentation qui découpe mal le signal et n'en fait qu'une fenêtre unique. Ces cas se retrouvent principalement dans la case SI, CI de notre tableau, qui est dans les deux cas l'erreur la plus fréquente. En effet, nous nous retrouvons, dans ce cas, avec une mauvaise segmentation du signal et une absence de prise de décision étant donnée que celle-ci a déjà été prise dans la fenêtre.

Nous pourrions très bien rechercher un second maximum dans la fenêtre, cependant, il

serait nécessaire d'inclure encore des conditions entre les deux maximums, en effet la transformée en ondelette peut donner un second maximum assez proche du premier mais distinct. Celui-ci sera plus faible, mais selon la forme de l'ondelette ce cas est tout à fait possible. Cela aurait plus pour effet de dégrader les performances que l'effet inverse.

4.6 Conclusion et discussion sur l'implémentation de la carte ACTIM6D

Ce chapitre présente la réalisation d'un système électronique embarqué autonome permettant un suivi cinématique de l'activité de la personne. Les résultats des tests nous amènent à réfléchir sur cette implémentation et sur les améliorations pouvant y être apportées par la suite. Cette section indique des pistes de réflexion sur les données cinématiques et l'amélioration de leur traitement.

4.6.1 Limitations matérielles

4.6.1.1 Bruit de mesure

Le bruit de mesure sur les accéléromètres est relativement faible et le moyen le plus efficace de l'améliorer serait de remplacer le convertisseur intégré au microcontrôleur (convertisseur à approximations successives) par un convertisseur de type sigma delta donnant une mesure plus fine (mais plus lente à obtenir). En ce qui concerne les magnétomètres, ils sont quant à eux assez sensibles à l'environnement électromagnétique très proche de la personne. Ainsi, ces magnétomètres sont difficilement utilisables dans un ascenseur ou dans une voiture par exemple. Pour la mesure au domicile, qui est notre but initial, ils sont par contre relativement adaptés et ne sont perturbés que lorsque le porteur s'approche à moins d'un mètre d'une cage d'ascenseur. L'élimination de ce type de bruit est par contre relativement délicat. En effet, le bruit n'est pas constant en énergie et le sujet va être mobile à sa proximité. Une piste possible, à explorer plus longuement, serait les algorithmes de séparation de sources, pour séparer ces deux signaux qui se trouvent dans la même bande de fréquences.

4.6.1.2 Positionnement du capteur

Le positionnement du capteur est un point qui mérite attention. En effet, les sujets étaient équipés du capteur, positionné dans une poche cousue sur le tee-shirt. Les algorithmes se basent sur le fait que la centrale crée une base en trois dimensions solidaire au sujet. Si le tee-shirt est trop ample, que le capteur effectue des mouvements indépendants, les algorithmes s'en trouvent perturbés. Dans ce cas, une ceinture doit maintenir le capteur afin que le changement de référentiel reste correct au cours du temps. Pour s'affranchir de tels problèmes, des solutions à base de patch seraient peut-être plus adaptées à notre utilisation.

De plus, les algorithmes présentés ne modifient pas la base de départ à chaque début d'expérimentation (qui nécessiterait d'être immobile pendant un certain temps). Selon les morphologies, la position du plan ($\mathbf{x}, \mathbf{z}$) peut se trouver légèrement modifiée. D'après nos expérimentations, ces légères modifications morphologiques n'ont pas d'influence sur la détection. Cependant un repositionnement des axes du capteur pourrait être fait pour prendre en compte des cas extrêmes, en se basant sur la mesure moyenne pendant cinq secondes du signal accélérométrique.

4.6.2 Logiciel

4.6.2.1 Intérêt des ondelettes

Le développement s'est axé sur les ondelettes pour tester leur adéquation avec le type de problème posé. Comme énoncé précédemment, les ondelettes permettent une analyse temps-fréquences du signal et la détection à l'aide de celles-ci permet de localiser le signal en temps et en échelle. Le temps nous permet de voir le moment auquel le signal se superpose au modèle, et l'échelle permet d'obtenir un critère de vitesse d'exécution de ce mouvement, connaissant celle du modèle. Ce paramètre n'est pas important pour la classification des activités présentée dans cette thèse. Cependant, il peut s'avérer intéressant de regarder l'évolution sur le long terme de la vitesse moyenne d'exécution de chaque types de mouvements afin, peut-être, de détecter l'apparition de certaines pathologies. Pour effectuer ce suivi, des mesures les plus longues possibles seront nécessaires à recueillir afin de voir l'influence et l'importance des variations de ces données au cours du temps.

Autre point, cet algorithme se base sur quatre mouvements, desquels il apprend les modèles, et qui sont ainsi facilement modifiables. Les étendre pourrait aussi permettre d'apprendre à détecter d'autre mouvements, cependant il faut trouver les paramètres les décrivant au mieux, puis vérifier que les classes sont bien séparables et trouver un paramétrage de la segmentation pour cette nouvelle classe. Ensuite, de ces modèles sont calculés des coefficients se basant sur la variance, puis la transformée en ondelettes détecte dans le signal des maximums sur toutes les ondelettes considérées. Ces maximums ne sont pas fixés par un paramètre, ils sont recherchés parmi les maximums locaux non négligeables du signal de la transformée (en comparaison au signal global). Cet algorithme est donc plus évolué que ceux calculant seulement des seuils sur l'angle ou l'accélération pour détecter les changements de posture.

4.6.2.2 Discussion sur la segmentation et la classification

La segmentation des données est un point critique pour beaucoup d'algorithmes de classification et notre algorithme ne déroge pas à cette règle. En effet, il est souvent impossible de tenter de classer l'ensemble des données et il est donc nécessaire de réduire la quantité de fenêtres à analyser en éliminant les zones dans lesquelles le signal ne possède pas les propriétés adéquates. En se basant sur l'écart-type de notre signal, nous avons tenté de réduire au minimum les non détections de signaux. Cependant, un mouvement qui serait

fait extrêmement lentement (exemple s'asseoir sur une chaise en dix secondes) perturberait l'algorithme et ne serait pas segmenté. Les paramètres de la segmentation sont relativement adaptés à des mouvements « normaux », effectués par des personnes saines.

En ce qui concerne la classification, les ondelettes donnent des résultats acceptables, surtout pour un mouvement compliqué tel que le décubitus. Cependant, les deux classes de signaux dépendent tout de même beaucoup de la segmentation pour être correctement séparées. En effet, le décubitus contient un mouvement debout-assis. Tenter de classer avec la classe « s'asseoir » un mouvement s'allonger donnera donc un maximum au début de la fenêtre d'analyse. Pour résoudre ce problème, une analyse mécanique plus poussée des deux mouvements et une séparation de ceux-ci (permettant de détecter le mouvement debout-assis, puis le décubitus, le tout à l'intérieur d'un mouvement debout-allongé) serait une piste de recherche permettant de corriger certaines mauvaises classifications.

4.6.2.3 Analyse temps-fréquences de la marche

La marche peut être analysée de plusieurs manières. Les analyses portent souvent sur le temps entre chaque pas (stride-to-stride) et son évolution lors de mesures effectuées sur le long terme [Hausdorff *et al.*, 2004]. Ces mesures peuvent être effectuées de plusieurs manières. Ainsi, les tapis de marche donnent ce type d'information. Des séances de marche non contraintes sur ce type de tapis peuvent donner des indications précises sur certains risques chez la personne. Les dispositifs accélérométriques peuvent également donner ce type d'information. Ainsi McRoberts, dans la centrale présentée précédemment, donne ces informations avec son accéléromètre posé sur le sacrum.

La base d'accéléromètres et de magnétomètres que nous possédons permet également ce type d'analyse, dans le plan temps-fréquences. Ainsi, lorsqu'une zone de marche est détectée, elle peut être analysée par une ondelette. La forme d'ondelette « le chapeau mexicain » a été retenue du fait qu'elle est représentative de l'évolution de la norme de l'accéléromètre dans le plan (x, z) .

La figure 4.29 nous montre la faisabilité de la détection, dans le plan temps-fréquences, de ces moments de la pose du pied. L'extraction des pics pourrait permettre une analyse sur le long terme de ces périodes de marche, et fait partie des pistes à explorer pour la suite des travaux avec ce circuit.

4.6.2.4 Perspectives

4.6.2.4.1 Correction des incohérences des données

L'analyse des données s'effectue sur une passe uniquement. Ainsi, les incohérences les plus grossières sont « corrigées ». Par exemple, la détection d'une longue fenêtre de marche alors que la personne est allongée donnera un changement de posture au moment du début de la marche, considérant que celui-ci a été manqué dans le signal. Cependant, pour effectuer des corrections plus poussées, un retour vers les précédents changements de posture pour une correction de ceux-ci pourrait apporter des informations plus précises (réaliser un

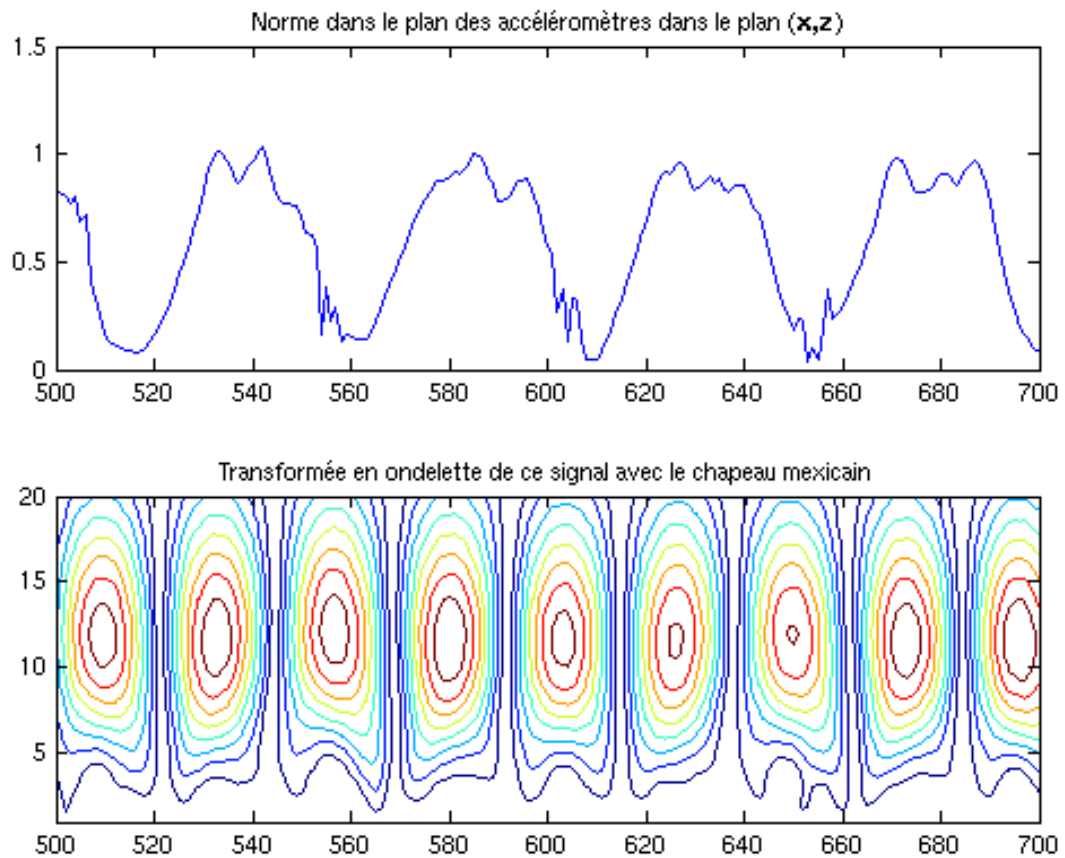


FIG. 4.29 – Transformée en ondelettes avec le chapeau mexicain du signal de marche des accéléromètres

système possédant un rebouclage en quelque sorte). Cela rendrait le traitement non causal mais améliorerait la classification.

4.6.2.4.2 Système embarqué?

Le système développé est embarqué sur la personne, cependant l'ensemble des calculs présentés ci-dessus est effectué hors-ligne. En effet, ce type de calculs est très gourmand en ressources et difficilement implémentable dans un processeur à architecture RISC cadencé à 4MHz. Un processeur plus rapide et munit d'un coprocesseur arithmétique pourrait aider à l'implémentation de tels algorithmes, cependant la consommation s'en ressentirait fortement. Pour le type de traitement fait dans cette thèse, une analyse hors-ligne des signaux n'est pas un handicap. Elle correspond également à ce qui est fait (avec des algorithmes différents) par les centrales présentées dans l'état de l'art.

Ces algorithmes fonctionnent de manière causale, par contre ils sont relativement difficiles à implémenter en embarqué en conservant une durée de vie de la pile suffisante. Une autre piste est de transmettre les données sans fil pour les traiter à la volée sur un ordinateur les recevant. La question à se poser avant de se lancer dans de tels projets est de savoir si l'analyse des signaux et des paramètres nécessite cet investissement ou si ce type de traite-

ment hors-ligne suffit aux besoins.

Les données sonores

Sommaire

5.1 Le son dans le HIS	101
5.1.1 Organisation globale du système de reconnaissance de sons et de la parole	102
5.1.2 Matériel utilisé	102
5.1.2.1 Microphones	102
5.1.2.2 Carte d'acquisition	102
5.1.2.3 Acquisition, traitement et stockage des données	102
5.1.3 Microphones : installation et réglages	103
5.1.3.1 Choix de l'emplacement des microphones	103
5.1.3.2 Réglage des gains	104
5.1.4 Acquisition des données sonores	105
5.1.5 Objectifs	105
5.2 Algorithmes de traitement des sons	106
5.2.1 Détection	106
5.2.2 Estimation du Rapport Signal sur Bruit (RSB)	107
5.2.3 Segmentation entre signal de parole et son de la vie courante	108
5.2.3.1 Modèles acoustiques	108
5.2.3.2 Segmentation parole/sons de la vie courante par classification	109
5.2.4 Classification des sons de la vie courante	109
5.2.5 Reconnaissance automatique de la parole	110
5.3 Bases d'apprentissage	110
5.3.1 Corpus des sons de la vie courante	110
5.3.2 Paroles	111
5.3.2.1 Construction des modèles de langage et des modèles acoustiques	111
5.3.2.2 Base de tests	112
5.4 Résultats en laboratoire	112
5.4.1 Détection des sons par calcul d'un arbre d'ondelettes	112
5.4.2 Segmentation parole/sons de la vie courante	113
5.4.3 Classification des sons	113
5.4.4 Reconnaissance des mots clés de détresse	114
5.5 Implémentation : le logiciel AuditHIS	114
5.5.1 Description générale	114
5.5.2 Données extraites	116
5.6 Exploitation des données	117
5.6.1 Extraction des données	117
5.6.2 Règles de classification	117

5.6.2.1	Parole	117
5.6.2.2	Sons de la vie courante	118
5.6.3	Fichier de sortie	118
5.7	Expérimentations préliminaires	118
5.7.1	Expérimentation 1	119
5.7.1.1	Protocole	119
5.7.1.1.1	Reconnaissance sonore	119
5.7.1.1.2	Reconnaissance de la parole	120
5.7.1.2	Résultats expérimentaux	120
5.7.1.2.1	La parole	120
5.7.1.2.2	Les sons de la vie courante	121
5.7.1.2.3	Résultats globaux	121
5.7.2	Expérimentation 2	122
5.7.2.1	Protocole	122
5.7.2.1.1	Reconnaissance sonore	122
5.7.2.1.2	Reconnaissance de la parole	123
5.7.2.2	Résultats expérimentaux	123
5.7.3	Expérimentation 3	124
5.7.3.1	Protocole	124
5.7.3.2	Résultats expérimentaux	124
5.8	Analyse des résultats de la détection du son et de la parole	126

DEPUIS la collaboration entre le CLIPS-IMAG, équipe GEOD¹ et le TIMC-IMAG, équipe AFIRM, dans le cadre des projets RESIDE/HIS (REconnaissance de Signaux de Détresse dans l'Habitat Intelligent Santé) et DESDHIS (DÉtection de Situations de Détresse dans un Habitat Intelligent Santé), des capteurs sonores sont installés et en tests dans le HIS de la faculté de médecine de Grenoble. Les travaux du CLIPS, et maintenant du LIG (équipe GETALP), ont permis d'élaborer des modèles en laboratoire pour la reconnaissance de sons et de la parole, et de tester ces modèles avec des signaux sources captés dans l'habitat, avec différents rapports signal sur bruit. Cette thèse a été l'occasion d'élargir cette collaboration et d'effectuer des tests en situation réelle afin de déterminer la sensibilité et la spécificité de ces capteurs lorsqu'ils sont intégrés à l'environnement de la personne.

L'objet de ce chapitre sera dans un premier temps de décrire les travaux effectués par l'équipe GETALP du LIG, les algorithmes, le programme réalisé pour les implémenter ainsi que les données utilisées pour l'apprentissage. Nous décrirons ensuite les méthodes utilisées pour la classification des sons et la réalisation d'un fichier de sortie contenant les données nécessaires pour la fusion de données. Nous décrirons alors les protocoles expérimentaux utilisés pour la validation du système sonore, et finirons par les résultats de ces expérimentations sur des sujets jeunes.

5.1 Le son dans le HIS

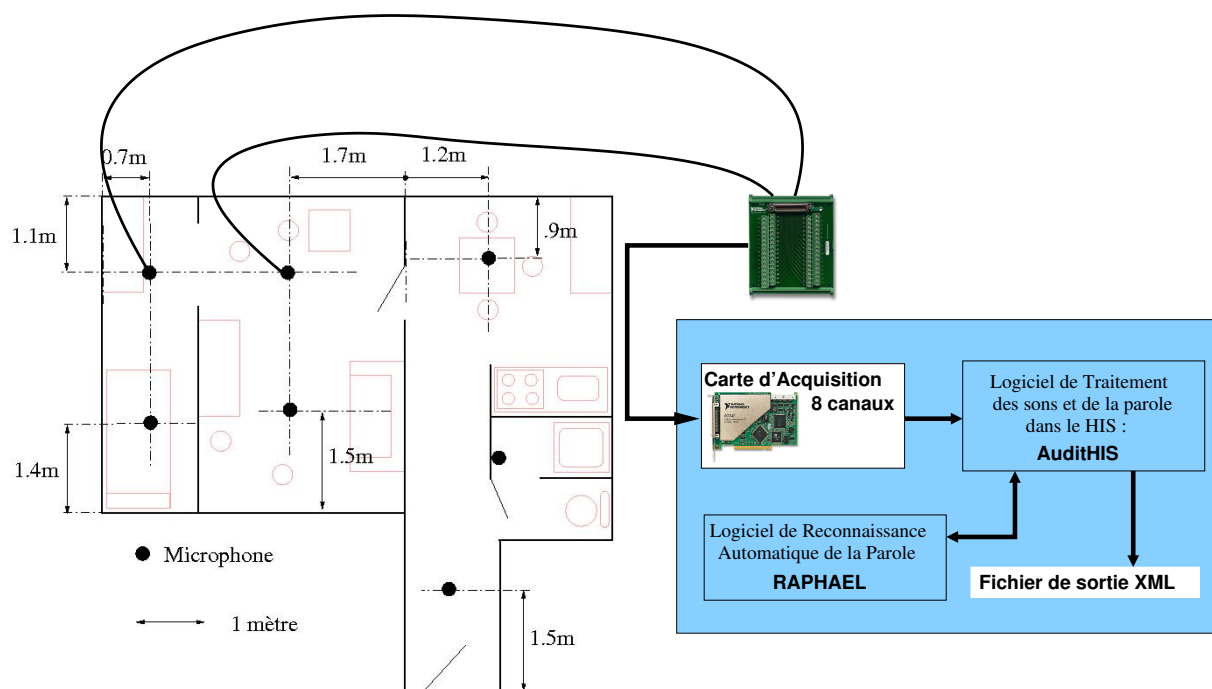


FIG. 5.1 – Organisation globale du système de reconnaissance de sons et de la parole

¹Désormais LIG, équipe GETALP

5.1.1 Organisation globale du système de reconnaissance de sons et de la parole

La figure 5.1 nous montre l'organisation globale du système. Sept microphones sont disposés dans l'appartement et reliés à une carte d'acquisition dans l'ordinateur. Une voie analogique de la carte d'acquisition est réservée pour une annulation éventuelle d'un signal parasite (télévision, radio, ...). Cet ordinateur contient l'ensemble des programmes en charge de la gestion de la carte, de la détection des sons, de la reconnaissance des types de sons et de la traduction de la parole. Les différentes parties de ce système seront décrites dans les sections suivantes.

5.1.2 Matériel utilisé

5.1.2.1 Microphones

Sept microphones ont été placés dans l'environnement de la personne. Ces microphones à électret sont omnidirectionnels (MCE-200, Panasonic). Les microphones à électret ont été choisis pour leur bon équilibre entre les caractéristiques, les performances, le prix et l'intégration. Il existe des capsules de microphones à électret de petite dimension (environ 3mm de diamètre), montables sur circuit imprimés, qui présentent un rapport signal sur bruit supérieur à 58dB et peuvent donc être utilisés dans les conditions normales d'un appartement (leur plage de température de fonctionnement étant de 0 à 40°C).

5.1.2.2 Carte d'acquisition

Les microphones sont reliés sur les entrées analogiques d'une carte au bus PCI (National Instrument NI-PCI6034E). Cette carte a une capacité d'acquisition de 200kEchantillons/s sur 16 bits, suffisante pour permettre une acquisition à 16kHz sur 8 voies. Cette carte d'acquisition est installée dans l'ordinateur contenant le programme « AuditHIS » qui est en charge de l'acquisition et du traitement des données sonores.

Un connecteur à 68 broches (TBX68, National Instrument) distribue les différents microphones (dont les câbles sont dissimulés dans les faux plafonds de l'appartement) sur les huit voies d'entrée de la carte d'acquisition analogique.

Dans une version plus récente du prototype, il est également possible d'utiliser des microphones HF (sans fil) afin de faciliter l'installation. Ce type de microphone peut être embarqué sur la personne.

5.1.2.3 Acquisition, traitement et stockage des données

Les données sont donc recueillies par le programme « AuditHIS ». Ce programme échantillonne les entrées de la carte d'acquisition pour capturer les données sonores à 16kHz. Ces données sont ensuite traitées au fil de l'eau et seuls sont stockés les échantillons correspondant à des sons de la vie courante ou de la parole.

5.1.3 Microphones : installation et réglages

5.1.3.1 Choix de l'emplacement des microphones

Les résultats de la reconnaissance sonore dans le HIS vont dépendre en grande partie de l'emplacement choisi pour les microphones. Les performances en laboratoire sont connues (voir section 5.4). En conditions réelles, nous ne pourrions ni égaler ni dépasser ces performances. Cependant, nous pouvons réduire l'écart entre performances en laboratoire et résultats dans l'appartement par ces deux variables que sont les réglages et le positionnement des microphones.

Les précédents réglages ont tous été revus suite à la nouvelle implémentation logicielle. De plus, lors des travaux précédents [Istrate, 2003], un ensemble de quatre microphones avait été disposés sur un même support en plexiglass et installé afin de capter les sons à la fois du salon et de la chambre. Ces quatre microphones devaient permettre la localisation dans l'appartement par triangulation. Les résultats de la localisation n'étant pas concluants, les travaux dans ce sens avaient été abandonnés mais les microphones n'avaient pas été déplacés. L'analyse des sons produits dans l'appartement avait montré qu'un son « continu » se superposait aux sons captés par ces microphones. Ce son était produit par la canalisation d'eau placée juste derrière l'ensemble de quatre microphones. De plus, selon le placement de la personne (sur le lit de la chambre, sur le canapé ou à côté de la fenêtre), les sons enregistrés lors des premières expérimentations par ces quatre microphones étaient quasiment toujours identiques, nous avons donc un mauvais rapport signal sur bruit lors de sons produits sur le lit alors qu'il était meilleur près de la fenêtre ; ce qui est logique du fait du placement des microphones.

Nous avons ainsi choisi, pour résoudre ces problèmes, de revenir à un quadrillage plus standard de l'appartement. Les sept microphones sont maintenant également répartis dans tout l'appartement, et chaque pièce est analysée par un ou deux microphones, selon sa taille. Le placement des microphones est montré sur la figure 5.2. Les microphones ont été intégrés à l'environnement et cachés dans le plafond de l'appartement, comme le montre la photo de droite.

Une autre modification a été apportée par rapport à la première installation. Deux microphones étaient placés dans la douche et les toilettes respectivement. Ces deux microphones avaient des performances diminuées du fait de leur placement dans une pièce présentant un écho non négligeable, et étaient extrêmement proches l'un de l'autre. Nous avons donc décidé d'éliminer l'un de ces microphones et de brancher, sur la huitième voie disponible, une référence de son qui sera utilisée pour tester en condition le module réalisé lors d'un stage au LIG qui a pour but de retrancher une source sonore (exemple une télévision ou une radio) des sons de la vie courante, en captant cette source avant de l'annuler sur chacune des autres voies. Cet algorithme sera présenté dans la suite de ce chapitre avec les autres algorithmes de traitement audio. Les premières expérimentations liées au son ont ainsi été effectuées avec les huit microphones tandis que l'expérimentation de fusion de données ainsi que l'expérimentation sur la validation du modèle d'annulation de référence ont été

réalisées avec sept microphones uniquement. Cependant, aucun autre réglage n'a été rectifié entre ces deux expérimentations, ce qui ne modifie pas les résultats que nous avons pu obtenir précédemment.

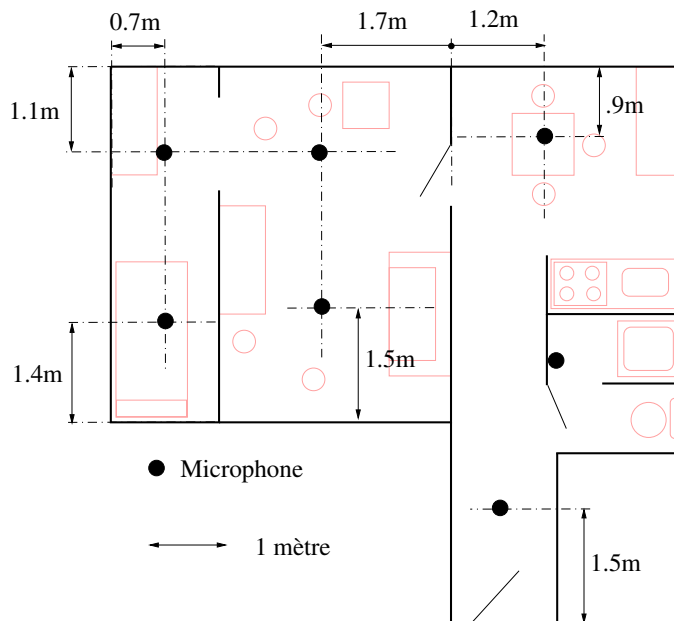


FIG. 5.2 – Répartition des microphones dans l'appartement HIS et intégration dans l'environnement

5.1.3.2 Réglage des gains

La précision des gains et du seuil de détection a été augmentée dans la version actuelle du logiciel. Cela permet un contrôle plus fin et plus précis du matériel et ainsi d'optimiser les résultats que nous pourrions obtenir en condition réelle.

Chaque microphone de l'appartement possède ses caractéristiques propres, et l'installation de chacun d'eux (proximité des sources d'émission des sons courants, intégration, ...) modifie d'autant plus les écarts que nous pouvons observer entre les microphones. Le réglage des gains de chacun des canaux ne peut donc se faire sur la base uniquement du gain du microphone donné par sa documentation technique.

Pour régler plus précisément ces gains, et déterminer les plages de capture nécessaires, nous avons utilisé la fonction de test proposée par l'interface, et généré des sons de la vie courante à l'endroit où ils seront produits dans l'appartement. Par exemple, nous avons effectué des sons de vaisselle avec une tasse et une cuillère sur la table de la cuisine. Ainsi, nous avons des exemples de sons réels et nous pouvons visualiser, avec la fonction de test, le signal obtenu et déterminer le gain nécessaire pour que le microphone capte de manière suffisante le son et sans saturation. Nous avons également réalisé des tests avec de la parole pour éviter les saturations dans ce cas qui est particulièrement critique.

Cet ensemble de tests nous a permis de définir un réglage adapté pour le HIS et de conserver celui-ci au cours des différentes expérimentations.

À noter que toute autre installation aura des caractéristiques différentes, nous devons donc effectuer ce type de tests pour chaque installation afin que les réglages soient op-

timums. Un réglage basique, défini à l'avance par rapport aux caractéristiques des microphones, ne rendrait pas le système inopérant mais diminuerait fortement ses performances.

5.1.4 Acquisition des données sonores

Les données sonores relatives à chaque expérimentation sont ensuite stockées dans un répertoire du compte utilisateur utilisé pour lancer le logiciel. Le nom de ce répertoire se compose de la date et de l'heure du début de l'expérimentation ainsi que d'un identifiant que nous pouvons modifier dans le logiciel.

Pour chaque évènement sonore, le système conserve un ensemble de données complètes :

- Le fichier WAV correspondant au signal de son ou de la parole isolé.
- Le fichier SAM correspondant aux informations de découpe du fichier.
- Un fichier texte contenant des informations intéressantes relatives à la segmentation et à la classification.
- Un fichier au format XML qui rassemble l'ensemble des résultats de traitement depuis la détection jusqu'à la classification et à la reconnaissance : identifiant unique de l'évènement, horodatage, microphone ayant effectué la détection, RSB, probabilités et résultat de la segmentation entre parole et sons, et enfin probabilités et résultat de la classification dans le cas des sons, ou bien les cinq phrases les plus probables dans le cas de la parole.

Cet ensemble de fichiers est conservé pour des réapprentissages futurs (afin d'améliorer les modèles et d'augmenter les corpus). Dans le traitement qui est fait de ces données dans le cadre de la classification des activités, nous nous concentrons sur l'ensemble de fichiers XML qui est ensuite traité et importé sous Matlab comme expliqué dans la section 5.6.

5.1.5 Objectifs

L'objectif de la collaboration entre les deux laboratoires est double. Tout d'abord, l'installation déjà effectuée permet de tester, en condition réelle et sur des sujets en bonne santé, l'architecture globale de reconnaissance de sons et de la parole, et d'obtenir des résultats plus réalistes que ce qui peut être obtenu en laboratoire, et ainsi se rapprocher le plus possible des conditions réelles auxquelles le système serait confronté s'il était intégré dans un habitat intelligent à l'heure actuelle. Ceci permet, de plus, de continuer les travaux sur cette architecture, et de connaître les besoins spécifiques à cette application qui sont autant de sources de réflexions et de projets à développer.

Le second objectif est de préciser l'apport des données sonores dans la détection d'activités. En effet, les coûts d'installation de ce système sont non négligeables surtout du fait de la carte d'acquisition. Notre intérêt est donc de déterminer par nos travaux son efficacité et son utilité afin d'en justifier ou non l'installation.

De plus, le système peut très bien être implanté dans un ordinateur « classique » et relié à une carte son du marché. On peut ainsi imaginer de mettre en place un système assez réduit,

connecté à un microphone embarqué sur la personne, comme il a été envisagé au début de cette thèse.

5.2 Algorithmes de traitement des sons

Les différentes parties du programmes utilisé dans le HIS sont décrites dans les travaux effectués par Dan Istrate et Michel Vacher [Istrate *et al.*, 2006a, Vacher *et al.*, 2006]. Le traitement d'un évènement sonore s'effectue essentiellement en trois étapes. La première est la détection de l'évènement (début et fin), la seconde est la segmentation entre son de la vie courante ou la parole et la troisième est la classification pour les sons de la vie courante et la reconnaissance de la parole pour une phrase énoncée. Ces trois étapes vont être décrites dans les sections suivantes.

5.2.1 Détection

La fréquence d'acquisition des données est de 16kHz, ce qui autorise une fréquence de coupure haute de 8kHz, suffisante pour la reconnaissance de la parole (théorème d'échantillonnage de Nyquist-Shannon). La capture se fait sur des fenêtres de 64 ms (1024 échantillons). La détection des sons utilise deux fenêtres, soit 128 ms (2048 échantillons).

Pour réaliser cette détection, la première étape est de calculer la transformée en ondelettes discrète du signal. Cette transformée est calculée avec comme ondelette mère ψ l'ondelette de Daubechies possédant six moments nuls (DB6) [Daubechies, 1988]. Pour rappel, les ondelettes de Daubechies sont à support compact, DBn possède n moments nuls et 2n coefficients et la famille d'ondelettes est une famille orthogonale de $L^2(\mathbb{R})$. La transformée en ondelettes discrète est alors donnée par l'ensemble des coefficients :

$$f_{m,n} = \langle f | \psi^{m,n} \rangle \quad (5.1)$$

Avec $\psi^{m,n}$ définie par :

$$\left\{ \psi^{m,n} : x \mapsto \frac{1}{\sqrt{2^m}} \psi \left(\frac{x - 2^m n}{2^m} \right) \right\}_{(m,n) \in \mathbb{Z}^2} \quad (5.2)$$

Pour réaliser la détection, une analyse multi-résolution du signal est faite à l'aide des ondelettes de Daubechies. L'arbre utilisé est de profondeur 3 (voir figure 5.3). La détection se fait alors par seuillage sur l'énergie de l'arbre d'ondelettes, calculé par la somme des énergies de chacun des noeuds (somme pondérée des carrés des coefficients de la transformée en ondelettes contenue dans ce noeud). Le seuil utilisé est adaptatif, sa valeur est maintenue à jour en considérant les dix dernières fenêtres analysées et n'ayant déclenchées aucune détection. Cela permet de prendre en compte instantanément le bruit de fond présent dans un appartement. L'algorithme est décrit par la figure 5.4.

Pour réaliser la détection, le logiciel calcule donc une énergie de l'arbre correspondant à l'analyse multi-résolution du signal. Lorsqu'il n'y a pas de détection, la valeur de cette éner-

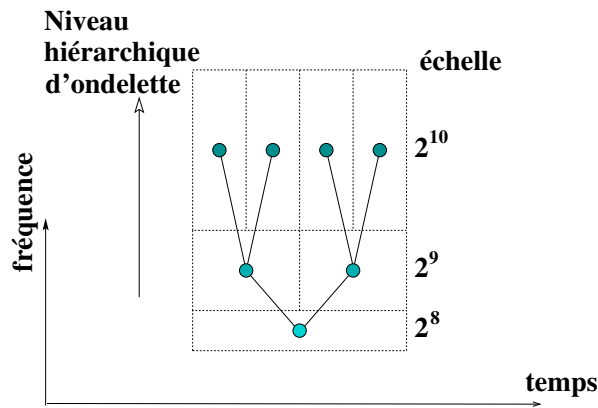


FIG. 5.3 – Construction de l'arbre pour l'analyse multi-résolution du signal avec les ondelettes de Daubechies [Vacher *et al.*, 2004]

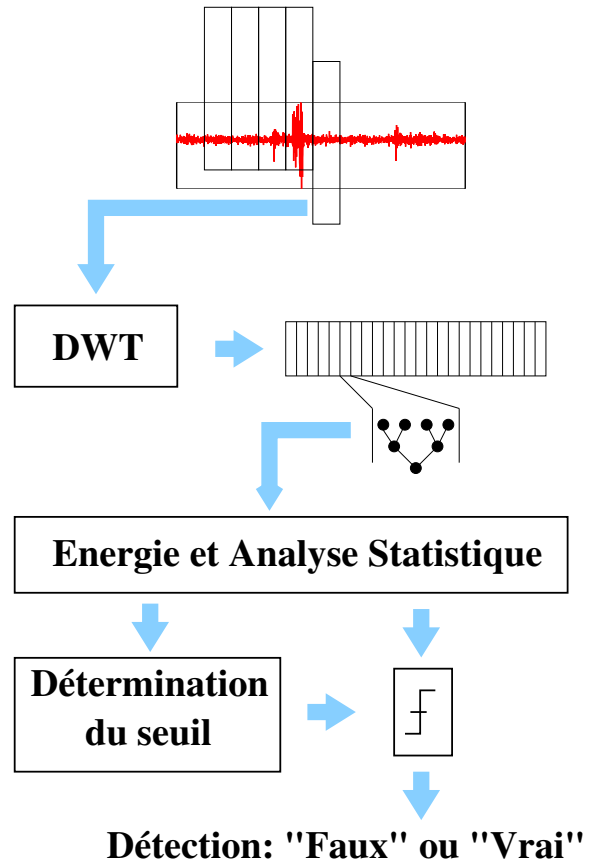


FIG. 5.4 – Algorithme de détection des sons [Vacher *et al.*, 2004]

gie permet de maintenir à jour un seuil adaptatif, qui conserve la mémoire des dix dernières fenêtres n'ayant pas provoquées de détection. La fonction de mise à jour de ce seuil est du type $th = \kappa + \sigma \mu$ avec μ la moyenne des dix dernières valeurs des coefficients et κ et σ deux coefficients fixés préalablement par analyse de performances expérimentales.

5.2.2 Estimation du Rapport Signal sur Bruit (RSB)

Le rapport signal sur bruit, ajouté au fichier de sortie, est un excellent indicateur de la « qualité du son » et donc de la pertinence de la décision qui a été prise. Il est ainsi pris en compte dans les algorithmes de tri des données décrits à la fin de ce chapitre.

Afin de calculer le rapport signal sur bruit, il faut tout d'abord estimer le bruit de fond présent dans la pièce. Ce bruit de fond est calculé par la puissance des dix dernières fenêtres sans détection (soit une durée de 640ms). La puissance de l'une des fenêtres est calculée par la formule :

$$P_b = \sum_{i=1}^{2048} ech_i^2 \quad (5.3)$$

Puis, dans le cas d'une fenêtre sans détection, la moyenne des dix dernières puissances est calculée :

$$P_{\text{BruitFond}} = \frac{1}{10} \cdot \sum_{i=1}^{10} P_{b_i} \quad (5.4)$$

Lors de la détection d'un évènement, la puissance sur la fenêtre utile du fichier WAV, ramenée à une fenêtre de 2048 échantillons, est donnée par :

$$P_e = \frac{2048}{N} \sum_{i=1}^N \text{ech}_i^2 \quad (5.5)$$

Avec N le nombre d'échantillons conservés dans la fenêtre utile. Le RSB est ensuite donné par la formule :

$$RSB_{dB} = 10 \cdot \log_{10} \left(\frac{P_e}{P_{\text{BruitFond}}} \right) \quad (5.6)$$

5.2.3 Segmentation entre signal de parole et son de la vie courante

5.2.3.1 Modèles acoustiques

Afin d'être classifié, le signal doit être modélisé. La modélisation retenue utilise les coefficients LFCC (Linear-scale Frequency Cepstrum Coefficients). Les coefficients sont calculés sur des fenêtres de 16 ms avec un recouvrement de 50%.

Le calcul se fait en quatre étapes :

- Transformée de Fourier discrète du signal s : $S(n) = \sum_{i=0}^{N-1} s(i) e^{-j2\pi i n / N}$, $n \in [0..N-1]$
- Filtrage de la transformée de Fourier avec un banc de filtres triangulaires dont les fréquences sont linéairement réparties. Le recouvrement entre chaque filtre est de 50% et le banc de filtres contient 16 fréquences (figure 5.5).
- Transformation de cosinus inverse du logarithme de l'amplitude du résultat du précédent filtrage.

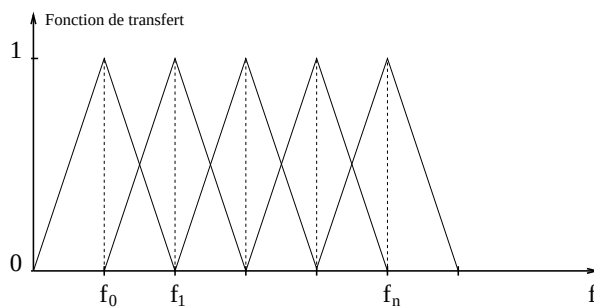


FIG. 5.5 – Banc de filtres pour le calcul des coefficients LFCC [Glasson, 2008]

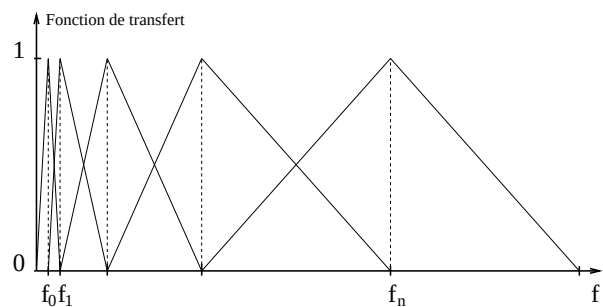


FIG. 5.6 – Banc de filtres pour le calcul des coefficients MFCC (échelle Mel) [Glasson, 2008]

Cette dernière étape nous donne les coefficients dits LFCC. Ceux-ci se basent sur une échelle linéaire pour le banc de filtres triangulaires. Une échelle logarithmique peut également être utilisée, appelée échelle Mel (figure 5.6), et donnant les coefficients dits MFCC (Mel Frequency Cepstral Coefficients). Cette seconde méthode est également implémentée dans le logiciel, mais cependant nous utilisons, lors des expérimentations, les coefficients LFCC, qui ont montré de meilleurs résultats dans la discrimination des sons de la vie quotidienne par rapport à la parole.

5.2.3.2 Segmentation parole/sons de la vie courante par classification

Les précédents coefficients servent à l'extraction de paramètres avant classification. Celle-ci se fait ensuite en utilisant les GMM (Gaussian Mixture Model – Modèle de Mélange de Gaussiennes) pour une classification supervisée (les différents échantillons étant étiquetés et le nombre de classes prédéfinies) entre deux classes : sons de la vie courante et paroles.

L'algorithme utilise un nombre fixe de Gaussiennes (24) pour modéliser les différentes classes selon les paramètres acoustiques précédemment définis. Le nombre optimal de Gaussiennes a été déterminé en traçant l'évolution du « Bayesian Information Criterion Coefficient » [Schwarz, 1978] en fonction du nombre de Gaussiennes.

Cet algorithme se réalise en deux étapes : l'apprentissage et la classification. Le but de l'apprentissage est de déterminer les caractéristiques et poids respectifs de chaque Gaussienne pour un ensemble donné et qui modélisent au mieux cet ensemble. L'initialisation consiste à fixer les paramètres des Gaussiennes de manière arbitraire en divisant l'ensemble d'apprentissage en sous-ensembles : il y a ensuite une phase d'optimisation des valeurs de chaque paramètre par un algorithme EM (Expectation Maximisation) avec vingt itérations sur une partie du corpus tirée au sort puis vingt itérations sur l'ensemble du corpus.

Une fois cet apprentissage effectué, l'étape de classification intervient. Une nouvelle donnée sera attribuée à la classe la plus probable après calcul de la vraisemblance pour son vecteur acoustique dans chacune des classes. La décision est ensuite prise avec le maximum de vraisemblance.

5.2.4 Classification des sons de la vie courante

La classification des sons de la vie courante se fait par une méthode similaire, le nombre de classes est par contre différent. Les coefficients LFCC paramétrisent les sons et la classification se fait suivant le calcul du maximum de vraisemblance, chaque classe étant modélisée par un GMM. Le nombre de gaussiennes est fixé à 12 et le nombre de coefficients LFCC est de 24.

Le modèle contient actuellement 8 classes de sons de deux types :

- Des sons « normaux » : claquement de porte, sons de pas, serrures de porte, son de vaisselle, sonnerie de téléphone.
- Des sons « de détresse » : chute d'objet, cri, bris de verre.

Des classes supplémentaires devraient être ajoutées suite à la capture de certains sons (comme les sons d'eau).

Une seconde méthode, se basant cette fois-ci sur les HMM (Hidden Markov Models – Chaînes de Markov Cachées) est aussi implémentée dans AuditHIS mais n'a pas été utilisée dans cette expérimentation.

5.2.5 Reconnaissance automatique de la parole

La reconnaissance automatique de la parole est effectuée par l'outil RAPHAEL basé sur le projet Janus² [Akbar et Caelen, 1998]. Cette reconnaissance fonctionne en quatre étapes :

- Calcul des paramètres acoustiques sur des fenêtres de 16 ms (avec 50 % de recouvrement). Les paramètres utilisés sont 16 MFCC, l'énergie, ainsi que les dérivées premières Δ et secondes $\Delta\Delta$ des coefficients MFCC. Ces dérivées sont utiles pour tenir compte des variations dans le temps des MFCC lors du calcul de la vraisemblance.
- Extraction des phonèmes à partir du modèle acoustique (contenant les modèles de l'ensemble des phonèmes). Les modèles utilisés pour chaque phonème sont des Chaînes de Markov Cachées.
- Reconstruction des mots, à l'aide du dictionnaire phonétique donnant la transcription phonétique de chacun des mots du langage.
- Reconnaissance de la phrase prononcée à l'aide des 3-grams représentant le modèle de langage (c'est à dire la probabilité de rencontrer un mot après un autre).

Le logiciel de reconnaissance produit, à partir d'un signal sonore, les cinq phrases les plus probables suivant le modèle acoustique, le dictionnaire et le modèle de langage utilisé.

5.3 Bases d'apprentissage

Si l'on met à part le cas du système de reconnaissance automatique de la parole, qui sera traité séparément, l'apprentissage se fait sur deux bases spécifiques de sons et de paroles. Les bases d'apprentissage initiales ont été enregistrées en utilisant un microphone supercardioïde HF sans fil (eW500, Sennheiser). La première base contient des sons de la vie courante et la seconde des paroles, utilisées pour tester la reconnaissance et la classification en parole de type normale/détresse. L'apprentissage pour la segmentation se fait à partir de l'ensemble des sons de ces deux bases, chacune des bases constituant une classe d'apprentissage. La base de sons de la vie courante se compose de huit classes de sons qui sont utilisées de manière individuelle pour leur tâche respective d'apprentissage.

5.3.1 Corpus des sons de la vie courante

Le corpus des sons de la vie courante permet de décrire les huit classes de sons reconnues par le logiciel. Ces classes sont de deux types :

- Celles des sons dits normaux, au nombre de cinq : claquement de porte, sonnerie de téléphone, son de pas, son de vaisselle et son de serrure de porte.
- Celles des sons relatifs aux situations de détresse : les cris, les chutes d'objets et les bris de verre.

L'enregistrement de ce corpus a été réalisé au laboratoire CLIPS pour la première partie, quelques fichiers extraits du web, puis des sons enregistrés au LIG (61% du corpus). Le corpus est constitué de 1985 fichiers, chaque fichier contenant un son isolé et correctement

²<http://www.is.cs.cmu.edu/mie/janus.html>

découpé. La durée totale de ces enregistrements est de 35 minutes et 38 secondes. La répartition des sons dans le corpus est donnée par le tableau 5.1.

Classe	Corpus CLIPS	Corpus LIG	Internet	Nombre de Fichiers	Durée moyenne d'un son (ms)
Vaisselle	45%	50%	5%	363	606
Serrure de porte		100%		507	390
Claquement de porte	40%	60%		372	1022
Bri de verre	40%	50%	10%	118	269
Chute d'objet		100%		128	1039
Sonnerie de téléphone	15%	70%	15%	319	991
Cris	80%		20%	102	432
Sons de pas	10%	60%	30%	76	86
Global	29%	61%	10%	1985	276

TAB. 5.1 – Répartition des éléments de chacune des classes de la base d'apprentissage des sons

À noter que les expérimentations que nous décrivons dans la suite de ce chapitre seront également utilisées *a posteriori* pour améliorer ce corpus de sons. En effet, une fois triés, les sons peuvent être recoupés correctement et réinsérés dans le module d'apprentissage afin d'avoir des sons avec un RSB plus faible dans la base d'apprentissage. Elles permettent aussi d'augmenter le nombre de classes et la variabilité des sons reconnus. Par exemple, lors de la première expérimentation, nous avons été confrontés au problème que la base d'apprentissage, pour les sons de téléphone, ne contient que des sonneries « classiques ». Les sonneries multi-tons des téléphones DECT actuels sont donc moins bien reconnues (pour moitié comme cri et pour moitié comme sonnerie de téléphone) comme nous le verrons. Le même problème s'est posé avec la serrure de l'appartement, ce type de serrure n'étant pas dans la base de données.

Le HIS nous a également permis d'enregistrer d'autres sons. Nous avons profité de l'installation pour enregistrer divers bruits d'eau (robinets, chasse d'eau, douche). Ces sons collectés permettront, à terme, de constituer des classes supplémentaires.

5.3.2 Paroles

5.3.2.1 Construction des modèles de langage et des modèles acoustiques

Avant de reconnaître les différents mots et les différentes phrases, il est nécessaire de déterminer les séquences les plus vraisemblables de phonèmes dans la phrase analysée. Chaque phonème est modélisé par un HMM à trois états qui nécessite un apprentissage à partir de corpus de la langue française qui ont été étiquetés après enregistrement, l'ensemble de ces modèles constitue le modèle acoustique qui sera utilisé par le système de RAP (reconnaissance automatique de la parole). À cet effet, et pour permettre une bonne indépendance de la reconnaissance vis-à-vis du locuteur, nous avons utilisé un ensemble de corpus qui ont

été enregistrés par trois cents personnes différentes dans les laboratoires CLIPS (cent locuteurs, BRAF100) [Vaufreydaz *et al.*, 2000] et LIMSI³ (80 locuteurs, BREF80 puis 120 locuteurs, BREF120) [Gauvain *et al.*, 1990].

Pour reconnaître les phrases les plus probables, le logiciel de reconnaissance de la parole a besoin de connaître un « modèle de langage », c'est à dire non seulement les mots et leur constitution à partir des phonèmes grâce à un dictionnaire phonétique, mais aussi la probabilité d'observer une séquence de mots au moyen de 3-grammes. Le modèle de langage actuellement utilisé par RAPHAEL est constitué de 9958 mots français et est obtenu à partir de l'extraction d'informations textuelles du journal « Le Monde » et d'Internet. Il a ensuite été optimisé à partir de l'extraction d'informations dans une base de conversations en français, contenant 253 phrases énoncées lors de conversations téléphoniques (« Allô oui », « A demain », « j'ai bu ma tisane », « au revoir » ...), 66 phrases considérées comme anodines (situations normales de la vie courante) « Bonjour », « Où est le sel » ... et 60 typiques d'une situation de détresse (« Aouh », « Aïe », « Au secours », « Un médecin vite » ...). Des versions syntaxiquement incorrectes – mais couramment utilisées en français – de ces phrases ont également été introduites dans les corpus, telle que « Ça va pas bien ».

Le but recherché lors de la génération de ces différents modèles est la détection de situation de détresse par l'apparition de mots clés caractéristiques d'une situation anormale ou de détresse, et non une détection parfaite de la conversation de la personne dans l'appareil (ainsi que le ferait un système de reconnaissance de la parole permettant de dicter un texte à l'ordinateur).

5.3.2.2 Base de tests

La base de test utilisée pour l'évaluation de la détection des situations de détresse à partir de mots clés dans la parole reprend l'ensemble du corpus anodin/détresse utilisé pour la segmentation entre sons de la vie courante et parole. Cette base contient ainsi 2646 fichiers, chacun représentant une phrase, prononcée par un des 21 locuteurs (âgés de 20 à 65 ans).

5.4 Résultats en laboratoire

5.4.1 Détection des sons par calcul d'un arbre d'ondelettes

Le tableau 5.2 présente les résultats de la détection des sons avec l'algorithme utilisé (basé sur le calcul de l'énergie d'un arbre d'ondelettes), dans le cas d'un bruit capté dans le HIS et dans le cas d'un bruit blanc, superposé au sons de la base avec différents rapports signal sur bruit (0dB, 10dB, 20dB et 40dB).

Comme nous le montre ce tableau, la détection se fait sans erreur à partir de 10dB pour un bruit de fond capté dans le HIS et à partir de 20dB pour un bruit blanc. Les taux d'erreur sont toujours inférieurs à 10%.

³Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur

RSB	Taux d'erreur	
	Bruit HIS	Bruit Blanc
0 dB	7,3 %	6,1 %
10 dB	0 %	4%
20 dB	0 %	0 %
40 dB	0 %	0 %

TAB. 5.2 – Taux d'erreur de détection pour un bruit blanc et un bruit de fond capté dans le HIS, à différents rapports signal sur bruit

5.4.2 Segmentation parole/sons de la vie courante

Pour la validation des algorithmes de segmentation, les corpus de sons et de paroles ont été mélangés à un bruit enregistré dans le HIS. Ce bruit a été mélangé au son initial selon quatre rapports signal sur bruit différents (0dB, 10dB, 20dB et 40dB). Les résultats sont donnés, pour la classification utilisée en fonction de ce rapport signal sur bruit, dans le tableau 5.3. Ces résultats montrent l'importance significative de ce rapport signal sur bruit sur les résultats expérimentaux. Les résultats se stabilisent cependant à partir de +20dB. En effet, au dessus de cette limite, le taux d'erreur stagne à environ 3.8% alors qu'il atteint tout de même 17% pour un rapport signal sur bruit de 0dB. Ces résultats ont été obtenus à partir d'une validation croisée sur tout le corpus, dans laquelle on utilise séquentiellement 80% du corpus pour l'apprentissage et 20% pour le test.

Rapport Signal sur Bruit (dB)	0	+10	+20	+40
GMM (24), 16 LFCC	17.3%	5.1%	3.8%	3.6%

TAB. 5.3 – Taux d'erreur de segmentation Son/Parole

5.4.3 Classification des sons

Les performances de classification sont données par le tableau 5.4 (validation croisée). De nouveau, ces résultats sont présentés en fonction du rapport signal sur bruit, pour la méthode utilisée par le logiciel lors de nos expérimentations. Une nouvelle fois, le taux de bonne classification est fortement dépendant du rapport signal sur bruit.

Signal to Noise Ratio (dB)	0	+10	+20	+40
GMM, 24 LFCC	36.6%	21.3%	13%	9.3%

TAB. 5.4 – Erreur de classification entre les 8 classes de sons

5.4.4 Reconnaissance des mots clés de détresse

La dernière partie à tester est la détection des mots clés de détresse à l'intérieur d'une conversation. Cette partie a été testée sur les données enregistrées par 21 locuteurs du corpus du anodin/détresse (soit un total de 2646 tests). Les résultats sont donnés dans le tableau 5.5. Les deux taux qui sont donnés représentent les deux types d'erreurs possibles : (1) dans 0,5% des cas, pour une phrase dite anodine, un mot clé de détresse a été inclus dans la traduction, ce qui mène à une fausse alarme (faux positif) ; (2) dans 22% des cas, pour une phrase de détresse, le mot de détresse a été manqué ce qui mène à une alarme manquée (faux négatif). Ces alarmes manquées se produisent souvent sur des mots isolés (« Aïe », « SOS » . . .) ou sur des phrases syntaxiquement incorrectes en français, qui sont incluses dans le modèle de langage mais encore trop peu représentées. Le taux d'erreur de reconnaissance globale des mots clés dans ces conditions (non bruité) est donc de 11%.

Corpus	Erreur de reconnaissance	
Phrases « normales »	Taux de fausses alarmes	0,5%
Phrases de « détresse »	Taux d'alarmes manquées	22%
Erreur globale de reconnaissance		11%

TAB. 5.5 – Taux d'erreur de reconnaissance de la parole

Ce taux pourrait être amélioré de deux manières, en cours d'investigation au LIG : d'une part en améliorant le modèle de langage pour qu'il soit plus adapté au français spontané et parlé et non au français écrit (corpus du monde) et d'autre part en ajoutant des variantes phonétiques supplémentaires pour les différents mots du corpus.

5.5 Implémentation : le logiciel AuditHIS

Dan Istrate avait précédemment réalisé une application de détection et de classification des sons de la vie courante [Istrate, 2003]. Cette application tournait sous Microsoft Windows™ et Labwindows/CVI™ et envoyait ses données sur le bus CAN de la même manière que les autres capteurs intégrés au HIS. Pelayo Menenedez Garcia, lors d'un stage ingénieur, avait repris les travaux pour y ajouter la reconnaissance de la parole [Vacher *et al.*, 2006]. Le développement a été repris dans son intégralité sous GNU/Linux par Hubert Glasson [Glasson, 2008] lors de la réalisation de son mémoire CNAM dont les travaux ont abouti au logiciel AuditHIS, système « multithread » (processus légers) d'analyse en temps-réel des sons et de la parole.

5.5.1 Description générale

L'application se compose de plusieurs processus en charge des différentes tâches demandées au logiciel (figure 5.7) :

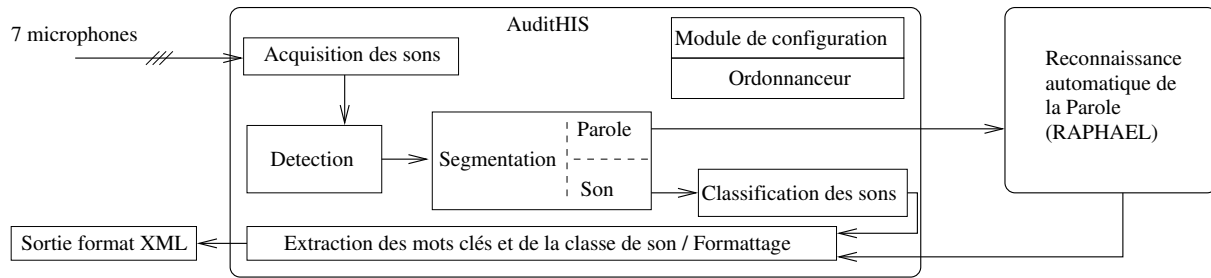


FIG. 5.7 – Organisation générale de l'application AuditHIS

- Gestion de la carte son, ainsi que de la détection des sons et du calcul du RSB.
- Suppression de sources sonores parasites.
- Détection des sons simultanés pour l'indiquer dans le fichier de sortie.
- Segmentation des sons en parole/son de la vie courante.
- Classification des sons de la vie courante dans l'une des huit classes.
- Gestion des échanges avec le logiciel RAPHAEL pour la récupération des phrases.
- Écriture des fichiers de sortie au format XML, rassemblant l'ensemble des informations sur le signal en cours d'analyse.

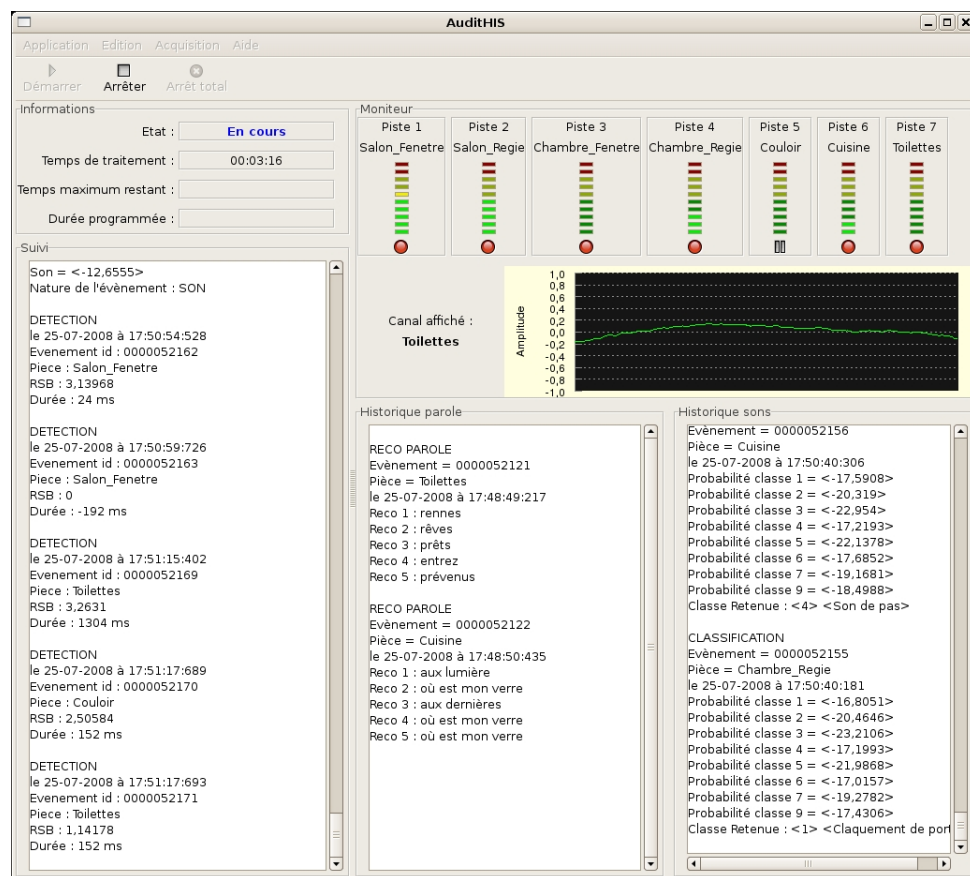


FIG. 5.8 – Capture d'écran de la fenêtre principale de l'application AuditHIS

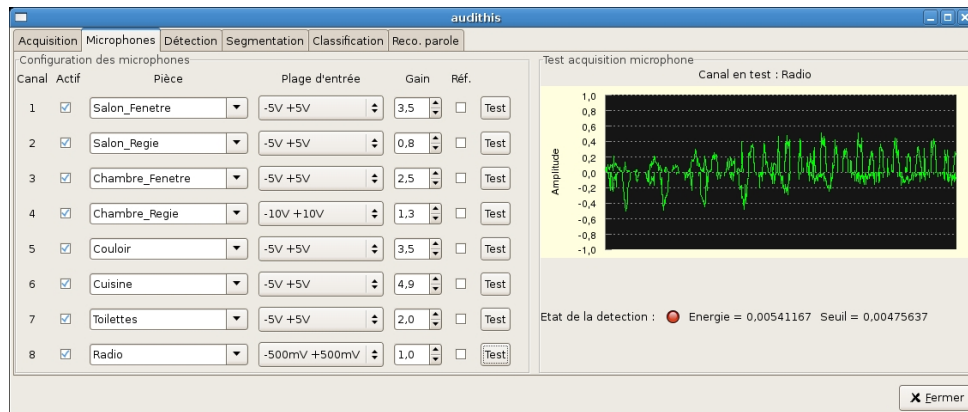


FIG. 5.9 – Capture d'écran de la fenêtre de configuration du logiciel AuditHIS

5.5.2 Données extraites

```
<?xml version="1.0" encoding="UTF-8"?>
<evenement id="0000009404">
  <fichier>0000009404_detect.wav</fichier>
  <rsb>10,1699</rsb>
  <detection>
    <piece>Salon_Haut_Gauche</piece>
    <dhm>2008-03-05 15:27:20:457</dhm>
    <debut>5120</debut>
    <fin>14464</fin>
    <duree>584</duree>
    <simultane>
      <evenement numero="1">0000009405</evenement>
      <evenement numero="2">0000009406</evenement>
      <evenement numero="3">0000009407</evenement>
      <evenement numero="4">0000009408</evenement>
    </simultane>
  </detection>
  <segmentation>
    <probabilite>
      <parole>-18,4389</parole>
      <son>-20,0514</son>
    </probabilite>
    <resultat>parole</resultat>
  </segmentation>
  <recoparole>
    <texte>non merci</texte>
    <texte>non merci</texte>
    <texte>non merci</texte>
    <texte>oui merci</texte>
    <texte>poésie</texte>
  </recoparole>
</evenement>
```

```
<?xml version="1.0" encoding="UTF-8"?>
<evenement id="0000009262">
  <fichier>0000009262_detect.wav</fichier>
  <rsb>14,445</rsb>
  <detection>
    <piece>Salon_Haut_Gauche</piece>
    <dhm>2008-03-05 15:25:23:767</dhm>
    <debut>5888</debut>
    <fin>45696</fin>
    <duree>2488</duree>
    <simultane>
      <evenement numero="1">0000009263</evenement>
      <evenement numero="2">0000009264</evenement>
      <evenement numero="3">0000009265</evenement>
    </simultane>
  </detection>
  <segmentation>
    <probabilite>
      <parole>-30,0085</parole>
      <son>-23,6304</son>
    </probabilite>
    <resultat>son</resultat>
  </segmentation>
  <classification>
    <probabilite>
      <classe numero="1">-62,5576</classe>
      <classe numero="2">-41,2478</classe>
      <classe numero="3">-36,0293</classe>
      <classe numero="4">-82,9985</classe>
      <classe numero="5">-39,305</classe>
      <classe numero="6">-74,7359</classe>
      <classe numero="7">-43,5397</classe>
      <classe numero="9">-42,4548</classe>
    </probabilite>
    <classeretenue>3</classeretenue>
    <description>Sonnerie téléphone</description>
  </classification>
</evenement>
```

Le format de sortie des fichiers de AuditHIS est le format XML. Ces fichiers contiennent l'ensemble des informations sur les différentes étapes du traitement d'un son :

- Sur la détection d'abord, avec l'horodatage, le numéro du microphone qui a servi à la détection, un identifiant unique, le rapport signal sur bruit, la durée et enfin l'ensemble des sons simultanés à celui-ci.
- Sur la segmentation, avec la décision prise et les probabilité d'appartenir à chacune des deux classes (parole et son de la vie courante).
- Sur la classification des sons avec la classe retenue ainsi que l'ensemble des probabilités pour chacune des classes.

- Sur la reconnaissance de la parole avec les cinq phrases les plus probables fournies par RAPHAEL.

L'ensemble de ces informations est indiqué dans un fichier indépendant pour chacun des sons, numéroté avec l'identifiant correspondant à la détection. Une copie du fichier WAV ainsi que du fichier SAM (indiquant la découpe du son) sont également conservées. Deux exemples de fichiers XML complets sont montrés ci-dessus, avec à gauche une détection de la parole et à droite une détection d'un son de la vie courante.

5.6 Exploitation des données

5.6.1 Extraction des données

Les données sont ensuite dépouillées sous MatlabTM. Chaque expérimentation contient un ensemble de dossiers (un dossier par microphone), chaque dossier contenant les différents fichiers produits par le logiciel (WAV, SAM et XML) pour chacune des détections. Nous nous intéressons principalement aux fichiers XML, qui contiennent l'ensemble des informations nécessaires à la reconstitution du « parcours sonore » de la personne dans l'appareil durant l'expérimentation.

L'ensemble de ces fichiers XML est chargé par le programme Matlab, puis traité pour produire un fichier texte unique, contenant l'ensemble des informations nécessaires pour les utilisations ultérieures dans un contexte de fusion de données.

5.6.2 Règles de classification

5.6.2.1 Parole

La classification de la parole se base sur la détection de mots clés de détresse. Nous avons mis en place plusieurs règles que nous avons testé durant les expérimentations. Trois classifications différentes ont été testées :

- La classification C1, associée à la segmentation S1, se base sur le microphone ayant le plus fort RSB et prend la décision en fonction de la première phrase de celui-ci. Si cette phrase contient un mot de détresse alors elle est classée comme telle.
- La classification C2, associée à la segmentation S1, se base également uniquement sur le premier microphone mais donne du poids aux trois premières phrases de celui-ci. Elle donne comme poids 1 à la première phrase, 0,75 à la seconde et 0,5 à la dernière et fait la somme des décisions entre anodin et détresse.
- La classification C3, associée à la segmentation S2, considère l'ensemble des microphones ayant un RSB suffisant. Sur un ensemble de sons simultanés, le RSB du microphone ayant le maximum est noté x . Est considéré comme RSB suffisant pour une décision tout microphone ayant un RSB minimum de $0,8 \cdot x$ où x est le RSB maximum de l'ensemble des microphones pour un son donné.

Pour la classification C3, la décision est prise ensuite par « vote majoritaire » entre les microphones. En cas d'égalité lors de ce vote, la décision Détresse est prise plutôt que la décision Anodine, par principe de précaution.

5.6.2.2 Sons de la vie courante

Les sons de la vie courante sont classés d'une manière équivalente à C3, précédemment présentée. Le système retient le RSB maximum x pour un ensemble de sons simultanés, puis sélectionne tous les microphones dont le RSB est supérieur à $0,8 \cdot x$. Ensuite, un vote majoritaire donne la classification finale. En cas d'égalité entre plusieurs décisions, celle donnée par le microphone ayant le plus fort rapport signal sur bruit sera retenue.

5.6.3 Fichier de sortie

Un fichier de sortie résume l'analyse des sons faite précédemment et concentre l'ensemble des informations des détections pour une expérimentation. Il contient, pour chacune des détections, le RSB maximum du son détecté ainsi que le microphone ayant effectué cette détection, l'horodatage, l'identifiant du fichier sonore ayant le RSB maximum, ainsi que la classe de son retenue après classification suivant les méthodes précédentes.

Ce fichier se présente comme suit :

12.9419	Chambre_Fenetre	2008-07-11/11:02:18	45035	Cris
23.285	Cuisine	2008-07-11/11:02:24	45044	Cris
12.6175	Toilettes	2008-07-11/11:02:29	45049	
Claquement de porte				
13.8855	Chambre_Regie	2008-07-11/11:02:38	45054	ParoleA
33.3434	Couloir	2008-07-11/11:02:42	45059	Cris
11.3888	Chambre_Fenetre	2008-07-11/11:02:47	45063	
Claquement de porte				
25.0563	Salon_Fenetre	2008-07-11/11:02:53	45066	Son de pas
23.6623	Chambre_Fenetre	2008-07-11/11:02:59	45072	Cris
20.9433	Chambre_Regie	2008-07-11/11:03:05	45078	Chute d'objets
23.2928	Chambre_Regie	2008-07-11/11:03:12	45083	ParoleA
22.5176	Couloir	2008-07-11/11:03:17	45089	Cris
25.1226	Couloir	2008-07-11/11:03:22	45093	
Claquement de porte				
15.7925	Chambre_Fenetre	2008-07-11/11:03:28	45099	ParoleD

5.7 Expérimentations préliminaires

Plusieurs expérimentations ont été menées, nous montrant chacune des problèmes à résoudre au sein du logiciel ou de la chaîne de traitement. Ces différentes expérimentations sont décrites dans les sections suivantes. Elles se sont déroulées dans le HIS du laboratoire, chacun des passages étant filmé afin de pouvoir faire une indexation a posteriori. Deux des

caméras vidéo du HIS acquièrent en même temps les données sonores grâce à leur microphone directionnel. L'une d'elle surveille le salon et la chambre et la seconde surveille le couloir et la cuisine. L'ensemble de ces expérimentations et de ces résultats est révélateur de la démarche suivie pour améliorer le système, raison pour laquelle ils sont tous présentés dans cette section. À noter également que les deux premières expérimentations se sont déroulées avec l'ancienne configuration au niveau de la position des microphones (donc quatre microphones placés ensemble sur une « antenne » dans le salon). Dans tous les cas lors de ces expériences, nous n'avions aucun contrôle sur la position de la personne, sur la manière de réaliser le scénario ou de réaliser les actions individuelles. Nous n'avions également aucun contrôle sur l'environnement extérieur. Par exemple, si un bruit important était effectué dans le couloir attenant (claquement d'une porte, bruit de perceuse, ...), l'expérimentation n'était pas interrompue et les résultats étaient tout de même pris en compte.

5.7.1 Expérimentation 1

5.7.1.1 Protocole

Cette première expérimentation visait à valider les deux systèmes de reconnaissance de la parole et de sons. Pour ceci, le protocole se divisait en deux parties distinctes, une partie dans laquelle le sujet effectuait des sons de la vie courante et une seconde dans laquelle il énonçait des phrases anodines et des phrases de détresses. 15 sujets (8 femmes, 7 hommes) se sont portés volontaires pour effectuer cette expérimentation. L'âge moyen de la population est de $27.6 \text{ ans} \pm 6$, pour un poids de $60.1 \text{ kg} \pm 13$ et une taille de $1.71 \text{ m} \pm 0.08$. Le temps d'exécution moyen du scénario a été de $7 \text{ min } 11 \text{ secs} \pm 1 \text{ min } 13 \text{ secs}$.

5.7.1.1.1 Reconnaissance sonore

Il était demandé au sujet d'entrer dans l'appartement, à l'aide de la clé, de refermer la porte derrière lui et d'effectuer la suite d'actions ci-dessous, sans contrainte d'ordre ni de manière de les exécuter :

- Marcher de la porte d'entrée de l'appartement jusqu'au lit, puis revenir jusqu'à la porte. Cette action était à répéter 3 fois dans l'ensemble du scénario.
- Fermer la porte de la salle de bains (3 fois).
- Crier « Aïe » (3 fois).
- Répondre au téléphone lorsqu'il sonne, et dire « Allô » (3 fois).
- Laisser tomber un livre de la table de la cuisine (3 fois).
- Remuer une cuillère dans la tasse à café, placée sur la table de la cuisine, durant quelques secondes (3 fois).
- Tirer la chasse d'eau.
- Faire couler l'eau au robinet de la salle de bain et de la cuisine.

5.7.1.1.2 Reconnaissance de la parole

En ce qui concerne la reconnaissance de la parole, les sujets disposaient de quatre feuilles avec les corpus de phrases anodines et détresse. Il leur était demandé d'énoncer trois phrases tirées au hasard de chacune des feuilles. Ils devaient donc énoncer six phrases de détresses ainsi que six phrases anodines. L'ensemble des corpus de phrases est donné en Annexe E.

5.7.1.2 Résultats expérimentaux

Pour l'ensemble des 15 sujets, 3580 sons ont été détectés par le système, dont 350 éliminés du fait d'un rapport signal sur bruit inférieur à 0 dB. Le rapport signal sur bruit de l'ensemble des sons retenus s'élève à $6.2 \text{ dB} \pm 6.6$.

5.7.1.2.1 La parole

Dans le traitement de cette expérimentation, seul le son le plus probable pour une parole donnée a été classifié, en considérant la première phrase sortie pour cette parole (la phrase la plus probable). Ainsi, si un mot clé de détresse est détecté dans la première phrase du microphone ayant le meilleur rapport signal sur bruit, la phrase sera retenue comme détresse. Dans le cas contraire, elle sera classée comme anodine. Le tableau 5.6 présente l'ensemble des résultats pour la reconnaissance de la parole. Il présente en effet les taux d'erreurs en détection, segmentation et classification. L'un des enseignements apporté par cette expérimentation a été sur le taux de détection. En effet, comme le tableau le montre, le taux de non détection pour la parole s'élève à 33%, ce qui influe énormément sur le taux d'erreur global. En effet, un tiers des phrases n'est jamais entré dans l'étape de segmentation et de classification. Certaines d'entre elles étaient éliminées du fait d'un rapport signal sur bruit trop faible, mais la majeure partie de ces non-détections n'étaient pas dues au rapport signal sur bruit (pour un rapport signal sur bruit trop faible, la segmentation et la classification sont abandonnées). Ensuite, comme nous pouvons le voir, le taux d'alarmes manquées est très élevé alors que le taux de fausses alarmes est acceptable. Ceci est dû au modèle de langage qui, du fait du bruit capté, provoque des erreurs de classification qui tendent à éliminer les mots de détresse plus souvent qu'à en créer (statistiquement cohérent du fait du faible nombre de mots de détresse).

Phrases	Taux d'erreur			Erreur de reconnaissance	
	Détection	Segmentation (parole)	Identification (normal/détresse)	Type	Taux
Global	33.3%	15.2%	37.9%		
Détresse	26.2%	11.3%	56.4%	Alarme manquée	71.4%
Normale	40.5%	20%	12.5%	Fausse alarme	12.5%

TAB. 5.6 – Performances globales du système de reconnaissance de la parole

5.7.1.2.2 Les sons de la vie courante

En ce qui concerne les sons de la vie courante, cette expérimentation nous a permis de conclure que les bruits de pas ont une incidence sonore très faible et que la distance entre le pas et le micro (du fait que ceux-ci sont au plafond) ne permet pas une détection de ces bruits de pas. Nous ne savons donc pas si ceux-ci seraient bien classés ou non. Il en est de même pour les bruits d'eau dans la cuisine. Augmenter le gain du microphone pour les détecter provoquerait une saturation dans tous les autres cas (donc très fréquemment, pour la parole, les sons de vaisselle, ...). Il n'est donc pas imaginable d'augmenter de manière démesurée le gain de ce microphone et de détériorer l'ensemble des détections pour détecter ces deux événements. Notre choix se porte sur une bonne détection des autres classes. Les résultats de détection du son (taux d'erreur par classe) sont présentés dans le tableau 5.7.

Sons de la vie courante	Taux d'erreur		
	Détection	Segmentation	Classification
Global	2.52%	6.32%	48.47%
Claquement de porte	0%	4.29%	56.72%
Sonneries de téléphone	0%	0%	50.73%
Son de vaisselle	15%	35.29%	21.21%
Serrure de porte	0%	0%	94.12%
Chute d'objet	0%	0%	15.46%
Cri	0%	2.22%	15.91%

TAB. 5.7 – Performances globales du système de reconnaissance de sons

Ce tableau nous montre les taux de bonne détections, de bonne segmentation puis de bonne classification pour chacune des classes connues par le système. Nous pouvons voir que par exemple, le bruit de la serrure de la porte est toujours mal classé alors qu'il est toujours bien détecté et bien segmenté. De même pour les claquements de porte qui sont mal classés une fois sur deux alors qu'ils sont toujours bien détectés et segmentés. Les sons de vaisselle quant à eux présentent un taux d'erreur plus faible (20%) mais ce taux d'erreur est à prendre en compte sur les signaux déjà bien segmentés. Or, dans 35% des cas, les sons de vaisselle sont mal segmentés et considérés comme de la parole. Les sonneries de téléphones sont mal reconnues (la base n'était pas adapté au type de sonneries utilisé), avec un taux d'erreur de 50%. Les chutes et les cris ne présentent par contre qu'un taux d'erreur de 15% et un taux de mauvaise segmentation nul ou très faible.

5.7.1.2.3 Résultats globaux

Le tableau 5.8 représente la matrice de confusion du système de reconnaissance de sons et de la parole. Les valeurs en gras (la diagonale) correspondent aux données correctement classifiées. Cette table met en exergue la proximité entre certaines classes, qui a pour conséquence, dans un environnement bruyé, de nous donner de fréquentes mauvaises classifications. Cette situation se retrouve par exemple pour la sonnerie de téléphone, sonnerie récente et non connue du système lors de l'expérimentation, qui se retrouvent dans la moitié

des cas classée en cri. De même les sons de vaisselle se retrouvent souvent mal segmentés et sont alors détectés comme de la parole. Enfin, la serrure de la porte du HIS n'était pas de même type que celles utilisées dans les bases d'apprentissage. Cela se retrouve dans cette matrice qui montre un fort taux d'erreur (plus de 80% au total) sur cette classe.

	Résultats								
	Claquement	Pas	Téléphone	Vaisselle	Serrure	Bris	Chute	Cri	Erreur segmentation
Action	Claquement de porte	42.03%	0%	0%	5.80%	0%	0%	49.28%	2.90%
	Sonnerie de téléphone	1,45%	0%	49.28%	0%	0%	0%	49.28%	0%
	Son de vaisselle	0%	0%	0%	50.98%	0%	0%	1.96%	43.14%
	Serrure de porte	11.59%	2.90%	0%	7.25%	5.80%	1.45%	71.01%	0%
	Chute d'objet	6,67%	0%	0%	0%	0%	6.67%	84.44%	2.22%
	Cri	0%	0%	2.27%	4.55%	0%	2,27de%	0%	79.55%

TAB. 5.8 – Matrice de confusion pour la reconnaissance des sons. Les valeurs en gras correspondent aux sons correctement classifiés.

Les sons de pas ainsi que les sons de bris de vaisselle n'apparaissent pas dans les actions soit parce qu'ils sont trop faibles pour être détectés (pour les premiers) soit parce qu'ils n'ont pas pu être généré (pour les seconds) pour des raisons matérielles.

5.7.2 Expérimentation 2

Au préalable de la réalisation de cette seconde expérimentation, le logiciel a été fortement modifié pour améliorer la détection. La première modification apportée a été de permettre le réglage d'un seuil de détection beaucoup plus bas. L'algorithme de calcul du bruit ambiant a été également modifié après identification et correction d'une erreur de programmation. Cette partie améliore également la détection (du fait que la détection se base sur ce calcul du bruit pour utiliser le seuil variable).

5.7.2.1 Protocole

5.7.2.1.1 Reconnaissance sonore

Le protocole pour la reconnaissance sonore a été modifié afin de prendre en compte les indications apportées par la première expérimentation. Nous avons donc réduit le nombre des actions à effectuer. Le sujet rentre toujours dans le HIS et referme la porte derrière lui, puis effectue le protocole suivant :

- Tirer la chasse d'eau des toilettes et se servir du robinet de la salle de bains (3 fois)
- Fermer la porte de la salle de bain (3 fois)
- Crier « Aïe » (3 fois)
- Remuer la cuillère dans la tasse à café (3 fois)
- Faire tomber un livre de la table de la cuisine (3 fois)

Une fois le scénario et la parole effectués, le sujet ressortait et refermait la porte du HIS ce qui marquait la fin de l'expérimentation en cours.

5.7.2.1.2 Reconnaissance de la parole

La reconnaissance de la parole a été également modifiée dans cette expérimentation afin d'augmenter le nombre de phrases prononcées. Dans celle-ci, trente phrases (vingt de détresse et dix anodines) ont été mélangées de cinq manières différentes. Quinze phrases étaient à prononcer dans la cuisine et quinze dans le salon. Toujours dans le salon, le sujet devait répondre au téléphone trois fois, et chacune des fois lire une petite conversation de cinq phrases. Pour ces conversations, cinq conversations plausibles différentes ont été créées et, pour chaque manière de générer les phrases précédemment, trois conversations parmi les cinq lui étaient associées.

Le sujet avait donc, dans ce scénario, un ensemble de vingt phrases de détresse et de vingt-cinq phrases anodines à énoncer.

5.7.2.2 Résultats expérimentaux

Cette expérimentation s'est faite avec 13 sujets jeunes, en bonne santé et volontaires (3 femmes et 10 hommes). Les caractéristiques de cette population étaient les suivantes : âge : 33 ± 12 ans, poids : 64 ± 20 kg, taille : $1,74 \pm 0,06$ m). Le nombre de sons collectés s'élève à 5417 (2399 d'entre-eux ont été éliminés car leur rapport signal sur bruit était inférieur à 5 dB), avec un RSB de $12,5 \pm 5,6$ dB. Après classification, nous avons conservé 1820 sons avec un RSB moyen de $13,6 \pm 6,5$ dB.

Le tableau 5.9 détaille les résultats de cette expérimentation sous la forme d'une matrice de confusion. Encore une fois, celle-ci nous montre les taux de bonne détection (en gras) ainsi que les classes proches (éléments élevés hors diagonale). Ces premières améliorations ont permis d'éviter les non détections des sons et de la parole. Par contre, nous avons toujours un taux d'alarmes manquées de 62% avec par contre un taux de fausses alarmes de seulement 3.49%. Encore une fois les bruits de vaisselle sont proches de la parole, ce qui amène un taux de mauvaise segmentation élevé (plus de 50%) pour la classe bruits de vaisselle. Nous remarquons également que dans 20% des cas, le système classe mal les sons de claquement de porte ou de chute d'objets, sons très proches (courts et forts). La classification de la parole ici se réfère à la classification C3 présentée précédemment.

		Résultats									
		Claquement	Pas	Téléphone	Vaisselle	Serrure	Bris	Chute	Cri	Parole normale	Parole détresse
Action	Claquement de porte	81.25 %	0 %	0 %	0 %	0 %	0 %	18.75 %	0 %	0 %	0 %
	Sonnerie Téléphone	0 %	0 %	100 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %
	Bruit de vaisselle	0 %	0 %	0 %	42.86 %	0 %	0 %	0 %	4.76 %	52.38 %	0 %
	Chute d'objet	19.05 %	0 %	0 %	4.76 %	0 %	0 %	76.19 %	0 %	0 %	0 %
	Cri	8.7 %	0 %	0 %	8.7 %	0 %	0 %	30.43 %	30.43 %	21.74 %	0.00 %
Parole normale		0.74 %	0 %	0.37 %	4.1 %	0 %	0 %	3.35 %	4.48 %	83.44 %	3.49 %
Parole détresse		0.74 %	0.37 %	0 %	2.4 %	0.37 %	0 %	3.35 %	0 %	62.92 %	29.85 %

TAB. 5.9 – Matrice de confusion pour la reconnaissance de sons et de la parole. Les valeurs en gras représentent les bonnes classifications.

Cette matrice nous montre l'impact des modifications effectuées. En effet, le seuil de détection influence fortement la détection du début du son ainsi que sa bonne segmentation

et classification. La détection des sonneries de téléphone donne de bien meilleurs résultats du fait de l'utilisation d'une sonnerie plus conventionnelle (mais d'un téléphone moderne). Les bruits de vaisselle restent toujours très mal segmentés. En ce qui concerne les cris, le taux de bonne classification chute énormément et de nombreux cris se retrouvent classés en chute. Cette matrice nous montre également une partie réservée à la parole, les deux dernières lignes, avec une très bonne classification pour la parole normale, par contre la parole de détresse est bien segmentée mais mal classée (30% de bonne classification uniquement). Ceci est dû aux modèles de langage utilisés qui sont améliorés continuellement.

5.7.3 Expérimentation 3

5.7.3.1 Protocole

Le protocole de cette dernière expérimentation était le même que la précédente, à l'exception que cette fois-ci toutes les phrases étaient énoncées dans le salon. Entre la seconde et la troisième expérimentation, le plus grand changement a été le positionnement des microphones. En effet, le placement sur l'antenne des quatre microphones pour la chambre et le salon était peu avantageux et nous espérons de bien meilleurs résultats en répartissant mieux les microphones dans les deux pièces. Le placement lors de cette expérimentation est donc conforme à ce qui est montré sur la figure 5.2. Nous avons également baissé les stores du salon pour réduire le phénomène d'écho qui se crée du fait de la disposition l'une en face de l'autre des deux vitres. Les actions à réaliser ainsi que le nombre de phrases à prononcer est par contre resté identique.

5.7.3.2 Résultats expérimentaux

Dix sujets en bonne santé et volontaires ont participé à cette dernière expérimentation, 3 femmes et 7 hommes âgés en moyenne de $37,2 \text{ ans} \pm 14$, et dont les mensurations étaient $69 \text{ kg} \pm 12$ et $1,72 \text{ m} \pm 0,08$). L'expérimentation a permis la collecte de 3164 sons (dont 2019 éliminés car leur rapport signal sur bruit était inférieur à 5 dB avec un rapport signal sur bruit moyen de $12,65 \pm 5,6 \text{ dB}$). Après classification, nous avons retenu 1008 sons avec un rapport signal sur bruit de $14,4 \pm 6,5 \text{ dB}$. La première remarque que nous pouvons faire sur ces statistiques est une augmentation flagrante du rapport signal sur bruit depuis le déplacement des microphones. Cela est dû à un meilleur positionnement et une meilleure isolation des microphones du salon et de la chambre.

	S1	C1	C2	S2	C3
Global	8.3 %	33.4 %	34.5 %	6 %	30.5 %
Anodin	9.6 %	10.4 %	10 %	6.9 %	9.6 %
Détresse	7 %	60.1 %	63.1 %	4.3 %	54.8 %

TAB. 5.10 – Taux d'erreur de segmentation/classification pour la reconnaissance des phrases anodine/détresse

Le tableau 5.10 présente les résultats expérimentaux liés à la classification de la parole. Dans ce tableau, nous présentons les résultats à l'aide des trois algorithmes de classification introduits précédemment, et donc les résultats de segmentation associés. S1 est associée à C1 et C2 alors que S2 est associée à C3. Ce tableau nous montre l'influence de la prise en compte des différentes phrases et des différents microphones sur la segmentation et la classification. En effet, le fait de prendre en compte plusieurs microphones réduit considérablement le taux de mauvaise segmentation, pour le passer de 8% à 6% en global, et améliore également la classification. En effet, passer de la classification C2 à la classification C3 permet de réduire de 10% le taux d'alarmes manquées, ce qui n'est pas négligeable. Ceci améliore également, plus faiblement, le taux de fausses alarmes.

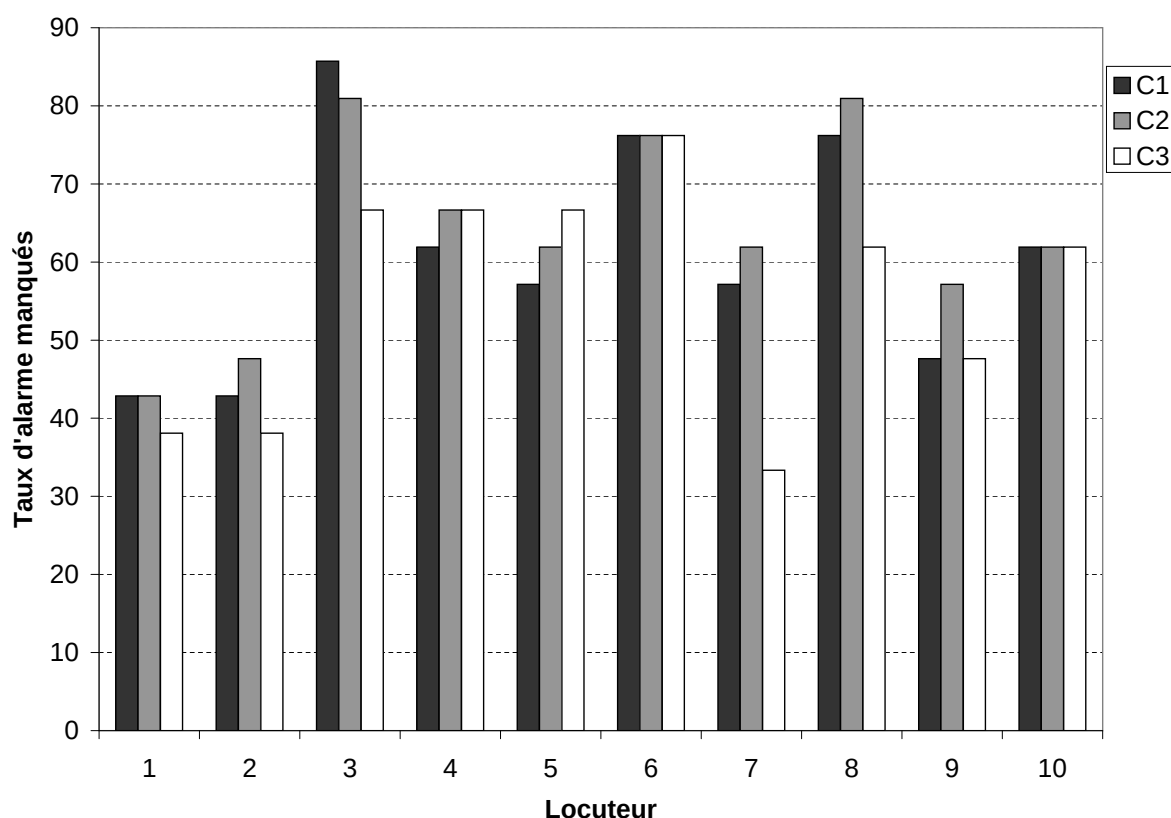


FIG. 5.10 – Résultat de la détection des mots de détresse par locuteur

Les algorithmes de traitement de la parole sont supposés indépendants du locuteur. Cependant, la figure 5.10 montre que, selon la personne, les taux de mauvaise classification sont plus ou moins élevés. Par exemple, les locuteurs 3 à 6 ont des performances moins bonnes que les locuteurs 1, 2 et 7. Ceci peut s'expliquer, en plus de certains « accents » que peuvent avoir les sujets (prosodie), par le bruit ambiant qui est différent d'une expérimentation à une autre.

Le tableau 5.11 nous présente la nouvelle matrice de confusion à la suite de ces expérimentations. Il n'y a pas vraiment de différence notable avec la précédente, excepté sur la partie reconnaissance de la parole, où l'amélioration est flagrante depuis que les microphones ont été déplacés.

		Résultats							
		Claquement	Téléphone	Vaisselle	Bris	Chute	Cri	Anodin	Détresse
Action	Claquement de porte	80.1%	0%	0%	0%	19.9%	0%	0%	0%
	Sonnerie de Téléphone	0%	100%	0%	0%	0%	0%	0%	0%
	Son de vaisselle	0%	0%	51%	0%	2.1%	3.5%	43.4%	0%
	Chute d'objet	21%	0%	0%	2%	77%	0%	0%	0%
	Cri	0%	2.4%	4%	2.4%	0%	78%	13.2%	0%
Parole normale		0.4%	0%	2.1%	0%	2.5%	2.9%	90.5%	1.6%
Parole détresse		1.4%	0%	1%	0%	1.4%	0%	51%	45.2%

TAB. 5.11 – Matrice de confusion pour la reconnaissance de sons et de la parole. Les valeurs en gras représentent les bonnes classifications.

5.8 Analyse des résultats de la détection du son et de la parole

Ces expérimentations ont permis d'obtenir des premiers résultats en condition réelle sur le système de reconnaissance de sons et de la parole. Il a permis également à l'équipe GETALP d'améliorer les algorithmes et l'interface graphique, mais aussi de mettre en place les traitements pour en extraire les données.

Les résultats sont prometteurs mais montrent aussi le travail restant à effectuer pour atteindre un système de détection de mots clés de détresse utilisable dans un environnement réel.

En ce qui concerne l'utilisation des données dans le cadre de la fusion multicateurs, ces résultats nous montrent deux choses :

- La première est que le système n'est pas parfait, même en fusionnant les informations venant de plusieurs microphones, nous arrivons toujours à un taux de bonne classification acceptable mais loin de 100%. Les autres capteurs vont permettre de corriger ces défauts pour détecter les bonnes activités.
- La seconde, que nous pouvons déduire de l'observation des matrices de confusion 5.9 et 5.11 est que même si le système se trompe parfois de décision, il est quand même « stable » d'une expérimentation à l'autre, c'est à dire qu'il donne des résultats identiques en taux de bonne détection et classification. Cela va permettre de faire, sur ces données associées aux autres capteurs, de la classification supervisée. Pour appliquer ce genre d'algorithmes, il faut en effet que le comportement du système de reconnaissance de sons et de la parole soit répétable dans le temps.

Ce système est en constante amélioration du fait des travaux de l'équipe GETALP. Des algorithmes d'annulation de sources parasites et d'amélioration du rapport signal sur bruit sont en développement. Les modèles de langage font également l'objet d'un travail constant qui permet à chaque fois d'améliorer les taux de bonne classification. Une meilleure adaptation du modèle de langage devrait voir le jour avec les futurs projets du laboratoire.

Un prélude à la classification de données : formatage et interprétation des données

Sommaire

6.1 Le HIS de Grenoble	130
6.1.1 Présentation générale	130
6.1.2 Positionnement des capteurs	130
6.1.3 Architecture, connectivité et acquisition des données	131
6.1.4 Matériel de la pièce technique	132
6.2 Les données actimétriques	133
6.2.1 Données brutes	133
6.2.2 Extraction des paramètres	133
6.3 Les données sonores	133
6.3.1 Données brutes	133
6.3.2 Extraction des paramètres	134
6.4 Les données de présence	134
6.4.1 Données brutes	134
6.4.2 Extraction des paramètres	135
6.5 Les données liées aux instruments de la vie quotidiennes	135
6.5.1 Données brutes	135
6.5.2 Extraction des paramètres	136
6.6 Les données environnementales	136
6.6.1 Données brutes	136
6.6.2 Extraction des paramètres	137
6.7 Conclusion	138

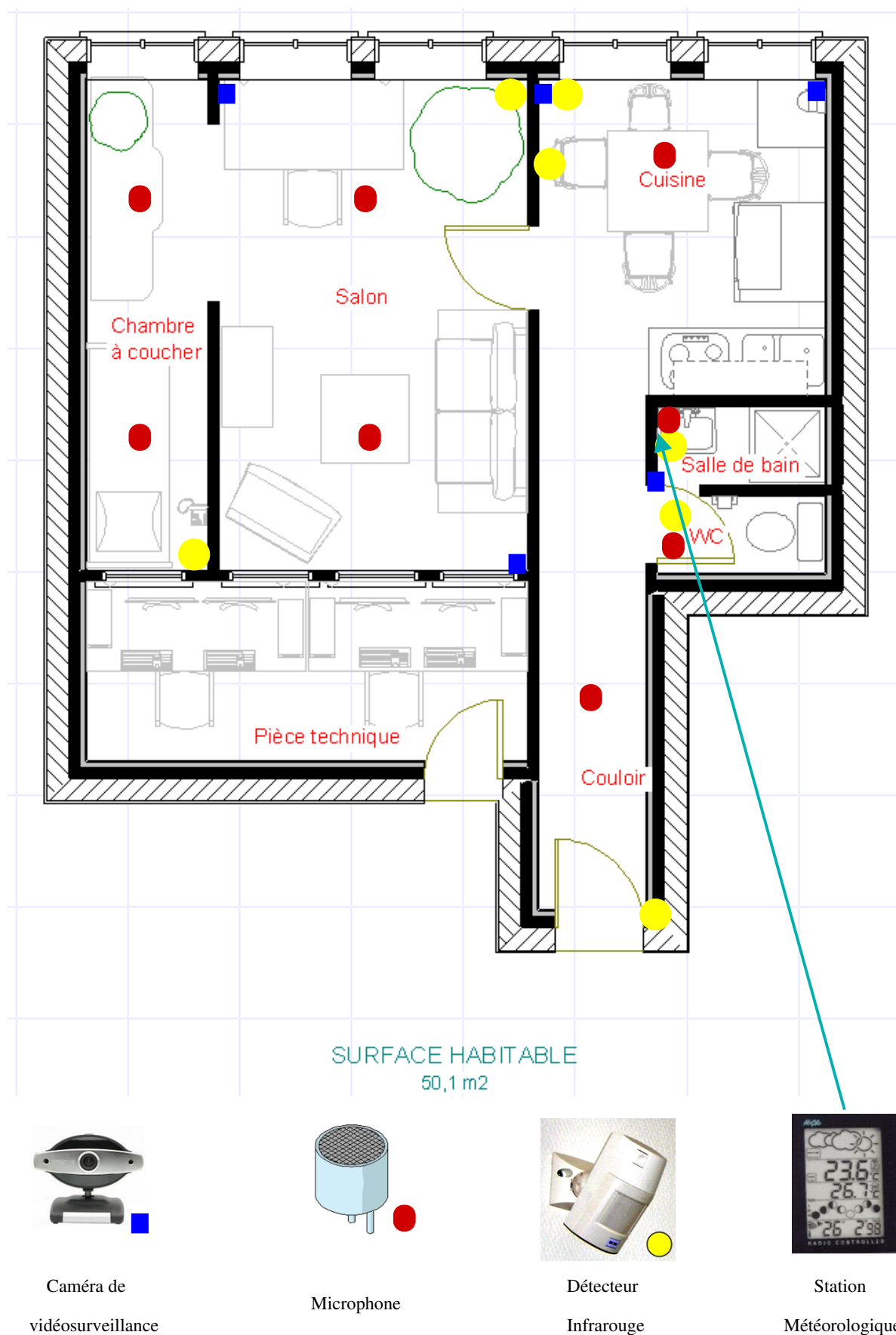


FIG. 6.1: Positionnement des différents équipements dans l'Habitat Intelligent pour la Santé de Grenoble

6.1 Le HIS de Grenoble

6.1.1 Présentation générale

Le HIS (Habitat Intelligent pour la Santé) du laboratoire TIMC-IMAG est installé à la Faculté de Médecine de Grenoble depuis 1999. Par ailleurs, le laboratoire, avec le projet AILISA, a installé plusieurs habitats à travers la France (Paris, Toulouse, Grenoble), que ce soit au CCAS¹ (Résidence Notre-Dame de Grenoble) ou dans des chambres hospitalières (à Grenoble à l'Hôpital Sud, à Toulouse et à Paris). Ces différentes chambres sont équipées principalement de détecteurs infrarouges et ont permis de mesurer des rythmes d'activités sur le long terme.

L'appartement HIS, dans lequel se sont déroulées les expérimentations décrites dans ce manuscrit, mesure 50m² et est équipé comme un appartement classique avec toutes les commodités. Il possède un salon (avec canapé, TV et table basse), une chambre (avec un lit et une commode), une cuisine (avec micro-onde, réfrigérateur, table et chaises, plaques de cuisson électriques, évier, rangements, couverts) ainsi que des toilettes et une salle de bain (lavabo et douche). Le plan d'aménagement de l'appartement avec les principales cotes est donné dans l'annexe C.

À l'intérieur de cet appartement, divers capteurs ont été intégrés afin de surveiller les activités de la vie quotidienne de l'occupant :

- Des détecteurs infrarouges, pour surveiller les endroits stratégiques de l'appartement correspondant à des activités (par exemple le lit pour l'activité « sommeil », le canapé pour l'activité « détente », ...).
- Huit microphones répartis dans l'appartement et reliés à un système de reconnaissance de sons et de paroles décrit au chapitre précédent.
- Un capteur de température et d'humidité de l'air, placé dans la salle de bain, et mesurant toutes les cinq minutes ces deux variables environnementales.
- Des capteurs d'ouverture et de fermeture de porte, placés sur le réfrigérateur, le placard de la cuisine ainsi que sur la commode dans la chambre.
- Un capteur embarqué, ACTIM6D, placé sur le sujet et enregistrant les données provenant des accéléromètres et des magnétomètres (données cinématiques).
- Cinq web-caméras grand-angle (85° d'angle d'ouverture), reliées à des ordinateurs de la salle technique (tous synchronisés par NTP) et qui acquièrent les images de l'ensemble de l'appartement et les horodatent lors de la compression. Ces caméras ne sont utilisées qu'à des fins d'indexation.

6.1.2 Positionnement des capteurs

La figure 6.1 donne un aperçu des différentes positions de chacun des capteurs. En plus des capteurs intégrés sur cette figure, le sujet porte, dans son tee-shirt, la carte ACTIM6D.

Le positionnement des capteurs a été choisi en fonction des critères suivants :

¹Centre Communal d'Action Social

- Les caméras de vidéosurveillance ont été placées de telle manière qu’aucune zone de l’appartement ne soit invisible, de façon à indexer correctement les actions. Cependant, pour des raisons éthiques évidentes, la salle de bain et les toilettes possèdent une caméra mais celle-ci ne filme que la porte (et aucune autre partie). Il a donc été décidé que selon sa manière de fermer la porte, le sujet est aux toilettes ou bien à la salle de bain.
- Un contacteur magnétique est placé sur la porte d’entrée afin d’en détecter l’ouverture.
- Les capteurs infrarouges sont au nombre de six. Ils ont été placés de manière à surveiller des « fonctions » de l’habitat. Ainsi, celui de la cuisine surveille principalement la table, un autre surveille le couloir, un autre surveille le lit, un autre surveille la partie détente (salon), un autre surveille les toilettes et le dernier surveille la douche.
- Les microphones ont été placés de telle manière à ce qu’ils quadrillent de manière homogène l’appartement dans sa globalité. Le chapitre 5 a déjà détaillé ce placement et les raisons des changements apportés au cours du temps par rapport au placement initial.
- La station météorologique est, quant à elle, placée de telle sorte qu’elle détecte des changement de température et d’humidité créés dans la salle de bain lors de tâches d’hygiène (la douche par exemple).
- Le capteur ACTIM6D est placé sur le sujet, sous son aisselle gauche. Ce placement a été expliqué dans le chapitre 4.
- Enfin les contacteurs de porte ont été placés à des endroits stratégiques correspondant aussi à des actions de la vie quotidienne. Ainsi, l’un se trouve dans la commode servant à entreposer les vêtements du sujet, le second se trouve sur la porte du réfrigérateur et le dernier sur la porte du placard servant au rangement des aliments.

6.1.3 Architecture, connectivité et acquisition des données

La connectivité des HIS, tels qu’installés dans le projet AILISA, se base sur le bus CAN (Controller Area Network, norme ISO 11898), un bus de terrain permettant une communication moyenne distance (20 ko/s à 1km). Ce bus est multi-maître et est surtout adapté à la communication de messages courts. La transmission sur le bus se fait en se basant sur un octet de priorité. Le plus fort octet peut émettre en premier lorsqu’il requiert la ligne en même temps qu’un autre émetteur plus faible. Un message d’alerte est alors émis de manière prioritaire. Ce type de bus montre donc un aspect déterministe.

Le bus CAN du HIS est montré sur la figure 6.2. Un récepteur radio se charge de recevoir les données des différents capteurs infrarouges puis les transmet à la carte SmartCAN2 (figure 6.3) que nous avons programmée de telle sorte qu’elle transmette les données reçues sur le bus CAN dès la commutation du relais. Le branchement de la SmartCAN est réalisé par sept relais, un pour chaque détecteur infrarouge, qui ne peuvent être activés simultanément. Le programme C à l’intérieur du PIC de la SmartCAN se résume donc à la scrutation bit à bit des valeurs de certaines de ces entrées numériques et à la détection des changements. La gestion du bus CAN permet d’avoir une détection d’erreur et une réémission dans

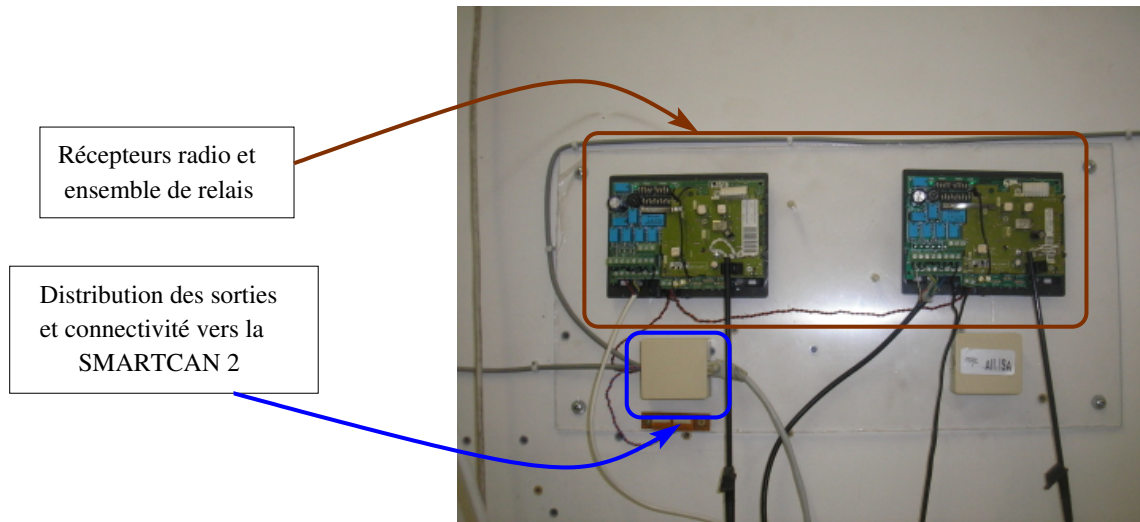


FIG. 6.2: Récepteur radio et bus CAN du HIS

ce cas. Aucune trame n'est donc perdue en cas de mauvaise transmission. De plus, un système d'acquittement permet de s'assurer que le message a bien été reçu.



FIG. 6.3: Carte SmartCAN 2 gérant la transmission des données des détecteurs infrarouges sur le bus CAN

Outre les détecteurs infrarouges, le pèse-personne sans fil ainsi que l'oxymètre de pouls se connectent également sur le bus CAN et sont gérés par le même logiciel. Chaque type de capteur possède son niveau de priorité différent afin de pouvoir transmettre ses trames.

6.1.4 Matériel de la pièce technique

Du fait de certains capteurs nécessitant des ressources importantes (les caméras et le son) le traitement des capteurs du HIS se retrouve réparti sur quatre ordinateurs. Ces quatre stations (classiques) sont maintenues à l'heure par le protocole NTP afin d'assurer une bonne synchronisation entre toutes les données et d'éviter tout problème d'interprétation.

6.2 Les données actimétriques

6.2.1 Données brutes

Le chapitre 4 a présenté l'acquisition des données cinématiques ainsi que leur traitement, permettant d'aboutir à un fichier de sortie décrivant les changements de posture de la personne et les périodes de marche. Grâce à ce traitement, nous disposons donc, dans le fichier de sortie, du début et de la fin de chaque fenêtre temporelle à l'intérieur de laquelle la personne se déplace, et de la position temporelle de chaque transition posturale : Assis-Debout, Debout-Assis, Debout-Allongé et Allongé-Debout.

6.2.2 Extraction des paramètres

Les paramètres qui nous importent pour ce capteur sont :

- Le temps passé dans chaque posture (assis, debout, allongé), à l'intérieur de la fenêtre temporelle considérée. Ce temps est comptabilisé en reconstruisant un signal, à partir du fichier de sortie, et en calculant le pourcentage de temps passé dans chaque position. Ce paramètre nous donne des informations sur le type d'activité effectué (qui demandent chacune d'être assis, debout ou couché pendant certaines périodes).
- Le pourcentage de temps passé à marcher à l'intérieur de la fenêtre d'analyse. Ce second paramètre nous donne la « mobilité » de la personne durant l'activité.

6.3 Les données sonores

6.3.1 Données brutes

Le chapitre 5 a présenté le système de reconnaissance de sons et de paroles du HIS. Ce système fournit en sortie un ensemble de fichiers qui est par la suite traité sous Matlab™ en vue d'obtenir un fichier unique contenant l'ensemble des informations nécessaires.

Pour rappel, ce fichier de sortie contient, pour chacun des sons détectés dans l'appareil, les informations suivantes :

- Le numéro du microphone ayant effectué la détection avec le meilleur rapport signal sur bruit,
- Le rapport signal sur bruit du signal détecté,
- La date et l'heure de la détection,
- L'identifiant unique du signal sonore (du microphone ayant le meilleur RSB),
- La classe du signal sonore après fusion des informations émanant des autres microphones. Dans cette expérimentation, la parole, qu'elle soit de type anodine ou de type détresse, appartiendra à la classe unique « Parole ». Nous avons donc un ensemble de neuf classes, les huit premières étant les classes de sons déjà présentés et la dernière étant la classe parole.

Ce fichier est ensuite utilisé pour l'extraction des paramètres. En effet, cette suite de sons n'aurait pas vraiment de sens au niveau de la classification. Nous allons donc extraire certaines informations pour chaque fenêtre temporelle afin d'avoir des données plus significatives à traiter.

6.3.2 Extraction des paramètres

Deux paramètres vont être significatifs pour la classification des activités. Le premier est la pièce dans laquelle un son a été le mieux détecté (qui sera dans la plupart des cas la pièce dans laquelle ce son a été produit), ainsi que la classe de ce son. Pour chaque fenêtre temporelle, à partir du fichier de sortie précédemment présenté, deux vecteurs sont ainsi construits :

- Le premier contient le nombre d'évènements sonores détectés sur chacun des microphones. Ce vecteur contient sept entrées : couloir, cuisine, salle de bain/toilette, salon fenêtre, salon régie, chambre fenêtre et enfin chambre régie.
- Le second contient le nombre d'évènements sonores détectés dans chacune des classes de son. Ce vecteur contient, quant-à-lui, neuf entrées : son de pas, son de vaisselle, claquement de porte, serrure de porte, bri de verre, paroles (regroupant les paroles anodine et de détresse), cri, sonnerie de téléphone et enfin chute d'objet.

Ces deux vecteurs vont ainsi nous donner l'ensemble des informations nécessaires sur le type de son et la localisation des sons dans la fenêtre temporelle analysée.

6.4 Les données de présence

6.4.1 Données brutes

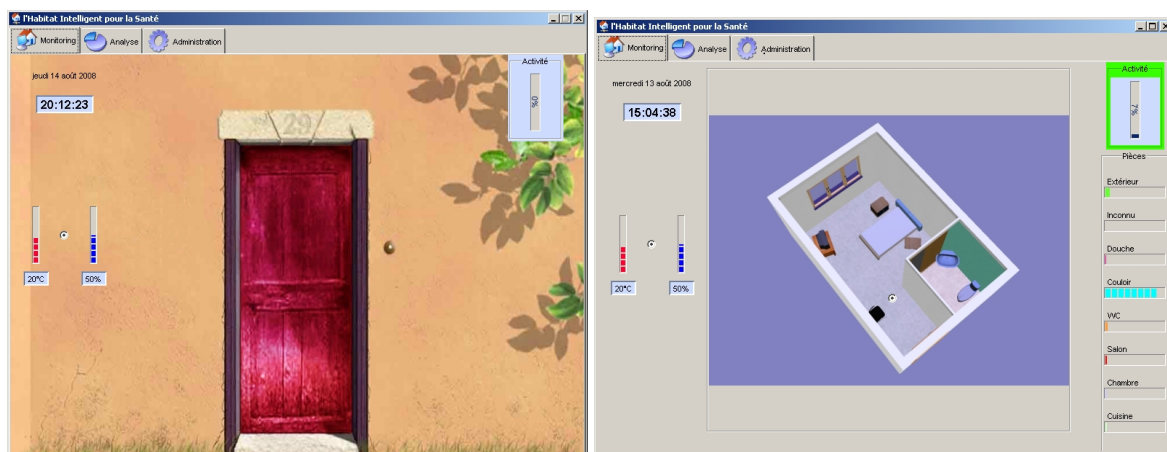


FIG. 6.4: Interface Homme-Machine du logiciel AILISA de recueil et de traitement des données de l'HIS

Les données de présence sont émises par les sept détecteurs infrarouges présents dans l'appartement. Ces détecteurs captent un changement de température dans leur cône de détection et, lors de ce changement, vont alors envoyer une trame au récepteur situé dans

la régie. La réception de cette trame va engendrer la mise à 1 d'un des bits en entrée de la SmartCan 2 et ainsi l'envoi d'une trame sur le bus CAN, en direction du logiciel AILISA.

Le programme embarqué dans la SmartCan 2 est relativement simple. Il détecte les changements de niveau logique sur chaque entrée de la carte et se charge, lors de la détection du front descendant, d'envoyer une trame contenant le numéro de la pièce avec un format spécifique décrit par le logiciel. En effet, les capteurs de présence vont cohabiter, sur le bus CAN, avec différents matériels tels que l'oxymètre ou le pèse personne sans fil.

Le logiciel AILISA, programmé par Christophe Villemazet [Villemazet, 2004], prend en charge l'enregistrement, le stockage et le traitement des données de l'habitat. La figure 6.4 présente l'interface homme-machine de ce logiciel. Celui-ci, installé dans les habitats du projet AILISA ainsi que dans le HIS de la faculté de médecine, se charge de l'enregistrement des détections et valeurs physiologiques dans une base de données MySQL et permet d'administrer cette base pour en extraire les données sous différentes formes. Le logiciel envoie également un mail par jour sous forme de compte rendu, avec les informations de la base pour les dernières vingt-quatre heures, ainsi que des mails de « secours » dans certaines situations (lorsqu'il n'y a aucune détection pendant plus d'une heure par exemple).

Le logiciel AILISA permet d'extraire des données de la base de données dans de multiples formats (texte, CSV, Matlab par exemple).

6.4.2 Extraction des paramètres

L'extraction des paramètres fournis par ces capteurs va se faire en construisant deux vecteurs à six éléments, pour représenter, d'une part le pourcentage de temps passé dans chacune des pièces pendant la fenêtre temporelle considérée, et d'autre part le nombre de détections à l'intérieur de cette pièce. En effet, une activité va concerner un nombre limité de capteurs (étant donné qu'elle se fait principalement dans une pièce).

Les pièces surveillées par le logiciel sont les suivantes : 1 : Entrée, 2 : Douche, 3 : Couloir, 4 : WC, 5 : Salon, 6 : Chambre, 7 : Cuisine.

Le premier capteur, un contacteur sur la porte d'entrée, est peu utile dans notre cas car il concerne surtout la détection de l'absence de la personne ou de l'entrée d'une seconde personne. Son utilisation intervient bien en amont de la classification des activités. Les autres détecteurs vont être révélateurs d'une ou plusieurs activités. La collecte des informations des autres capteurs permettra de donner plus précisément l'activité qui correspond à une occupation de la pièce. Le second vecteur va nous donner la « mobilité » de la personne induite par cette activité.

6.5 Les données liées aux instruments de la vie quotidiennes

6.5.1 Données brutes

Le réseau de détection de présence du HIS, tel qu'il est dessiné à l'heure actuelle, ne peut accueillir simplement plus de sept capteurs (le nombre de capteurs infrarouges ins-

tallés). Pour rajouter ces nouveaux capteurs, nous avons donc placé des caméras de vidéo-surveillance (dans les placards) et traité les images extraite de l'enregistrement vidéo pour obtenir des informations de type « Tout ou Rien ».

Chaque caméra enregistre l'intérieur de l'objet concerné. Les enregistrements se font à la vitesse et la résolution les plus basses possibles pour le modèle considéré (15 images par secondes en 320×240). Ces vidéos sont automatiquement datées au début de l'enregistrement, leur nom contenant l'heure précise du début. Il suffit donc d'utiliser cette heure pour connaître l'heure de commencement, et d'en déduire l'heure actuelle dans la vidéo grâce à la fréquence d'échantillonnage.

La détection des périodes d'ouverture et de fermeture est très simple. L'image est d'abord convertie de RGB en niveau de gris. Ensuite elle est binarisée avec un niveau de 0,2 déterminé par divers enregistrements faits avec des éclairages différents. Ce niveau est exprimé sur l'échelle des niveaux de gris, avec 0 pour le noir et 1 pour le blanc, et 256 niveaux entre deux. Une fois cette binarisation effectuée, dans le cas où plus de la moitié des pixels de l'image sont à 1, la porte est déclarée comme ouverte, fermée sinon. Selon l'état précédent, le programme calcule l'heure à laquelle l'image a été tournée et écrit dans le fichier de sortie la date et le nouvel état de la porte. Le fichier de sortie d'une des expérimentation effectuée pour le réfrigérateur a la forme suivante :

06-08-2008-09-17-17	0
06-08-2008-10-11-23	1
06-08-2008-10-11-28	0
06-08-2008-10-26-12	1
06-08-2008-10-26-16	0

6.5.2 Extraction des paramètres

L'extraction des paramètres dans le cas de ce capteur va être relativement basique. L'information désirée est de savoir si la porte considérée est utile ou non dans l'activité en cours. Pour ceci, nous allons donc comptabiliser le nombre d'ouvertures de cette porte. Un deuxième paramètre, qui va nous permettre de savoir si la porte a été mal refermée ou non par exemple, va être de calculer la position la plus utilisée dans la fenêtre. Par exemple, si le réfrigérateur a passé les trois quart de son temps ouvert, ce deuxième paramètre sera mis à 1. Cela va permettre de corriger d'éventuels cas d'oubli de fermeture de porte dans des activités qui ne nécessiteraient pas l'ouverture de celui-ci.

6.6 Les données environnementales

6.6.1 Données brutes

Les données environnementales sont recueillies par une station météorologique (figure 6.5, WS8610R, La Crosse Technologie) qui donne une mesure, toutes les cinq minutes, de

deux paramètres : la température ainsi que le taux d'humidité dans l'air. Les données brutes se présentent dans un fichier de la manière suivante :

Date	Indoor Temperature	Indoor Humidity
04.02.2008 15:55	19.3	44
04.02.2008 16:00	19.3	44
04.02.2008 16:05	19.3	44
04.02.2008 16:10	19.3	44
04.02.2008 16:15	19.3	44
04.02.2008 16:20	19.3	44
04.02.2008 16:25	19.3	44

Cette station permet l'enregistrement des valeurs de température et d'hygrométrie. Il faut par contre la connecter régulièrement à un ordinateur (par port série) pour récupérer les enregistrements.



FIG. 6.5: Station météorologique utilisée pour l'enregistrement des données environnementales dans le HIS

6.6.2 Extraction des paramètres

Les paramètres surveillés par cette station météorologique sont à variation très lente et vont surtout nous renseigner sur l'utilisation de la salle de bain (lors de la prise de douche par exemple, le taux d'humidité va fortement augmenter). De plus, ces paramètres dépendent fortement de la période de la journée, de l'année et de la localisation. Ainsi, en plein mois d'août à Grenoble, la température et l'humidité seront très différentes de la mesure à la même période au Havre. Il nous faut donc une représentation qui ne dépende ni de la période de l'année ni de la localisation géographique. L'intérêt pour nous se porte surtout sur la variation au cours du temps de ces deux mesures. Nous allons donc sommer le carré des différences des trois dernières mesures de ces deux paramètres et les utiliser comme variables

dans notre cas. En temps normal et sans utilisation forte de la salle de bain, les variations devraient être négligeables pendant la période de temps considérée (15 minutes).

6.7 Conclusion

Ce court chapitre présente la mise en forme des données et les choix effectués ainsi que les raisons de ces choix. Certains ont été pris empiriquement, à partir de deux points de réflexion :

- Le premier point est de savoir ce que nous désirons extraire de cette donnée et pourquoi nous l'avons mise en place à l'intérieur du HIS.
- Le second point est de savoir quelles types de données seront répétables au cours du temps pour une activité donnée, et ainsi prendra un sens pour cette activité. C'est le passage d'une représentation abstraite des données sous forme textuelle à une signification plus forte de cette donnée à l'intérieur d'une activité, et qui va, on le souhaite, être caractéristique de cette activité seule.

Le chapitre suivant va introduire la méthode de classification utilisée avant que la dernière partie ne décrive les expérimentations menées pour valider l'ensemble de ces choix.

La classification de données : de l'apprentissage aux expérimentations

La vraie logique du monde est celle du calcul des probabilités.

James Clerk Maxwell

Sommaire

7.1 Les activités de la vie quotidienne	141
7.1.1 Définition	141
7.1.2 Critères sur les activités	141
7.1.3 Classification et multimodalité des activités de la vie quotidienne . . .	142
7.2 Rappel sur les Séparateurs à Vaste Marge	143
7.2.1 Le problème de la classification	143
7.2.1.1 La classification : un problème ancien	143
7.2.1.2 Classification non supervisée ou supervisée	143
7.2.2 Les séparateurs à vaste marge pour la classification binaire	144
7.2.2.1 Classification avec séparation linéaire	144
7.2.2.2 Séparateur non-linéaire : Kernel Trick	147
7.2.3 Classification à N classes ($N > 2$)	149
7.2.3.1 Méthode Un-Contre-Tous	150
7.2.3.2 Méthode Un-Contre-Un	150
7.3 Apprentissage et tests dans le HIS	152
7.3.1 Indexation des données	152
7.3.2 Validation croisée	153

LES travaux de cette thèse portent sur la classification des activités de la vie quotidienne d'une personne à son domicile. Les précédents chapitres ont présenté successivement un état de l'art des solutions proposées, la mise en place des différents capteurs et les algorithmes qui y sont associés, et enfin le formatage des données, préalable à la classification. Dans ce chapitre, nous allons introduire la notion d'activité de la vie quotidienne, puis présenter l'algorithme de classification utilisé ainsi que sa mise en place dans le cadre du HIS.

7.1 Les activités de la vie quotidienne

7.1.1 Définition

Ces activités sont effectuées chaque jour machinalement sans que la personne les réalisant n'y porte attention, et pourtant, pour travailler sur les données les décrivant, nous devons définir précisément ce que nous considérons comme « activité de la vie quotidienne ». Les dictionnaires médicaux (anglais) définissent ces activités comme étant « the things we normally do in daily living including any daily activity we perform for self-care (such as feeding ourselves, bathing, dressing, grooming), work, homemaking, and leisure », ce qui nous donne en français « les choses que nous faisons normalement dans la vie de tous les jours, incluant les activités que nous effectuons pour notre propre bien-être (tel que se nourrir, se laver, s'habiller, se maquiller), travailler, effectuer des tâches ménagères et les loisirs ».

En se basant sur ces définitions et sur les critères ADL de Katz (voir chapitre 3), nous avons concentré nos travaux sur les activités de la vie quotidienne suivantes :

- Se nourrir, qui inclut également le fait de préparer le repas puis de nettoyer la cuisine (faire la vaisselle par exemple),
- Dormir, c'est à dire effectuer sa nuit de sommeil ou faire une sieste en plein après-midi,
- S'habiller,
- Se détendre, activité plus vague qui peut se faire en lisant un livre, regardant la télévision, écoutant la radio, ...
- Éliminer, en se rendant aux toilettes régulièrement,
- Soigner son hygiène, en se lavant les mains, se lavant les dents, ...
- Communiquer avec l'extérieur, à l'aide du téléphone.

7.1.2 Critères sur les activités

Lors de nos travaux, aucun critère sur les activités et la manière de les réaliser n'a été donné aux sujets. Le seul critère que nous nous sommes fixés pour les algorithmes de classification est de considérer des fenêtres temporelles à classer d'une durée de trois minutes. Ce critère se base sur les activités les plus courtes que nous pourrions rencontrer et qui vont en moyenne avoir cette durée. Cela permet aussi d'avoir un nombre suffisant de fenêtres pour les activités plus longues.

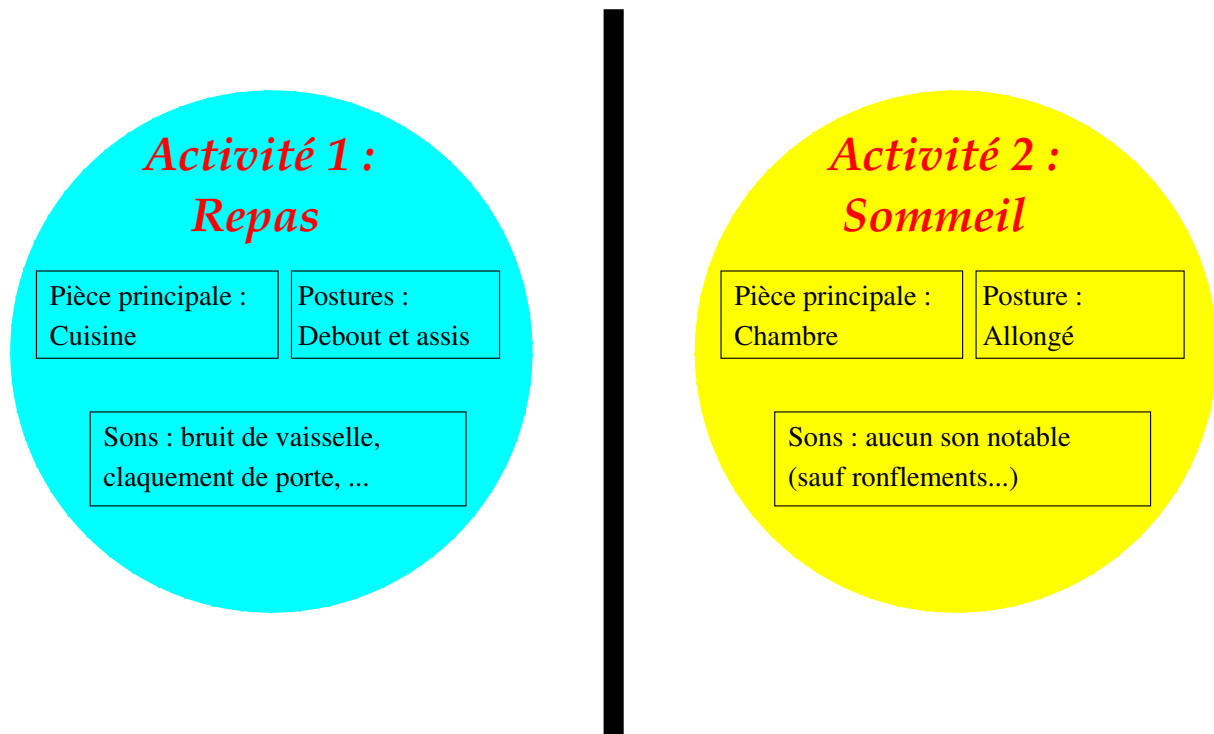


FIG. 7.1: Représentation sous forme de plusieurs modalités de deux des activités de la vie quotidienne

En considérant une personne âgée « en bonne santé », les activités les plus courtes seront pour elle de s'habiller, d'aller aux toilettes et les tâches d'hygiène. S'habiller et les tâches d'hygiènes durent bien au minimum trois minutes, le temps de choisir les vêtements pour s'habiller puis enlever les vêtements précédents, ou bien rien que le temps de se laver les dents pour l'hygiène dure les trois minutes. En ce qui concerne aller aux toilettes, le temps d'effectuer l'activité et ensuite de se laver les mains remplit encore la condition.

7.1.3 Classification et multimodalité des activités de la vie quotidienne

Les activités de la vie quotidienne sont donc des activités complexes demandant de nombreuses actions à la personne. Elles vont aussi induire un ensemble de valeurs différentes pour les capteurs installés dans l'appartement. La figure 7.1 illustre ces propos en retenant trois modalités. Ainsi, pour une première activité qui est la prise d'un repas, la pièce la plus occupée sera logiquement la cuisine, la personne sera debout ou bien assise selon les instants de la réalisation de l'activité, et les sons captés vont inclure des bruits de vaisselle, des claquements de porte pour les placards... Au contraire, l'activité numéro deux qui est le repos se déroulera principalement dans la chambre, la personne sera allongée et en toute logique (sauf pour les personnes ronflant ou parlant dans leur sommeil), aucun son ne sera capté durant cette activité.

Le but de nos algorithmes est donc de prendre en compte les informations de l'ensemble des capteurs afin de trouver une séparation entre les différentes activités. Le chapitre précédent donnait la définition des modalités que nous utiliserons pour la classification. La sec-

tion suivante va se concentrer sur les aspects mathématiques de la définition des séparations entre les différentes activités et sur les algorithmes permettant d'avoir un calcul optimal de celle-ci afin que toute nouvelle fenêtre d'une activité soit classée dans la zone correspondante.

7.2 Rappel sur les Séparateurs à Vaste Marge

Nos travaux se basent sur l'utilisation de Séparateurs à Vaste Marge, aussi appelés Machines à Vecteurs de Support (Support Vector Machines – SVM). Cette section a pour but un rappel mathématique sur ces classifieurs. Ce rappel mathématique va aussi justifier la raison pour laquelle cette méthode de classification a été préférée aux autres existantes.

7.2.1 Le problème de la classification

7.2.1.1 La classification : un problème ancien

Le problème de la classification apparaît aux yeux de chaque enfant lorsque ses parents essaient de lui apprendre pour la première fois la différence entre le bien et le mal. Cette notion de différenciation entre deux catégories d'objets ou de notions n'est pas nouvelle. Avec le développement de l'informatique ces dernières années, de nombreux algorithmes ont vu le jour pour essayer de donner aux ordinateurs la capacité que nous avons dans notre cerveau à « ranger » une donnée dans la case correspondante.

D'un point de vue mathématique, définissons deux ensembles X et Y . X est l'ensemble des observations et Y l'ensemble des classes. La classification revient à définir une application de X dans Y qui associera pour chaque élément de X une classe de Y . Dans les problèmes de classification, Y est à cardinal fini ($Y = \{1, 2, \dots, N-1, N\}$). Dans le cas contraire ($Y \subset \mathbb{R}$), on parle de modèle de régression. En d'autres termes, un modèle de classification est le fait d'attribuer une étiquette à une donnée.

Ces algorithmes peuvent être utilisés dans de très nombreuses applications. Ainsi, dans le domaine médical, la classification peut servir à différencier de manière automatique, en fonction d'un certain nombre de paramètres physiologiques (température, ...), deux maladies ayant des symptômes proches. Dans des domaines plus éloignés elle peut aussi servir à classer de manière automatique des sons ou des vidéos afin d'en « deviner » le contenu. Enfin, le problème de la classification des sons de la vie quotidienne présente le même type de travaux (avec des algorithmes différents). À partir du calcul des coefficients LFCC d'un son, le système tente de déterminer automatiquement à quelle classe de son le signal enregistré appartient.

7.2.1.2 Classification non supervisée ou supervisée

Les solutions de classification sont de deux types pour correspondre à des problèmes et des besoins différents. Le premier est la classification non supervisée. Dans ce type de problèmes, l'appartenance à chacune des classes des éléments de la base d'apprentissage n'est

pas connue, et le nombre de classes ne l'est pas forcément non plus. On part d'un ensemble d'apprentissage $X_1 \in X$. Dans le cas d'un nombre de classes connu, l'algorithme tente de découper l'ensemble d'apprentissage en ces N classes pour étiqueter ainsi chacun des éléments $x \in X_1$ et réaliser la fonction $f(x) = y$ associant un élément de la base d'apprentissage à sa classe déterminée automatiquement. La méthode des K-moyennes peut par exemple être utilisée pour créer différentes classes dans un ensemble de départ non étiqueté. L'algorithme va ensuite calculer les centroïdes de chacun des ensembles comme paramètre. Une nouvelle donnée sera classifiée en fonction de sa proximité au centroïde créé.

Au contraire, la classification supervisée, utilisée dans notre cas, s'appuie sur un avis « d'expert ». Cet avis d'expert nous donne une base d'apprentissage initiale (x_i, y_i) , $x_i \in X$, $y_i \in Y$, $i \in [1; N]$ qui représente des couples dont le premier élément est la donnée et le second élément la classe à laquelle il appartient. À partir de cet ensemble d'apprentissage, la classification supervisée cherche à construire une fonction de prédiction permettant, en considérant le modèle appris grâce aux données présentées, de classer au mieux une nouvelle donnée.

Cet apprentissage doit être « raisonnable » dans le sens où l'algorithme doit apprendre au mieux à séparer les données, sans apprendre par cœur un comportement. Ce problème, qu'il soit dû à un sous-apprentissage comme à un sur-apprentissage va engendrer ensuite de mauvais résultats de généralisation, lorsque de nouvelles données (inconnues) seront présentées à l'algorithme en vue de classification. Ceci est très bien décrit dans [Burges, 1998] par une analogie avec la botanique, disant qu'un algorithme ayant appris par cœur les données (et donc avec une bonne « capacité¹ ») sera comme un botaniste avec une mémoire photographique excellente à qui l'on présenterait un nouvel arbre et qui arriverait à la conclusion que ce n'est pas un arbre du fait qu'il a un nombre de feuilles différent de tous ceux qu'il a déjà rencontré alors qu'un algorithme avec une trop faible « capacité » serait comme le frère feignant du botaniste, qui concluerait que c'est un arbre parce qu'il est vert. Aucun des deux ne pourrait généraliser correctement le problème. Tout le problème des apprentissages supervisés est donc de trouver un bon compromis entre classification correcte de l'ensemble des points d'apprentissage et classification correcte de nouvelles données.

7.2.2 Les séparateurs à vaste marge pour la classification binaire

7.2.2.1 Classification avec séparation linéaire

Nous nous intéressons ici tout d'abord au cas le plus simple d'un classifieur recevant des données d'apprentissage linéairement séparables. La figure 7.2 montre un tel exemple (dans $\mathbb{R}^2$ pour des simplicités de représentation), avec une séparation voulue entre les points blancs et les points gris. Ce cas est relativement simple. Dans $\mathbb{R}^2$, on recherche un hyperplan de séparation et donc ici une droite (espace de dimension 1). En considérant les données initiales, il existe une infinité de ces hyperplans (quatre d'entre eux sont d'ailleurs représentés

¹La notion de capacité ainsi que la notion de dimension VC sont détaillées par V. Vapnik dans [Vapnik et Chervonenkis, 1971] dans sa théorie de l'apprentissage statistique que nous ne détaillerons pas dans ce manuscrit

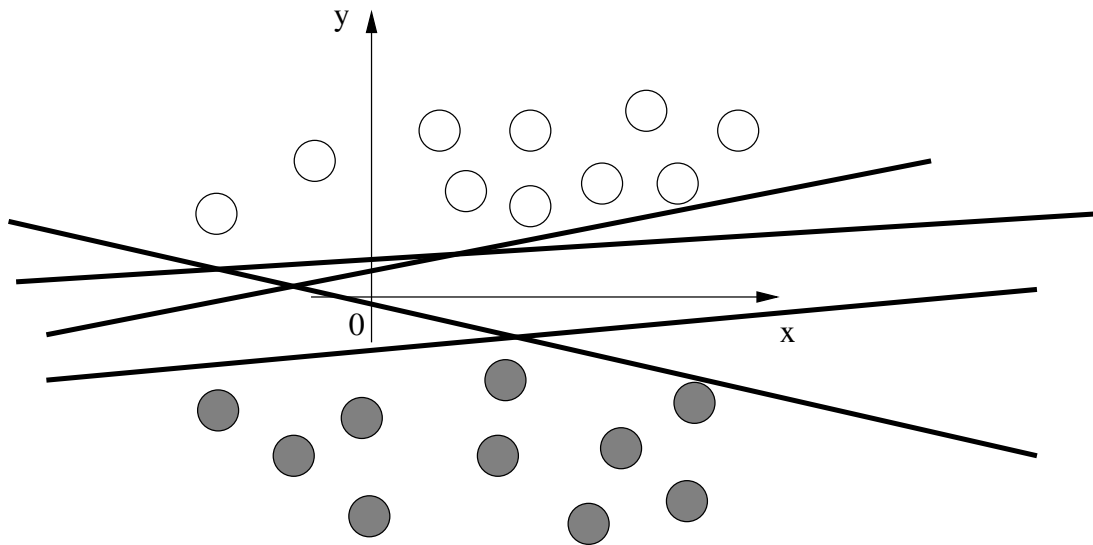


FIG. 7.2: Cas linéairement séparable. Un ensemble de points blancs et gris du plan. Des exemples de séparation linéaires possibles sont également représentées

sur cette même figure).

En considérant ces données comme étiquetées $+1$ pour les blancs et -1 pour les gris, un perceptron à un neurone va, pour trouver cet hyperplan, calculer un vecteur de poids $\mathbf{w}$ tels que :

$$\mathbf{w}^T \cdot \mathbf{x} + w_0 = \begin{cases} \geq 0 \Rightarrow f(x) = +1 \\ < 0 \Rightarrow f(x) = -1 \end{cases} \quad (7.1)$$

$\mathbf{x}$ est le vecteur décrivant l'entrée (dans notre exemple un point du plan, donc deux dimensions). $\mathbf{w}$ est la matrice des poids sur chaque entrée (les coordonnées du plan x et y) et w_0 est une donnée d'initialisation. Par des algorithmes d'optimisation, le perceptron aura donc pour but de trouver un ensemble de poids w_i pour chaque entrée i de telle sorte que les données d'apprentissage soient classées au mieux. Dans le cas de notre problème sur la figure 7.2, nous ne pouvons prédire la sortie du perceptron, qui va dépendre de plusieurs paramètres (w_0 , nombre d'itérations...). La fonction f qui est la sortie du perceptron aura pour valeur $+1$ (et donc classera la donnée dans la classe des points blancs) pour $\mathbf{w}^T \cdot \mathbf{x} + w_0 \geq 0$ et -1 dans le cas contraire. Tester une nouvelle donnée avec ce critère donnera donc sa classification.

Le perceptron est un bon modèle pour de la classification et fonctionne relativement bien sur certains problèmes. Cependant, comment être sûr que l'hyperplan obtenu est optimal pour notre problème? Afin de résoudre ce problème, [Boser *et al.*, 1992] ont mis en commun un ensemble de notions (hyperplan de séparation, marge maximum, noyau, ...) afin de créer les machines à vecteur de support. Contrairement au perceptron, la machine à vecteur de support va créer un unique hyperplan de séparation qui ne va dépendre que des données d'apprentissage. La figure 7.3 montre la construction de cet hyperplan.

Tout d'abord cette figure nous montre la séparation retenue entre les deux espaces, qui est en fait ici l'axe des abscisses, mais aussi deux droites passant par respectivement deux et trois points de la base d'apprentissage. La droite retenue en séparation est construite de ma-

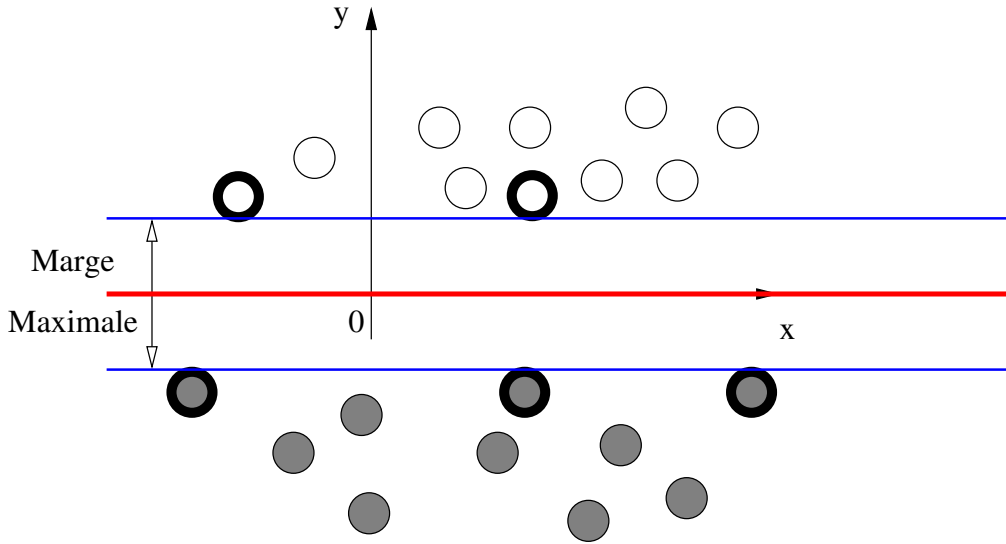


FIG. 7.3: Cas linéairement séparable résolu avec les SVM

nière à maximiser la distance entre elle-même et les points de chaque ensemble d'apprentissage les plus proches. Ces points d'apprentissage les plus proches (mis en évidence sur la figure) sont appelés les vecteurs de support. La marge maximale est représentée par une flèche. En résumé, cette construction va éloigner le plus possible l'hyperplan de séparation des deux classes et créer ainsi une séparation linéaire qui va classer au mieux les nouvelles données pour un ensemble d'apprentissage décrivant bien l'espace de départ.

L'équation 7.1 nous montrait que le perceptron construisait sa fonction de sortie f à partir de la fonction signe sur le produit scalaire des poids et des entrées (à un facteur w_0 près). Dans le cas des SVM, on va avoir $f(x) = \mathbf{w}^T \cdot \mathbf{x} + w_0$ et non plus $f(x) = \text{signe}(\mathbf{w}^T \cdot \mathbf{x} + w_0)$. L'hyperplan séparateur aura pour équation $\mathbf{w}^T \cdot \mathbf{x} + w_0 = 0$. On se donne une fonction u_i tel que :

$$\begin{cases} \mathbf{w}^T \mathbf{x}_i + w_0 > 0 \Rightarrow u_i = 1 \\ \mathbf{w}^T \mathbf{x}_i + w_0 < 0 \Rightarrow u_i = -1 \end{cases} \quad (7.2)$$

L'hyperplan séparateur sera trouvé par la condition suivante (pour un espace de départ $X \subset \mathbb{R}^d$ et n points dans la base d'apprentissage) :

$$\arg \max_{w, w_0} \min_{i=1..n} \left\{ \|x - x_i\| : x \in \mathbb{R}^d, \mathbf{w}^T \cdot \mathbf{x} + w_0 = 0 \right\} \quad (7.3)$$

La marge entre l'hyperplan et le point le plus proche vaut $\frac{1}{\|\mathbf{w}\|}$. Maximiser l'inverse de la norme revient à minimiser celle-ci (que l'on peut élever au carré pour éviter la racine), d'où la formulation primale de ce problème :

$$\begin{cases} \text{Minimisation de la quantité } \frac{1}{2} \|\mathbf{w}\|^2 \\ \text{Contrainte : } u_i(\mathbf{w}^T \mathbf{x}_i + w_0) > 1, i = 1..n \end{cases} \quad (7.4)$$

La résolution de ce problème d'optimisation passe par les multiplicateurs de Lagrange mais implique le réglage de $d + 1$ paramètres dans l'espace de dimension d . La formulation

duale est le plus souvent utilisée pour résoudre ce problème, étant donné que le nombre de ses paramètres ne dépend lui que du nombre d'échantillons de l'espace d'apprentissage et se trouve bien inférieur à d dans de nombreux problèmes en grandes dimensions. L'hyperplan séparateur en considérant la forme duale est alors donné par :

$$f(x) = \sum_{i=1}^n \alpha_i^* u_i(\mathbf{x} \cdot \mathbf{x}_i) + w_0 \quad (7.5)$$

dans laquelle les coefficients α_i^* sont les multiplicateurs de Lagrange optimaux.

En 1995, [Cortes et Vapnik, 1995] ont amélioré les SVM pour y inclure une notion de variable ressort, utile dans les cas de mauvais étiquetage de certaines données par exemple. La contrainte donnée par 7.4 est alors réécrite $u_i(\mathbf{w}^T \mathbf{x}_i + w_0) > 1 - \xi_i, \xi_i > 0$. Il faut alors minimiser la quantité :

$$\frac{1}{2} \|\mathbf{w}\|^2 - C \sum_{i=1}^n \xi_i, C > 0 \quad (7.6)$$

Avec C un paramètre donné par l'utilisateur et qui donne une indication à l'algorithme d'optimisation sur un compromis entre la largeur de la marge et le nombre d'erreurs de classification tolérées sur l'ensemble d'apprentissage.

Vladimir Vapnik a montré que l'augmentation de la marge améliorait le classifieur et sa capacité de généralisation. Ce classifieur a donc la marge maximale grâce aux vecteurs de support choisis. Selon lui, sa méthode permet de régler directement le compromis entre l'adéquation aux données et la limitation de l'espace des hypothèses (qui est donnée par l'ensemble d'apprentissage fini). Dans de nombreux exemples résolus par les SVM et dans ceux cités par [Burges, 1998] entre autre, il s'avère que cette méthode donne toujours des résultats au moins aussi optimaux que d'autres méthodes, voir meilleurs. Ceci explique la forte utilisation de cette méthode ces dernières années. Cependant, le fait qu'il soit également adapté à la résolution de problèmes non-linéairement séparables explique également leur succès. Nous verrons ce point dans la section suivante.

7.2.2.2 Séparateur non-linéaire : Kernel Trick

La section précédente nous montre la manière dont les SVM résolvent le problème de la classification binaire dans le cas de classes linéairement ou presque linéairement (avec les variables ressort) séparables. L'ensemble de ces problèmes ne représente cependant qu'une infime partie des problèmes de classification. Supposons maintenant que notre but est de faire apprendre à notre classifieur à trouver la séparatrice d'équation $x^2 + y^2 = 1$ (cercle unité) à partir de points tirés au hasard tels que ceux de la figure 7.4.

Sans modification de l'algorithme, ce cas ne donnerait pas de solution avec une machine sans variable ressort et de très mauvaises performances avec des ressorts. Autant la séparation linéaire n'est pas possible en restant dans $\mathbb{R}^2$, autant il est probable qu'il soit possible d'en trouver une dans un espace de dimension supérieure, voir infinie. Pour ceci, il suffit de passer dans un espace dit de redescription, noté Z et d'appliquer une transformation non linéaire aux points de l'espace de départ pour passer de X à Z , notée ϕ , dont l'image est

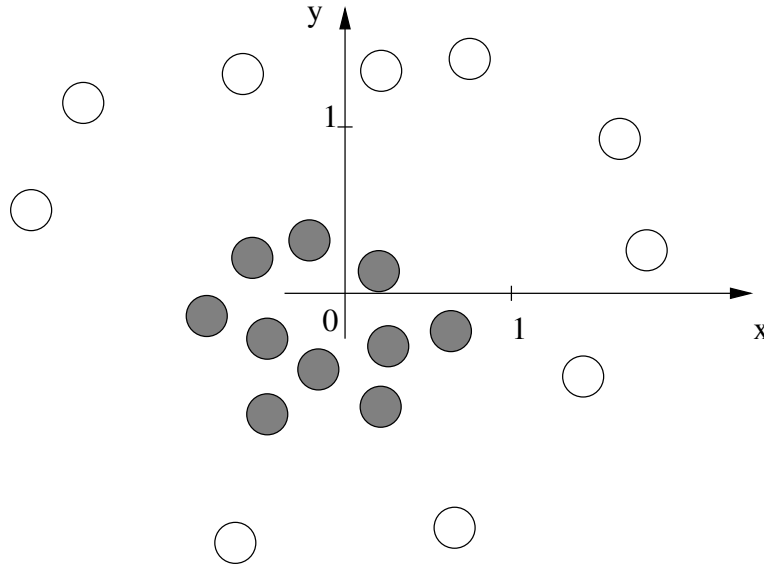


FIG. 7.4: Cas non linéairement séparable de classification binaire

inclue dans Z :

$$\begin{aligned}\phi: X \subset \mathbb{R}^d &\rightarrow Z \subset \mathbb{R}^m \\ \mathbf{x} = (x_1, \dots, x_d)^T &\mapsto \phi(\mathbf{x}) = (\phi_1(\mathbf{x}), \dots, \phi_m(\mathbf{x}))^T\end{aligned}$$

À noter que l'espace de redescription est de dimension $m \geq d$. On peut avoir un espace de redescription de même dimension que l'espace de départ mais dans lequel la séparatrice sera « linéaire » (comme c'est le cas pour notre cercle).

L'ensemble de la formulation des SVM repose sur le produit scalaire. En incluant cet espace de redescription, l'équation de la fonction f devient alors :

$$f(\mathbf{x}) = \mathbf{w}^T \cdot \phi(\mathbf{x}) + w_0 \quad (7.7)$$

Et l'on cherche l'équation de l'hyperplan qui satisfait

$$u_i(\mathbf{w}^T \phi(\mathbf{x}_i) + w_0) > 0 \quad (7.8)$$

Pour tout point x_i de notre ensemble d'apprentissage. Cet hyperplan va séparer linéairement notre espace de redescription.

L'équation 7.5 ne fait intervenir qu'un produit scalaire pour l'hyperplan de séparation. Dans cet espace de redescription, nous pourrions réécrire :

$$f(x) = \sum_{i=1}^n \alpha_i^* u_i(\phi(\mathbf{x}) \cdot \phi(\mathbf{x}_i)) + w_0 \quad (7.9)$$

Qui impose par contre deux contraintes :

- Le calcul, dans l'espace de redescription, d'un produit scalaire. Dans un espace de redescription grand ou infini, cela est coûteux en calcul, voire impossible à réaliser.
- La connaissance obligatoire de la fonction ϕ , fonction non linéaire qui transformera notre espace de départ en espace linéairement séparable. Celle-ci est souvent difficile

à expliciter.

Ces contraintes sont dépassées par l'utilisation d'un noyau, technique connue sous le nom de Kernel Trick et décrite par [Aizerman *et al.*, 1964]. L'idée de base de cette méthode est d'utiliser une fonction noyau K qui serait équivalente, dans un espace à grande dimension, à un produit scalaire. Nous aurions donc :

$$K(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi(\mathbf{x}_i)^T, \phi(\mathbf{x}_j) \rangle \quad (7.10)$$

Pour que la fonction noyau corresponde à ce produit scalaire dans un espace à grande dimension, elle doit respecter le théorème de Mercer, qui explique qu'une telle fonction est continue, symétrique et définie non-négative [Mercer, 1909].

Les noyaux utilisés en général dans les SVM sont les suivants :

- le noyau linéaire, qui revient à la classification initiale sans passage par un espace de redescription :

$$K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \cdot \mathbf{x}_j \quad (7.11)$$

- le noyau polynômial :

$$K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^T \cdot \mathbf{x}_j + 1)^p \quad (7.12)$$

- le noyau gaussien (ou noyau à base radiale) :

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right) \quad (7.13)$$

- le noyau à base de fonction sigmoïde :

$$K(\mathbf{x}_i, \mathbf{x}_j) = \tanh(a(\mathbf{x}_i \cdot \mathbf{x}_j - b)) \quad (7.14)$$

Il est également possible de construire sa propre fonction noyau qui respecterait les conditions données par le théorème de Mercer et serait plus adaptée à un problème donné.

L'ensemble de ces algorithmes, qu'ils concernent les résolutions linéaires ou non linéaires de problèmes de classification, ont été présentés pour des problèmes à deux classes. Notre problème de classification des activités n'entre pas dans cette catégorie de problèmes. Nous allons donc décrire, dans la suite, comment s'étend cette théorie afin de résoudre des problèmes non binaires.

7.2.3 Classification à N classes ($N > 2$)

L'ensemble des méthodes décrites précédemment donnent une classification binaire, c'est à dire que nous restreignons l'espace d'arrivée Y à $\{-1; 1\}$. Cependant, notre problème, comme de nombreux problèmes de classification, sont des problèmes à N classes, ce qui donnera un espace d'arrivée fini $Y = \{1, \dots, N\}$, $1 < N < +\infty$ (le cas $N = 2$ étant le cas précédent, avec une classification binaire, on rappelle que le cas N non fini est un modèle de régression). Pour surpasser cette limitation, deux techniques classiques existent et sont dé-

crites dans les sections suivantes. Ces techniques se basent sur la construction de plusieurs classifieurs binaires.

7.2.3.1 Méthode Un-Contre-Tous

La première méthode pour faire des SVM un algorithme de classification multiclass (et valable également pour d'autres algorithmes de classification binaires), est nommée Un-Contre-Tous (de l'anglais One-Versus-All ou One-Versus-The-Rest selon les articles). Cette méthode est montrée à la figure 7.5. Cette figure montre un exemple à trois classes, pour la relative simplicité de la représentation qui en découle.

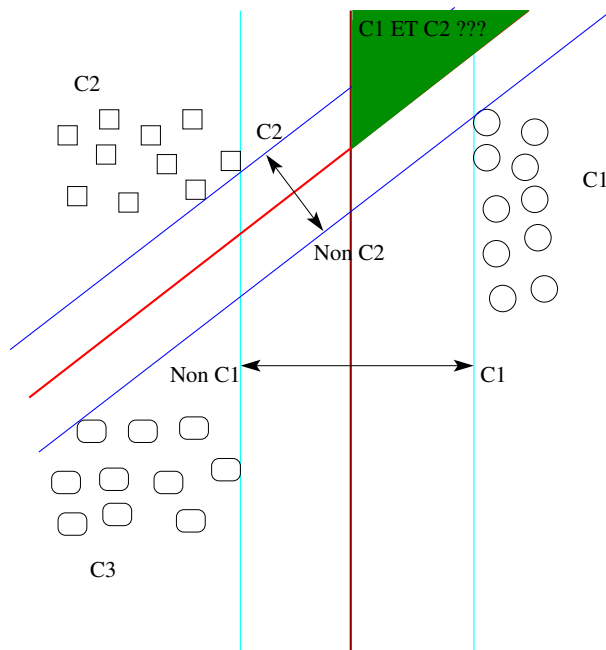


FIG. 7.5: Méthode Un-Contre-Tous pour les SVM multiclass

Cette méthode, pour une classification à N classes, se propose de construire $N-1$ classifieurs binaires, chacun des $N-1$ classifieurs noté B_i confrontant la classe C_i avec les $N-1$ autres classes. Ce classifieur aura donc deux réponses différentes d'un côté et de l'autre de l'hyperplan de séparation : C_i ou $\text{Non-}C_i$, créant ainsi un pavage du plan. La dernière classe (C_3 sur notre représentation) est déterminée par la réponse négative de chacun des classifieurs (donc ici Non-C1 et Non-C2). La zone verte donnée sur la figure représente une zone de réponse ambiguë, dans laquelle chacun des deux classifieurs de cet exemple répond que le point appartient à sa classe de détection. Ce type d'ambiguïté est résolu en considérant comme bonne la réponse du classifieur ayant la plus forte marge (qui est le paramètre exprimant notre « confiance » en la décision prise), ici la classe C_1 .

7.2.3.2 Méthode Un-Contre-Un

La méthode Un-Contre-Un (de l'anglais One-Versus-One) utilise beaucoup plus de classifieurs pour diminuer la zone d'ambiguïté et améliorer les performances sur la décision prise présentée par la méthode Un-Contre-Tous. En considérant un ensemble de classes

$C_i, i \in [1; N]$, pour chacune des classes C_i , on construit $N - 1$ classifieurs permettant de différencier la classe C_i d'une classe C_j ($j \neq i, j \leq N$). Le classifieur opposant C_i à C_j étant identique à celui opposant C_j à C_i , on construit donc $\frac{N \cdot (N-1)}{2}$ classifieurs binaires. La figure 7.6 montre la réalisation de ce type de classification sur un exemple à trois classes dans le plan.

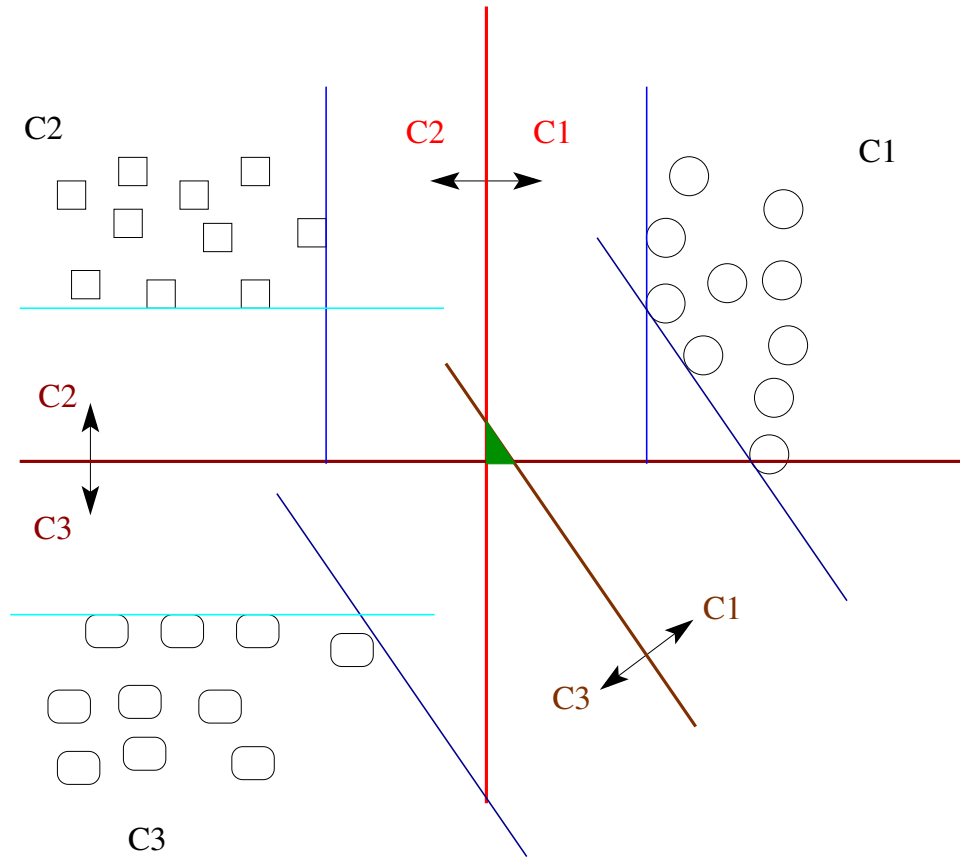


FIG. 7.6: Méthode Un-Contre-Un pour les SVM multiclass

Dans cet algorithme, la donnée $\mathbf{x}$ est testée avec chacun des classifieurs donnant ainsi un ensemble de réponses $\{y_{i,j}\}$, $i, j \in [1; N], j < i$. La décision est prise par vote majoritaire sur cet ensemble de réponses afin de diminuer au mieux les ambiguïtés. Formellement ceci peut s'écrire :

$$C = \max_{k=1..N} \text{Card}(\{y_{i,j}\} \cap \{k\}) \quad (7.15)$$

Ce second algorithme réduit la zone d'ambiguïté. Il reste par contre une petite zone (montrée en vert sur la figure 7.6) dans laquelle les trois classifieurs ont donné une réponse différente. On se retrouve donc avec une égalité pour C . La décision se prend encore une fois sur la marge maximale entre les classes pour lesquelles il existe une ambiguïté. Cependant, même si une zone reste ambiguë dans ce second algorithme, elle est réduite au minimum contrairement à l'algorithme précédent. Ceci se fait par contre au détriment des performances car l'entraînement, puis la phase de test, se fait ensuite sur un plus grand nombre de classifieurs. Nous opérons dans le cadre du HIS pour cette seconde méthode.

7.3 Apprentissage et tests dans le HIS

Afin de valider l'utilisation des SVM dans notre cas bien précis de la classification des activités de la vie quotidienne mais également d'obtenir des résultats expérimentaux sur ce type de données, nous avons testé la classification de ces données sous MatlabTM. L'expérimentation que nous avons réalisée sera décrite plus précisément dans le chapitre 8. Dans cette section nous allons développer la manière dont s'est fait l'apprentissage sur nos données ainsi que la méthode de validation de l'algorithme, dont les résultats seront présentés dans le chapitre 9.

7.3.1 Indexation des données

Dans tout algorithme de classification, la première étape est la phase d'apprentissage. Pour réaliser celle-ci, nous avons besoin de l'avis d'un expert qui étiquette les données expérimentales afin de connaître l'activité en cours pour une fenêtre temporelle donnée. Le HIS possède un ensemble de caméras nous permettant de faire cette indexation manuellement.

La classification des activités va se faire selon sept classes :

- Classe 1 : activité repos
- Classe 2 : activité détente
- Classe 3 : activité habillage/déshabillage
- Classe 4 : activité cuisine (préparer et prendre un repas)
- Classe 5 : activité élimination (aller aux toilettes)
- Classe 6 : activité hygiène (se laver les mains, se brosser les dents, ...)
- Classe 7 : activité communication avec l'extérieur.

Afin de permettre un apprentissage, il est nécessaire d'indexer les données de telle manière que l'algorithme en charge de la construction et du formatage des données, que nous avons présenté dans le chapitre précédent, puisse connaître la classe à laquelle appartient une fenêtre temporelle. Pour ceci, nous avons construit, pour chacune des sessions expérimentales, un fichier XML décrivant l'exécution du scénario par la personne. Ces fichiers sont chargés par l'application Matlab qui effectue ensuite la construction des fenêtres temporelles pour chacun des capteurs puis la mise en forme les données telle que décrite dans le chapitre précédent.

Dans ce fichier, chacune des activités est décrite avec son heure de début et son heure de fin. Cet ensemble de descriptions de chacune des activités est précédé par un ensemble d'informations sur les noms des fichiers pour chacun des types de capteurs. De plus, pour ce qui est de la carte ACTIM6D, il faut que le système soit capable de dater les événements qui sont présentés par le fichier de sortie sous la forme d'un numéro d'échantillon. Un champ est donc, pour cette utilisation, réservé pour donner la synchronisation de la carte ACTIM6D et de ses données avec l'horodatage de l'ensemble des autres capteurs. Ainsi, l'ensemble des données possède une cohérence temporelle du fait de la synchronisation de l'heure de chacune des machines avec un serveur de temps unique.

Les fichiers XML sont écrits à la main lors de l'indexation des vidéos. Les fichiers générés sont de la forme suivante :

```
<?xml version="1.0" encoding="UTF-8"?>
<Expe id="1">
    <dossier>ManipFusion\01</dossier>
    <ACTIM6D>ACTIM6D.out</ACTIM6D>
    <Frigo>Frigo.txt</Frigo>
    <Commode>Commode.txt</Commode>
    <Placard>Placard.txt</Placard>
    <meteo>Temperature.temp</meteo>
    <son>son.txt</son>
    <IR>HIS.his</IR>

    <ACTIM6D_synchro>
        <pos> 0 </pos>
        <date> 22-07-2008-14:12:01 </date>
    </ACTIM6D_synchro>
    <Activite id="2">
        <debut>"22-07-2008-14:12:15"</debut>
        <fin>"22-07-2008-14:17:55"</fin>
    </Activite>
    ...
    <Activite id="1">
        <debut>"22-07-2008-14:22:20"</debut>
        <fin>"22-07-2008-14:38:40"</fin>
    </Activite>
    ...
</Expe>
```

7.3.2 Validation croisée

Les résultats présentés par la suite sont issus de validation croisée. Notre ensemble d'apprentissage n'est pas excessivement grand (13 sujets sont passés sur l'expérimentation), nous allons donc utiliser la méthode « leave-one-out ». Cette méthode est adaptée aux cas comme le nôtre d'un ensemble d'apprentissage restreint, pas assez grand pour faire une base de test également. Ainsi, avec cette méthode, nous retirons un élément de notre ensemble de données étiquetées, nous apprenons sur l'ensemble restant puis nous testons sur la donnée retirée. Nous bouclons afin de retirer chacune des données une fois. Cela nous donne une première statistique sur les performances de notre algorithme. La validation croisée peut également se faire sur des répartitions différentes du corpus. Pour de plus gros corpus, il est possible d'utiliser 90% du corpus en apprentissage et 10% en test. Ce type de tests est plus

adapté au corpus importants.

Le chapitre suivant va revenir plus en détail sur l'expérimentation qui a été menée ainsi que les conditions dans lesquelles elle s'est déroulée.

Troisième partie

**... Aux expérimentations en conditions
réelles**

Déroulement de l'expérimentation

Sommaire

8.1 But de l'expérimentation	159
8.2 Matériel, équipement	160
8.3 Protocole expérimental	160
8.3.1 Contraintes imposées au sujet	160
8.3.2 Activités à réaliser	161
8.3.2.1 Se reposer	161
8.3.2.2 Se préparer et prendre un petit-déjeuner	161
8.3.2.3 S'habiller et se déshabiller	161
8.3.2.4 Se détendre	162
8.3.2.5 Activité d'hygiène	162
8.3.2.6 Élimination	162
8.3.2.7 Communication avec l'extérieur	162
8.4 Conclusion	163

LA théorie ayant été introduite dans les chapitres précédents, avec la présentation des différents capteurs, l'interprétation des différentes données et une introduction sur la classification des données par l'intermédiaire des SVM, nous allons maintenant nous concentrer sur la présentation des expérimentations effectuées dans le HIS de Grenoble pour ce chapitre suivie d'une présentation des résultats obtenus dans le chapitre suivant, avant de discuter plus longuement de l'ensemble des résultats présentés dans ce manuscrit.

8.1 But de l'expérimentation

Au cours de ce manuscrit, nous nous sommes efforcés de présenter un travail permettant la classification des fenêtres temporelles lors des activités de la vie quotidienne. Cette classification doit se faire le plus automatiquement possible. Notre but est également que celle-ci dépende de peu de paramètres liés, soit aux conditions expérimentales, soit au sujet. Pour ceci, nous avons investigué les techniques de classification supervisées, dans lesquelles un algorithme apprend un comportement sur un certain nombre de fenêtres pour être capable de le généraliser ensuite sur de nouvelles données.

La capacité de généralisation de l'algorithme va dépendre de certaines inconnues préalables aux expérimentations :

- La bonne séparabilité (linéaire ou non) des différentes classes données. L'absence de cette condition est le premier frein à toute généralisation. En effet, un ensemble de données difficilement séparable sera plus à-même de donner de mauvais résultats lors de tests sur des nouvelles données, du fait de l'imperfection de la séparation calculée.
- La représentativité des données d'apprentissage. Les données d'apprentissage suivent une loi associant à tout élément d'un ensemble X une classe de l'ensemble $Y = \{1..N\}$. Idéalement, le tirage des échantillons doit être indépendant et identiquement distribué. En obtenant une quantité suffisante de données, nous devrions statistiquement être dans ce cas.

Nous avons donc établi un protocole expérimental représentatif de l'ensemble des activités que nous voulions détecter automatiquement, ce protocole nous a permis de collecter des données qui sont nécessaires pour apprendre des comportements et ensuite valider les techniques de classification supervisées. Ce protocole expérimental, mis en œuvre avec un ensemble de sujets jeunes, en bonne santé et volontaires, nous fournira un ensemble de données diversifiées (du fait du nombre de sujets) nous permettant de tester l'apprentissage automatique et sa capacité de généralisation, à l'aide d'une validation croisée. Le but de cette expérimentation est donc de construire l'ensemble d'apprentissage nécessaire à notre algorithme de classification. Ces expérimentations devaient, par construction, se rapprocher au plus de la « vie réelle » des sujets, afin de travailler sur un type de données différent de celles d'un simulateur ou d'un petit scénario réalisé sous de nombreuses contraintes.

8.2 Matériel, équipement

Afin de se dérouler dans des conditions les plus proches possibles de la réalité, ces expérimentations se sont tenues dans le HIS de la faculté de médecine de Grenoble, présenté au chapitre 6. Le HIS était, pour l'occasion, totalement équipé, tous les capteurs présents étant en fonctionnement afin de pouvoir enregistrer l'ensemble des données. Nous avons également tenté de le rendre le plus « vivable » possible, afin qu'il soit proche d'un appartement classique dans lequel la personne aurait pu venir vivre.

Pour ceci, l'ensemble des commodités étaient présentes dans le HIS et nous avons également prévu de la nourriture en vue de se préparer un petit-déjeuner, et même des vêtements pour se changer ...

La personne portait également la carte ACTIM6D afin d'enregistrer ses postures et sa marche. Cette carte était intégrée dans un tee-shirt que le sujet devait porter (par dessus le sien ou en enlevant son propre tee-shirt selon les choix). Si le tee-shirt était trop ample et ce afin de diminuer les mauvaises détections, nous avons prévu une ceinture permettant de mieux maintenir le capteur au niveau du torse.

8.3 Protocole expérimental

Le protocole expérimental, pour arriver au but fixé, devait donc comporter la réalisation de l'ensemble des activités de la vie quotidienne par le sujet. Nous allons tout d'abord décrire brièvement les contraintes demandées au sujet, puis décrire le protocole en lui-même.

8.3.1 Contraintes imposées au sujet

Le sujet avait très peu de contraintes sur la réalisation des activités. En effet, il réalisait les activités dans l'ordre qu'il désirait, le nombre de fois voulu et l'expérimentation n'avait pas de durée minimum (selon la disponibilité de la personne et son implication dans l'expérimentation).

L'expérimentation était filmée dans son ensemble afin d'indexer les données et créer l'étiquetage, avec pour avantage de ne créer aucune contrainte et de laisser la personne vivre à sa guise. Les seules contraintes imposées au sujet étaient de :

- Réaliser l'ensemble des activités demandées au moins une fois.
- Effectuer un mouvement permettant la synchronisation des données. Ce mouvement, capté par les accéléromètres et les magnétomètres, devait être effectué par le sujet devant l'une des caméras du HIS dès qu'il y était entré.
- Rentrer dans le HIS et n'en ressortir qu'une fois qu'il considérait l'expérimentation terminée. Pendant ce temps, il ne devait pas avoir de communication avec l'extérieur (sauf dans l'activité correspondante, mais avec le téléphone fourni), et personne d'autre ne rentrait dans le HIS. En cas de problème, il lui était demandé de continuer comme si tout se passait correctement ; l'indexation permettrait ensuite d'éliminer toute autre forme de problème rencontré.

- Les activités d’hygiène et d’élimination ne sont bien sûr pas filmées et ce pour un respect évident des règles de la vie privée. Dans cette activité, la caméra présente dans la salle de bain vise la porte d’entrée de celle-ci et ne voit que cette porte. La personne vivant dans l’appartement devait laisser entrouverte la porte lorsqu’elle se rendait aux toilettes et la fermer complètement lorsqu’elle se rendait au contraire à la salle de bain.

En dehors de ces quelques exigences, la personne était supposée vivre dans le HIS comme chez elle, sans se préoccuper des caméras ni des capteurs environnants.

8.3.2 Activités à réaliser

Nous avons précédemment défini sept activités différentes que nous voulons différencier automatiquement à l’intérieur du HIS, en utilisant l’ensemble des capteurs installés dans l’appartement. L’ensemble des gestes et déplacements à réaliser devait donc couvrir l’ensemble de ces activités. Avant le début de l’expérimentation, le HIS était présenté à la personne de manière qu’elle n’y cherche rien et y prenne ses repères. Cet appartement devait lui sembler familier afin que l’expérimentation soit le plus réaliste possible. Après ceci, il était demandé dans le scénario d’effectuer les activités de la vie quotidienne suivantes :

8.3.2.1 Se reposer

Un lit est disponible dans le HIS. La personne pouvait s’y installer et y rester le temps qu’elle voulait. Le capteur embarqué n’est pas gênant dans ce cas et la personne est prévenue qu’elle ne doit pas avoir peur de « l’écraser » si jamais elle doit se tourner sur le côté. Si la personne réussissait à s’assoupir l’activité n’en était que plus crédible, cependant ce n’était pas une obligation.

8.3.2.2 Se préparer et prendre un petit-déjeuner

Les expérimentations n’avaient pas toujours lieu le matin (ceci était dépendant du sujet principalement). Cependant les différents ingrédients ainsi que le matériel disponible permettait à tout un chacun de se faire le petit-déjeuner qu’il appréciait. Différentes boissons ainsi que le matériel nécessaire pour les préparer était disponible. Le sujet n’avait aucune contrainte sur son alimentation, le temps utilisé pour prendre le repas ou bien la manière de le préparer. Il lui était seulement demandé de s’alimenter, en se servant dans l’ensemble des provisions disponibles, puis de nettoyer les couverts utilisés ainsi que la table pour laisser l’endroit propre (comme nous le ferions dans notre propre appartement).

8.3.2.3 S’habiller et se déshabiller

Un ensemble de vêtements était disposé dans la commode du HIS. Ces vêtements, d’une grande taille, sont supposés pouvoir être portés par toutes les personnes effectuant l’expérimentation et permettait aussi d’être portés par dessus les vêtements actuels. La personne pouvait se servir parmi ces vêtements et s’habiller (en les mettant par-dessus les siens par exemple) puis se déshabiller par la suite.

8.3.2.4 Se détendre

Lors de l'annonce, les personnes étaient au courant du type d'activité qui leur serait demandé. Entre autre, il leur était proposé d'amener de la lecture ou toute autre objet qui pourrait leur servir à effectuer cette activité de « détente ». Cette activité est en fait relativement vague et contient un peu tout ce que l'on peut faire en dehors des autres activités. Ainsi, les personnes pouvaient lire (un article, un livre ou bien l'un des nombreux magazines présent dans le HIS), jouer au sudoku, se mettre la télévision, s'asseoir sur le canapé, ... Cette activité est laissée à l'appréciation de la personne et doit refléter ce qu'elle pourrait faire chez elle, si elle n'avait rien de spécifique à faire.

8.3.2.5 Activité d'hygiène

Cette activité est un peu plus compliquée à réaliser. En effet, même si le HIS possède une douche, il est relativement difficile de demander et de convaincre des personnes volontaires pour une expérimentation de prendre une douche dans un appartement qu'ils ne connaissent pas et dans lequel un ensemble de capteurs est présent. Pour faire cette activité, les personnes avaient à leur disposition du savon à côté du robinet de la salle de bain, et il leur était proposé de feindre de se laver les dents.

8.3.2.6 Élimination

De même, cette activité est en général plus compliquée à demander aux personnes. Lors de la visite du HIS, les toilettes étaient présentées aux personnes comme étant fonctionnels. Dans le cas où ils avaient envie de s'y rendre ils le pouvaient. Cependant, ils pouvaient aussi feindre de réaliser cette activité en prenant la position assise puis en tirant la chasse d'eau par la suite. La personne était libre de réaliser cette activité telle qu'elle le désirait, la seule contrainte étant de ne pas oublier le « code » avec la porte permettant une indexation correcte de l'activité.

8.3.2.7 Communication avec l'extérieur

Enfin, la dernière activité proposée était la communication avec l'extérieur. Dans ce cas, la personne recevait des appels mais n'en émettait pas. Un ensemble de quatre conversations (une longue et trois courtes) était fourni au début de l'expérimentation. Cet ensemble était choisi parmi cinq ensembles différents générés à partir des corpus anodin et détresse et réalisés de sorte à former des conversations téléphoniques classiques. La personne n'était pas en réelle communication avec l'extérieur (nous faisons uniquement sonner le téléphone, ensuite elle énonçait les conversations), cependant il lui était demandé de réaliser cette conversation le plus naturellement possible, en effectuant des pauses comme si elle avait une autre personne à l'autre bout du fil.

8.4 Conclusion

Ce chapitre présente les différentes activités à classer. Ces activités ont donc été réalisées par un ensemble de personnes jeunes et en bonne santé, qui se sont portées volontaires pour cette expérimentation. La durée de l'expérimentation était variable, entre trente minutes pour les plus rapides jusqu'à plus de deux heures pour les plus lents. L'ensemble des données recueillies va ensuite être utilisé afin de classer les activités de la vie quotidienne et obtenir des premiers résultats (à l'aide de la validation croisée) sur les performances de l'algorithme.

Résultats

Sommaire

9.1 Population	167
9.2 Méthode d'obtention des résultats	167
9.3 Résultats expérimentaux	168
9.3.1 Noyau polynômial	168
9.3.2 Noyau gaussien	170
9.3.3 Comparaison des performances	171
9.4 Discussion	172

AFIN d'avoir des résultats intéressants sur l'apprentissage et le test avec un algorithme supervisé, nous avons réalisé une expérimentation sur des sujets jeunes et en bonne santé. Cette expérimentation a permis la collecte de données pour les différentes activités de la vie quotidienne que nous tentons de classer. Ce chapitre présente la méthodologie d'obtention ainsi que les résultats provenant de cette collecte de données puis du traitement par les algorithmes mis en place.

9.1 Population

Cette expérimentation s'est déroulée à la Faculté de Médecine de Grenoble, dans l'habitat HIS spécifiquement aménagé pour l'accueillir et totalement équipé. L'ensemble des capteurs présentés était en place afin de collecter et de formater les données de la manière présentée au chapitre 6.

La population cible de ces expérimentations était des sujets jeunes et en bonne santé, volontaires pour l'expérimentation et acceptant toutes les conditions de celle-ci (y compris d'être filmés dans un but d'indexation). Quinze sujets ont participé à l'expérimentation, cependant pour deux d'entre-eux, au moins l'un des capteurs a eu une défaillance pendant une longue période empêchant la synchronisation globale. Nous avons donc retenu treize sujets au total. Dans l'ensemble, ces personnes étaient âgées en moyenne de $30,4\text{ans} \pm 5,9$ (maximum 43 ans, minimum 24 ans), de taille moyenne $1,76\text{m} \pm 0,08$ (taille maximum : 1,92m, taille minimum 1,62m), pour un poids moyen de $69\text{kg} \pm 7,42$ (pour un maximum de 80 kg et un minimum de 57 kg). La répartition par sexe était de six femmes et sept hommes.

9.2 Méthode d'obtention des résultats

La consigne qui était donnée au sujet réalisant l'expérimentation était de se comporter comme chez eux dans notre appartement et d'y réaliser les sept activités de la vie quotidienne que nous voulions classer. Par la suite, l'ensemble des enregistrements vidéo était visualisé afin d'indexer les activités (trouver l'heure de début et de fin de chacune d'entre elles). À partir de ces indexations, un fichier XML décrivant l'expérimentation a été créé (voir les explications au chapitre 7).

Les fichiers XML de chacun des sujets sont ensuite chargés dans un programme Matlab qui crée de nouvelles données et les formate correctement. Les activités n'avaient pas de durée imposée pour le sujet. Nous avons nous, par contre, choisi de créer des fenêtres temporelles de trois minutes (qui correspondent à l'activité la plus courte), comme indiqué au chapitre 7. Lorsqu'une activité durait plus de trois minutes, elle était alors divisée en fenêtres de trois minutes, chacune d'entre elles étant étiquetée avec la classe de l'activité à laquelle elle appartient. Si l'une des activités durait moins de trois minutes (les activités d'hygiène et de toilettes étant, dans les conditions du laboratoire, plus rapides à réaliser), nous considérons la fenêtre uniquement si elle dépassait une minute afin de disposer tout de même d'un ensemble de données suffisant.

Classe	Nombre de fenêtres
C1 (Repos)	49
C2 (Détente)	75
C3 (Habillage/Déshabillage)	16
C4 (Repas)	45
C5 (Élimination)	16
C6 (Hygiène)	14
C7 (Communication)	17

TAB. 9.1 – Répartition des fenêtres temporelles dans la base d'apprentissage

Dans ces conditions, nous avons élaboré une base d'apprentissage des fenêtres à tester, étiquetées par un avis « expert ». La constitution de cette base d'apprentissage est donnée par le tableau 9.1.

Comme nous le voyons sur ce tableau 9.1, la base d'apprentissage est assez déséquilibrée en nombre de fenêtres par classes. En effet, autant il est aisé pour quelqu'un de dormir pendant vingt minutes, autant faire l'activité élimination pendant vingt minutes est plus compliqué et peu naturel. Nous avons donc des classes beaucoup plus représentées que d'autres dans notre base. Nous pouvons aussi noter que cette base reste de taille relativement modeste. En effet, demander aux sujets volontaires de rester des heures durant dans le HIS est contraignant. C'est pour cela que nous utiliserons la méthode leave-one-out pour tester la capacité de généralisation et les performances de l'algorithme.

9.3 Résultats expérimentaux

Les performances des classifieurs SVM vont aussi dépendre du noyau utilisé. Ce noyau, comme décrit dans le chapitre 7, sert à projeter les données dans un espace de redescription, dans lequel la séparation entre les différentes classes est linéaire. Pour des données dont la dimension est supérieure à trois (et nous sommes dans ce cas), il est difficile de les tracer afin de trouver un espace de redescription approprié, ce sont donc les tests sur notre base d'apprentissage qui vont pouvoir nous donner des indications sur le meilleur noyau à adopter et les paramètres les plus appropriés à notre classification [Chapelle, 2004].

9.3.1 Noyau polynômial

Le noyau polynômial est de la forme :

$$K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^T \cdot \mathbf{x}_j + 1)^d \quad (9.1)$$

Il nous suffit donc de fixer un paramètre, d , le degré du polynôme. Nous avons donc fait varier ce degré du polynôme afin d'observer les différences de performances de l'algorithme. Les résultats pour $d = 1$ et $d = 2$ sont présentés sous forme de matrices de confusion dans

		Résultat de la classification						
		C1	C2	C3	C4	C5	C6	C7
Activité	C1	89,8% (44)	6,1% (3)	4,1% (2)	0% (0)	0% (0)	0% (0)	0% (0)
	C2	17,3% (13)	73,3% (55)	4% (3)	1,4% (1)	0% (0)	0% (0)	4% (3)
	C3	25% (4)	6,25% (1)	62,5% (10)	0% (0)	0% (0)	0% (0)	6,25% (1)
	C4	0% (0)	0% (0)	2,25% (1)	95,5% (43)	0% (0)	2,25% (1)	0% (0)
	C5	6,25% (1)	12,5% (2)	0% (0)	6,25% (1)	56,25% (9)	18,75% (3)	0% (0)
	C6	14,3% (2)	7,1% (1)	0% (0)	14,3% (1)	21,3% (3)	50% (7)	0% (0)
	C7	0% (0)	23,6% (4)	11,7% (2)	0% (0)	0% (0)	0% (0)	64,7% (11)

TAB. 9.2 – Matrice de confusion pour un noyau polynômial avec $d = 1$. Pour l'ensemble des matrices de confusion données, les éléments en gras font ressortir les classifications correctes (diagonale), et les chiffres entre parenthèse représentent le nombre d'éléments classés et qui représentent le pourcentage donné

		Résultat de la classification						
		C1	C2	C3	C4	C5	C6	C7
Activité	C1	89,8% (44)	4,1% (2)	6,1% (3)	0% (0)	0% (0)	0% (0)	0% (0)
	C2	13,3% (10)	77,3% (58)	4% (3)	0% (0)	0% (0)	1,4% (1)	4% (3)
	C3	25% (4)	6,25% (1)	62,5% (10)	0% (0)	0% (0)	0% (0)	6,25% (1)
	C4	2,2% (1)	4,4% (2)	0% (0)	82,3% (37)	2,2% (1)	6,7% (3)	2,2% (1)
	C5	0% (0)	12,5% (2)	0% (0)	6,25% (1)	68,75% (11)	12,5% (2)	0% (0)
	C6	7,1% (1)	14,3% (2)	0% (0)	14,3% (1)	21,3% (3)	50% (7)	0% (0)
	C7	0% (0)	17,6% (3)	5,9% (1)	0% (0)	0% (0)	0% (0)	76,5% (13)

TAB. 9.3 – Matrice de confusion pour un noyau polynômial avec $d = 2$

les tableaux 9.2 et 9.3 respectivement. Nous ne pouvons utiliser de degré supérieur sur nos données, les matrices d'optimisation étant ensuite singulières (aucune solution).

Comme nous pouvons l'observer, les différences ne sont pas significatives, et le taux de bonne classification global est quasiment identique dans les deux cas. Pour le cas $d = 1$, nous arrivons à un taux global de bonne classification de 71,31% contre un taux de 71,71% dans le second cas.

Le tableau 9.2 est riche en informations. Les éléments en gras dans ces matrices nous montrent les taux de bonne classification pour chacune de nos classes. Les classes donnant les meilleurs résultats sont les classes repos, détente et repas, qui sont, comme nous l'avons énoncé précédemment, les mieux représentées dans notre base d'apprentissage. Nous pouvons également observer des taux de mauvaise classification élevés dans plusieurs cas :

- 25% des éléments de la base d'apprentissage de la classe Habillage sont classés en Repos. Ce résultat est le fruit du fait que, dans la chambre (dans laquelle se déroule également l'activité habillage), le repos possède beaucoup plus d'exemples d'apprentissage alors que la classe habillage est relativement restreinte. Dans de nombreux cas, des vecteurs de support seront donc mis en test, ce qui aura pour conséquence de modifier le classifieur et le rendre moins optimal pour détecter la classe habillage.
- Dans 19% des cas, une activité élimination et classée hygiène et dans 21% l'activité

hygiène est classée élimination. Ce résultat s'explique également par la faiblesse de la taille des deux bases d'apprentissage qui s'ajoute à l'unité des lieux, dans notre cas, pour ces deux activités.

- 23% des fenêtres temporelles correspondant à la classe communication avec l'extérieur sont classées en détente. Ceci s'explique également par l'unité des lieux et par le fait que, si l'élément discriminant (la sonnerie de téléphone puis la parole) est manqué ou mal classifié par le système de reconnaissance de son et de parole, le classifieur pourra commettre une erreur « légitime » de classification.

Pour ce qui est de la matrice 9.3, avec un noyau polynômial de degré 2, nous pouvons noter des évolutions par rapport au tableau précédent, même si le taux de bonne classification global n'est pas significativement modifié. Tout d'abord nous avons une baisse significative des performances dans le cas de la classe Repas mais une augmentation dans le cas de la classe communication. La forme et le degré du noyau vont être plus ou moins adaptés à réaliser une séparatrice linéaire pour nos classes. Le taux de mauvaise classification de la classe Habillage en détente restent dans ce cas inchangé. Par contre, la différenciation des classes détente et communication et des classes hygiène et élimination sont fortement réduites. Ceci, ajouté au fait que les mauvaises classification de la classe Repas se retrouvent réparties dans l'ensemble des autres classes peut nous amener à déduire que ce noyau, avec un degré 2, donne moins d'importance à la modalité « présence dans la pièce » et plus de poids aux autres modalités (son et actimétrie principalement).

9.3.2 Noyau gaussien

Le noyau gaussien est, quant à lui, défini par l'équation :

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right) \quad (9.2)$$

Il dépend d'un paramètre σ qui représente l'écart-type de la gaussienne calculée. Nous avons fait varier cet écart-type afin d'obtenir une courbe permettant de tracer le taux d'erreur global en fonction du paramètre σ . Cette courbe est présentée figure 9.1.

Pour la valeur de σ nous donnant les résultats les plus performants, qui est ici $\sigma = 28$, la matrice de confusion est donnée par le tableau 9.4. Dans ce cas, le taux de bonne classification global s'élève à 78,8%. En comparaison avec les matrices de confusion du noyau polynômial, nous pouvons remarquer plusieurs points :

- 19% des éléments de la classe Habillage sont encore classés en Détente, et ce pour les mêmes raisons que précédemment.
- 21% des éléments de la classe Hygiène sont toujours classés en Élimination, alors que dans le cas contraire, le taux de mauvaise classification d'éléments de la classe Élimination en hygiène a été réduit à 12,5%
- La différenciation entre les classes communication et détente est beaucoup mieux réalisée dans ce cas.

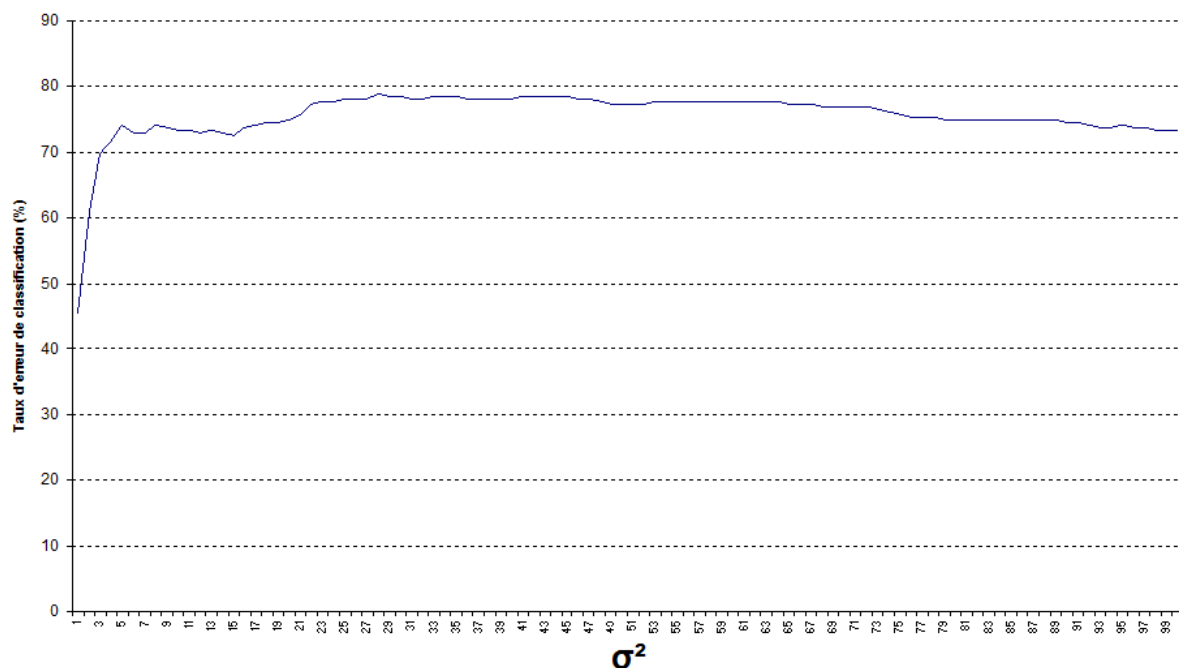


FIG. 9.1: Taux d'erreur global de l'algorithme de classification en fonction du paramètre σ du noyau

		Résultat de la classification						
		C1	C2	C3	C4	C5	C6	C7
Activité	C1	93,9% (46)	4,1% (2)	2% (1)	0% (0)	0% (0)	0% (0)	0% (0)
	C2	12% (9)	84% (63)	0% (0)	1,3% (1)	0% (0)	0% (0)	2,7% (2)
	C3	18,75% (3)	6,25% (2)	68,75% (11)	0% (0)	0% (0)	0% (0)	0% (0)
	C4	0% (0)	2,2% (1)	0% (0)	95,6% (43)	0% (0)	0% (0)	2,2% (1)
	C5	0% (0)	6,25% (1)	0% (0)	6,25% (1)	75% (12)	12,5% (2)	0% (0)
	C6	7,1% (1)	7,1% (1)	0% (0)	7,1% (1)	21,5% (3)	57,2% (8)	0% (0)
	C7	0% (0)	5,9% (1)	5,9% (1)	0% (0)	0% (0)	0% (0)	88,2% (15)

TAB. 9.4 – Matrice de confusion pour un noyau gaussien avec $\sigma^2 = 28$

Nous voyons que le noyau Gaussien donne de meilleurs résultats sur l'ensemble des classes de notre base d'apprentissage, y compris les classes de petite taille (dans la base d'apprentissage).

9.3.3 Comparaison des performances

Nous avons, dans les deux sections précédentes, deux modèles différents qui donnent des taux de bonne classification respectifs de 71,7% et 78,8%. Ces taux de bonne classification ont été obtenus en sélectionnant la meilleure valeur du paramètre, sur les 251 tests effectués. Le tableau 9.5 compare les taux de bonne classification entre chacune des classes pour les deux paramètres.

Comme ce tableau le montre, les classes bien reconnues par le noyau Gaussien le sont aussi par le noyau polynômial. La meilleure classification est obtenue cependant, dans notre

Classe	Noyau Polynômial	Noyau Gaussien
C1	89%	93,88%
C2	77,3%	84%
C3	62,5%	68,75%
C4	82,2%	95,56%
C5	68,75%	75%
C6	50%	57,14%
C7	76,47%	88,24%
Global	71,71%	78,88%

TAB. 9.5 – Comparaison des taux de bonne classification pour chacune des classes avec un noyau polynômial et un noyau gaussien

cas, avec un noyau Gaussien.

9.4 Discussion

Les résultats présentés précédemment montrent un bon taux de classification des activités de la vie quotidienne avec les SVM. Ces taux de bonne classification, du fait du faible nombre d'exemples dans notre base d'apprentissage, ont été obtenus par la méthode du leave-one-out.

Ces taux de bonne classification font cependant ressortir des résultats disparates entre les classes. Ces résultats s'expliquent par la taille de certaines des classes dans la base d'apprentissage. En effet, autant les classes Repos et Détente sont bien représentées, autant les autres classes contiennent un faible nombre d'éléments. Lors de la construction des marges et des séparatrices, les SVM se préoccupent principalement d'un sous-ensemble des éléments de la base d'apprentissage appelés exemples critiques. Ces exemples, qui deviennent les vecteurs de support, sont importants pour la construction de la séparatrice. Retirer un élément, qui n'appartient pas à ces vecteurs, de la base d'apprentissage, n'aura aucune influence sur la construction de la machine. Le cas contraire modifiera par contre la machine et les vecteurs de support, et donc la séparatrice. Sur les classes avec un nombre faible de vecteurs d'apprentissage, la méthode du leave-one-out va donc engendrer de nombreuses modifications des séparatrices d'une construction à l'autre.

Par contre, la méthode du « leave-one-out » va nous permettre d'effectuer un nombre de tests égal au nombre de vecteurs de notre base d'apprentissage. En ce sens, elle aura l'avantage de nous donner des résultats significativement plus intéressants sur des petites bases.

Ces résultats présentés, la prochaine partie va discuter plus généralement de l'ensemble des travaux de cette thèse avant de conclure et d'ouvrir sur les suites possibles.

Quatrième partie

Discussion et Conclusion

Discussion Générale

Sommaire

10.1 Apports de ces travaux	177
10.1.1 ACTIM6D	177
10.1.2 Système de reconnaissance de sons et de paroles	178
10.1.3 Classification des activités	179
10.1.3.1 Base d'apprentissage	179
10.1.3.2 Ajout d'information <i>a priori</i>	180
10.1.3.3 Le temps comme modalité	180
10.1.3.4 Classe supplémentaire nécessaire : une classe de rejet	181
10.2 Validité des résultats	181
10.2.1 « Approche de la réalité par l'expérimentation » versus « Réalité »	181
10.2.2 Généralisation des modèles à la personne âgée ?	182

APRÈS avoir introduit la problématique soulevée par ces travaux de thèse (chapitres 1, 3) puis décrit les algorithmes utilisés dans ces travaux pour la détection des postures (chapitre 4), des sons et de la parole (chapitre 5) puis des activités de la vie quotidienne (chapitre 7), le tout au sein d'un Habitat Intelligent pour la Santé, il s'agit désormais, dans ce chapitre, de discuter de l'apport des différentes parties de ces travaux ainsi que de la validité et de la généralisation des résultats obtenus.

10.1 Apports de ces travaux

10.1.1 ACTIM6D

L'actimétrie embarquée est un domaine de recherche sujet à de nombreuses investigations qui ont produit de nombreuses publications ; chacune des équipes expérimentant des méthodes et des capteurs différents pour résoudre les problèmes. Nous nous étions fixé comme objectif de mettre au point un système embarqué permettant de détecter la posture de la personne (assis, debout, allongé), et les périodes de marche. Ce système devait non seulement être de petite taille mais aussi autonome afin d'être utilisable sur la personne en situation routinière.

Ces objectifs ont été atteints et les premiers résultats sont encourageants. La discussion de la fin du chapitre 4 a déjà commenté ces résultats. Ceux-ci nous montrent que la méthode des ondelettes est capable de nous donner des informations fiables, en utilisant nos propres ondelettes, à partir de nos signaux accélérométriques et magnétométriques.

Cependant, le travail n'est pas terminé sur cette carte et la suite devra porter sur deux points importants :

- Le test en scénario écologique des algorithmes et sur des populations différentes.
- Le développement d'algorithmes embarqués.

Le premier de ces points permettra de compléter les tests que nous avons effectués au chapitre 4. Nos tests ont porté sur un scénario prédéfini, qui est la toute première étape de validation de ce type de système, afin de comparer la réalisation de ce scénario par la personne avec la sortie de l'algorithme de classification et nous permettre ainsi de calculer des pourcentages de bonne classification. Des mesures plus longues et dans des scénarios écologiques, c'est à dire des scénarios de vie non contraints, seront plus informatifs sur les performances que nous pouvons attendre de notre méthode. De plus, évaluer le système sur des populations différentes (typiquement les personnes âgées) pourrait aussi nous indiquer si, comme les autres algorithmes permettant de détecter ces changements de posture, et qui se basent eux sur des seuils d'accélération par exemple, nous devons modifier les réglages de l'algorithme afin d'avoir des résultats équivalents avec les personnes âgées.

Le second point, le développement d'algorithmes embarqués, est important pour l'évaluation de l'utilisabilité de ce système. Il serait intéressant de permettre au système de transmettre sans fil et en continu des informations sur la posture de la personne vers une base centralisatrice. Cette base se chargerait de traitements plus complexes en regroupant un

ensemble d'autres données et ainsi, par une analyse de ces différentes données, pourrait prendre la décision de générer une alarme ou non. Cependant, pour effectuer ce genre de traitement, il faut qu'un minimum de données ait été pré-traitées à l'intérieur même de la carte ACTIM6D. Pour ce faire, une optimisation de la taille des algorithmes et une amélioration de ceux-ci afin de les rendre embarquables sont nécessaires. On peut aussi envisager un changement d'architecture. Cela peut même passer par une combinaison des deux, prendre une architecture un peu plus puissante mais cependant réduire grandement le coût de calcul des algorithmes, afin de diminuer l'augmentation prévisible de la consommation due au nouveau processeur.

10.1.2 Système de reconnaissance de sons et de paroles

Le système de reconnaissance de sons et de paroles de l'équipe GETALP du LIG a été présenté et testé en conditions réelles (chapitre 5). Ce système permet la détection de signaux sonores dans un habitat intelligent équipé d'un ensemble de microphones, et ce avec un seuil adaptatif prenant en considération le bruit ambiant, puis leur segmentation en son de la vie courante ou bien parole, et enfin leur classification en l'une des classes de son pour la première catégorie (bruit de vaisselle, sonnerie de téléphone, claquement de porte...) ou bien la retranscription de la parole la plus probable dans la seconde catégorie. Les sons représentent l'une des modalités qui caractérisent une activité de la vie quotidienne et ce système nous est grandement utile dans l'algorithme de classification afin d'apporter d'autres informations.

La collaboration avec l'équipe GETALP au cours de cette thèse a permis :

- de caractériser, en environnement réel, cette modalité.
- de détecter certains problèmes, de les corriger, et de montrer les limites actuelles de telles détections.

La caractérisation en environnement réel nous donne une indication sur les taux de bonne classification dans le cas de la reconnaissance du son et de la parole. En effet, l'ensemble des tests sur le système avait jusqu'à présent été effectué en laboratoire, puis avec un mélange entre un bruit capté dans le HIS et les sons avec des rapports signal sur bruit différents (voir chapitre 5 pour ces résultats). En se mettant en situation réelle, nous avons été confrontés, entre autre, à des problèmes de non-détection des événements, problème rapidement corrigé, mais également à des rapports signal sur bruit plus bas et des sons perturbateurs plus diversifiés d'une part et plus imprévisibles d'autre part (l'hélicoptère qui décolle de l'héliport proche du laboratoire, une porte qui claque dans le couloir du bâtiment, ...). Cependant, ces conditions écologiques représentent celles que nous aurions dans le cas d'une installation dans un habitat et nous devons prendre en compte ce type de contraintes. L'équipe GETALP améliore continuellement son logiciel et travaille à l'heure actuelle sur l'annulation de sources de sons parasites (typiquement la télévision ou la radio).

Comme énoncé précédemment, tout problème dans le logiciel est en général vite corrigé et nous permet de reprendre les expérimentations. Cependant, nous nous sommes confrontés également à des limites moins facilement dépassables, les corpus. En effet, les corpus

d'apprentissage sont complets mais ne peuvent contenir, pour une serrure de porte, l'ensemble des types de serrures utilisées. Certaines serrures (dont la nôtre) produisent des sons particuliers qui n'étaient pas connus. C'est le cas aussi pour le téléphone avec les sonneries multi-tons. En plus des sons de la vie courante, la parole est également un problème. En effet, il faudrait une réadaptation des modèles de langage afin de générer des sorties plus proches de la réalité. Ces modèles vont être reconstruits par l'équipe GETALP dans les projets à venir afin d'améliorer les taux de bonne détection des algorithmes.

De même, la détection des sons et des situations d'urgence dans un habitat (avec des mots clés par exemple) reste un défi difficile. De nombreuses améliorations et un travail continu dans les années à venir vont rendre les résultats de plus en plus acceptables afin de parvenir, un jour, à un système implantable chez tout un chacun.

10.1.3 Classification des activités

10.1.3.1 Base d'apprentissage

L'objectif de ce travail était de trouver un ensemble minimal de capteurs permettant de caractériser l'ensemble des sept activités (repos, détente, préparation et prise d'un repas, hygiène, élimination, communication avec l'extérieur et habillage/déshabillage) de la vie quotidienne et de les classer correctement. Cela nous a amené dans un premier temps à lister un certain nombre de capteurs et à formater les données correctement afin qu'elles soient reproductibles pour une activité donnée (chapitre 6) puis à implémenter et tester les Séparateurs à Vaste Marge dans ce cas (chapitres 7 pour l'implémentation, 8 pour l'expérimentation et 9 pour la présentation des résultats). Nos résultats montrent un taux de bonne classification de 78,8% dans le meilleur des cas. Ces résultats ont été obtenus par « leave-one-out » du fait du nombre de fenêtres disponibles, afin d'avoir la meilleure idée possible de la capacité des SVM à classer ce type de données. Cela ajoute une catégorie de plus à l'éventail, déjà large, de problèmes auxquels cette méthode est relativement bien adaptée.

Ce taux de bonne classification est relativement élevé, en considérant que nous partons de sept activités différentes, effectuées par treize personnes différentes, libres de se comporter comme elles le souhaitent. La bonne reproductibilité montre que les paramètres choisis sont conformes à une bonne description de l'activité. Cependant, les taux de bonne classification dans certaines des classes sont moins bons (57% pour la classe hygiène et 68% pour la classe Habillage/deshabillage). Ceci s'explique principalement par la faiblesse du nombre de vecteurs dans la base d'apprentissage pour ces classes, qui a pour conséquence de modifier les exemples critiques lors de la création de la séparatrice alors que nous retirons un élément et ainsi de donner des mauvaises classifications. Ces nombres d'exemples faibles dans la base ont également une autre conséquence ; une mauvaise correction des erreurs des capteurs. En effet, ACTIM6D peut créer quelques fausses détections ou non-détections. Sur une base de grande dimension, les erreurs des différents capteurs sont automatiquement intégrées par le système d'apprentissage. En effet, le capteur cinématique n'étant pas parfait, il peut manquer un transfert ou en ajouter un autre. Sur une petite base, les éléments mal

formés (avec des erreurs) ont statistiquement plus de chance de faire partie des exemples critiques que sur les grandes bases.

Afin d'être encore plus représentative, cette expérimentation nécessiterait d'être complétée en augmentant le nombre de réalisations de chacune des activités et le nombre de sujets. Cela nécessiterait d'enregistrer l'activité des personnes volontaires à la journée dans le HIS. Une autre solution serait d'équiper, de la même manière que le HIS (mais sans les caméras vidéo), des appartements en résidence de personnes âgées et de demander, pendant une semaine, aux personnes, de noter les horaires approximatifs de leurs activités. De plus grandes bases d'apprentissage permettraient également de confirmer les choix pour le noyau et les différences entre les noyaux observés sur ces tests.

10.1.3.2 Ajout d'information *a priori*

Les travaux précédents de l'équipe, et notamment ceux de P. Barralon [Barralon, 2005], ont permis une première classification de certaines activités avec un capteur cinématique, les capteurs de présence et le temps. Cette classification s'effectuait à l'aide de l'inférence Bayésienne. Après avoir fixé un ensemble de seuils, la méthode consistait à calculer la probabilité qu'une fenêtre donnée appartienne à une activité donnée. La plus grande probabilité donnait l'activité la plus probablement en cours.

Dans ces travaux, au contraire, nous n'avons en aucun cas inclus de connaissances *a priori*, afin de créer une classification automatique des activités. Toute la connaissance du système d'apprentissage vient de la base. Cependant, par exemple, les détecteurs à infra-rouges nous donnent une bonne connaissance *a priori* sur un sous-ensemble d'activités possible. Il faudrait réfléchir à inclure différents *a priori* dans le système afin d'en améliorer encore la décision. Pour ceci, par contre, il faut une base d'apprentissage plus importante.

De même, une autre perspective de recherche pourrait envisager l'évolution vers des méthodes Bayésiennes en incluant un apprentissage des différentes densités de probabilité (avec un algorithme de type « Markov Chaines Monte-Carlo », MCMC). Cependant, pour avoir une bonne construction de ces différentes densités de probabilité, le nombre de représentant de chaque classe doit être plus important (d'autant que MCMC nécessite de sacrifier un certain nombre d'exemples au début de la base d'apprentissage, appelés « Burn-In samples »).

10.1.3.3 Le temps comme modalité

Nous n'avons inclus aucune connaissance *a priori* dans notre apprentissage, et nous n'avons également pas inclus la notion de temps. Cette décision est justifiée par le fait que notre expérimentation (qui s'est déroulée à toutes heures de la journée, heures de disponibilité des personnes), simulait une matinée (dormir, réveil, petit-déjeuner, élimination, hygiène, habillage). Inclure cette donnée dans le système d'apprentissage aurait nécessité de fixer, nous-mêmes, un *a priori* sur un comportement que nous aurions jugé normal. Nous aurions par exemple pu poser cet *a priori* sur l'heure de réveil et régler l'ensemble du timing

de l'expérimentation sur celui-ci. Cependant, ce sont nos propres connaissances que nous aurions incluses dans le système.

De réelles expérimentations, en conditions écologiques et qui se dérouleraient sur une journée, permettraient de mettre cette donnée à l'intérieur du système. Celle-ci pourrait se réaliser avec {matin, matinée, midi, après-midi, soir, nuit} comme découpage de la journée. Certaines activités se trouveraient localisées sur un ensemble réduit de ce découpage. Dans ce cas, l'horloge représenterait une réelle modalité et non une donnée simulée, et aurait vraiment un sens dans le système d'apprentissage, autre que celui que nous voudrions lui fournir.

10.1.3.4 Classe supplémentaire nécessaire : une classe de rejet

Le but final est tout de même une analyse de la journée de la personne. Cet algorithme de classification n'inclue aucun découpage des activités mais se base sur un temps de fenêtre donné.

Pour surveiller une journée dans son ensemble, il serait nécessaire d'inclure une nouvelle classe dans le système d'apprentissage, la classe de rejet. Cette classe serait en charge de classer correctement les fenêtres de transition entre chacune des activités. Ceci nécessite encore une fois des expérimentations complémentaires afin de créer justement des fenêtres représentatives de cette classe.

10.2 Validité des résultats

10.2.1 « Approche de la réalité par l'expérimentation » versus « Réalité »

Cette première expérimentation nous a donc donné des résultats sur la classification des activités de la vie quotidienne à l'intérieur du HIS de la faculté de médecine de Grenoble, et ce sur treize sujets jeunes et en bonne santé. Ces résultats, issus de validation croisée, montrent la capacité de « généralisation » de l'algorithme de classification utilisé.

Qu'entendons-nous par « capacité de généralisation » ? Celle-ci se définit, en général, pour un algorithme de classification, comme sa capacité à classer correctement une nouvelle donnée ne lui ayant jamais été présentée, et donc qui est en dehors de son corpus d'apprentissage.

Cependant, ce terme de capacité de généralisation pourrait se révéler trompeur. Il faut prendre en compte un facteur supplémentaire que nous avons pu observer durant les sessions expérimentales : l'effet de consigne. L'ensemble des consignes, ainsi que la visite de l'appartement, étaient au départ toujours effectués de la même manière (méthode la plus basique pour n'oublier aucun détail) et dans le même ordre. Ainsi, les activités à réaliser étaient présentées dans le même ordre, les exemples donnés sur la manière d'effectuer l'une ou l'autre de ces activités étaient identiques... En observant la manière dont étaient réalisées ces premières expérimentations, nous avons pu observer des schémas se reproduire : les nôtres. En effet, la personne, soucieuse de bien réussir l'expérimentation, réalise l'activité

telle qu'elle lui a été décrite, ce qui a pour conséquence de réduire la disparité inter-individus dans la réalisation des activités de la vie quotidienne. Bien sûr, il y a un facteur inhérent à la personne d'une part mais aussi dépendant de l'environnement d'autre part, qui fait que d'une expérimentation à l'autre nous aurons des manières d'effectuer une activité un peu différente. Cependant cet effet de consigne est très visible. Aussi, pour les expérimentations suivantes, les consignes étaient données de manière différente afin de rendre un peu plus aléatoires les scénarios et ainsi obtenir des résultats plus intéressants.

L'accent était aussi beaucoup plus marqué sur le fait qu'il leur était demandé d'effectuer ces actions telles qu'ils les effectueraient chez eux, sans se soucier de notre manière de les présenter, mais simplement du fait de les effectuer toutes.

Il n'en reste pas moins que ces données ont été obtenues uniquement à l'intérieur d'un appartement, celui de la faculté de médecine. Le second point qui peut influencer la manière de réaliser une activité est l'environnement dans lequel nous nous trouvons. Nous pouvons penser que, si nous nous positionnons dans un habitat équipé de manière identique en ce qui concerne les capteurs, les données que nous recueillerions dans cet habitat seraient comparables à celles obtenues à la faculté et ne nécessiteraient pas de réapprentissage pour avoir des performances similaires. Cependant plus d'expérimentations sont nécessaires pour que cette croyance, basée sur la connaissance des algorithmes d'apprentissage, se fonde sur des données réelles.

10.2.2 Généralisation des modèles à la personne âgée ?

Les expérimentations présentées ont toutes été réalisées sur des personnes jeunes et en bonne santé alors que l'introduction de ce manuscrit revenait en détail sur notre population vieillissante et les différentes solutions proposées aux personnes âgées. Les tests sur personnes jeunes et les réglages des algorithmes dans ces conditions représentent la première étape avant d'effectuer les expérimentations sur des personnes âgées. En effet, pour avoir des statistiques correctes, le nombre de sujets doit être assez important et réaliser une telle expérimentation sur des personnes âgées va demander un certain temps. Avant de réaliser une telle expérimentation, il était nécessaire d'obtenir des résultats sur une autre population afin de s'apercevoir des différents problèmes méthodologiques et algorithmiques.

Les tests faisant intervenir des personnes âgées représentent la poursuite de ces travaux de thèse. Un modèle de chacune des activités étant construit, il suffit d'utiliser cette base d'apprentissage pour classer les données nouvelles de quelques personnes âgées et voir ainsi les différences de capacité de généralisation de l'algorithme.

Cependant, en quoi les données des personnes âgées pourraient-elles être différentes de celles obtenues chez les personnes jeunes ? En effet, l'ensemble de cette classification se base sur la posture, les sons captés, l'utilisation de placards ou de certaines portes et sur la présence dans une pièce. L'ensemble de ces indicateurs ne devraient pas donner des mesures différentes sur les personnes âgées, si la carte ACTIM6D donne des performances équivalentes pour les personnes de cette population.

En ce qui concerne la reconnaissance de la parole pour la personne âgée, il est par contre

nécessaire de définir des modèles acoustiques adaptés du fait d'une prononciation modifiée par l'âge (cigarette, posture voutée...) ainsi que des modèles de langage plus adaptés à la parole spontanée de la personne âgée (alors que la base actuelle est plus adaptée au français écrit)

L'avenir et l'expérimentation pourront confirmer ou réfuter ce raisonnement. Les partenariats de l'équipe AFIRM et son implication dans l'expérimentation devraient permettre d'obtenir des résultats permettant de répondre à cette question.

Conclusions et perspectives

Sommaire

11.1 Conclusions	187
11.2 Perspectives	188
11.2.1 Correction des données	188
11.2.2 Réalisation d'activité et réalisation effective de celle-ci	188
11.2.3 Transmission de ce type de données au médecin ou à la famille	188
11.2.4 HIS et modèle économique	188

C E travail doctoral n'est bien sûr pas un travail isolé mais s'inscrit dans les activités du laboratoire et de l'équipe AFIRM sur le long terme (avant et après ces travaux). Elle a été dans la continuité de travaux précédemment réalisés et présentés dans le chapitre 3 et la thématique continuera à évoluer avec les prochaines thèses et les prochains travaux du laboratoire. Ce dernier chapitre va présenter les conclusions de ce manuscrit ainsi qu'une ouverture sur les perspectives à venir.

11.1 Conclusions

Ces travaux de thèse ont présenté :

- La réalisation d'un capteur cinématique embarqué, permettant de détecter les transferts posturaux et la marche chez la personne.
- La validation en condition réelle d'un système de classification de sons et de paroles réalisé par l'équipe GETALP du LIG.
- La classification de l'ensemble des données du HIS afin d'arriver à une observation de plus haut niveau : l'activité de la vie quotidienne effectuée par la personne. Cette classification a été validée par des expérimentations dans notre habitat.

Ces travaux dans leur ensemble montrent donc la construction d'un ensemble de données, représentatif pour une activité de la vie quotidienne, puis sa classification avec des séparateurs à vaste marge.

Nous avons équipé l'habitat des éléments suivants :

- Détecteurs de présence infrarouges pour la localisation de la personne
- Contacteurs de porte sur certaines portes du mobilier (réfrigérateur, commode de la chambre, placard de la cuisine)
- Station météorologique dans la salle de bain (température et hygrométrie)
- Capteur cinématique d'activité sur la personne (ACTIM6D)
- Microphones intégrés au plafond
- Caméras de vidéosurveillance pour indexation des activités.

Nos résultats nous ont permis de montrer que la classification de sept activités différentes était réalisable à l'intérieur d'un habitat, et nous avons obtenu un taux de bonne classification global de 78% sur cet algorithme. Nous avons minimisé l'équipement nécessaire à cette classification afin qu'il soit facilement, et à coût acceptable, implantable dans d'autres habitations.

Afin de valider les algorithmes et de construire de bien meilleurs classifieurs, de plus amples expérimentations seront nécessaires. Cela nécessiterait de séjourner plus longtemps dans l'appartement afin d'avoir une bonne représentation d'une journée complète.

Le but de cette recherche était, à terme, d'avoir une vue d'ensemble de la journée de la personne grâce à cette classification. Une fois la base d'apprentissage construite, il est alors nécessaire de classer les données réelles d'une personne sur une journée complète. Ces travaux restent à effectuer par la suite dans le laboratoire.

11.2 Perspectives

11.2.1 Correction des données

Certaines fenêtres pourraient être mal classées au milieu d'une activité, lors de l'analyse d'une journée avec une base d'apprentissage bien définie. Pour corriger ces mauvaises classifications, une correction pourrait se faire par une analyse sémantique de la journée avec des connaissances *a priori* données par une grammaire. Cette analyse interviendrait suite à la classification de chacune des fenêtres d'une journée, en analysant la suite des symboles représentée par la sortie du système de classification pour chacune des fenêtres de trois minutes de la journée.

Cette correction est nécessaire avant de présenter les données sous forme graphique à toute autre personne. En effet, le but de nos recherches est de pouvoir présenter, au médecin ou à la famille, un ensemble de données caractéristiques de la personne âgée. Ces données doivent être lissées et corrigées avant d'être présentées.

11.2.2 Réalisation d'activité et réalisation effective de celle-ci

Nous détectons les différentes activités telles que prendre un repas, ou bien éliminer. Pour l'exemple de ces deux activités, notre détection retrouve un schéma caractéristique de l'activité dans les données de la fenêtre, et classe donc cette activité comme repas ou comme élimination par exemple. Cependant, cela ne donne aucune indication sur la réalisation réelle de cette activité. La personne a-t-elle bien ingéré le repas ? A-t-elle éliminé ?

Pour corriger ce type de problème, certaines équipes, notamment en Asie, analysent, lorsque la chasse d'eau est tirée, les selles de la personne. Cela permettrait d'avoir une indication sur son alimentation. Est-ce une solution acceptable ? Quelles seraient les autres solutions pour détecter la réalisation effective d'une telle activité ? La suite des travaux apportera probablement des réponses à ces questions.

11.2.3 Transmission de ce type de données au médecin ou à la famille

L'intérêt de telles données est qu'elles soient partagées par le médecin, la famille et la personne. Ces trois types de public ne devront pas obtenir les mêmes informations tirées de l'ensemble des capteurs utilisés. En effet, la famille pourra préférer un graphique simple, lui présentant avec un chiffre unique par exemple l'évolution de l'état de son aîné. Le médecin va par contre avoir besoin de plus de détails sur les niveaux d'activité, la réalisation de chacune des activités de la vie quotidienne ... Ce type de demande reste à être formulé par les médecins gériatres lorsque l'installation de ce type d'appartement sera effective.

11.2.4 HIS et modèle économique

Le modèle économique actuel ne permet pas d'installer simplement des appartements, que ce soit en France ou dans le monde. En effet, les recherches bibliographiques montrent

souvent des expérimentations sur le terrain avec un ou deux appartements (voir le chapitre 2) mais rarement un passage à la commercialisation suivie d'une implantation à grande échelle. Ceci est dû au fait que l'installation de ce type d'appartement a un coût et que, tant que les données qui en ressortent ne sont pas assez riches en information pour être utilisées par un médecin, son installation est difficile à prescrire.

De plus, ce type d'installation nécessite une infrastructure autour de l'habitat, avec un centre d'appel et d'aide d'urgence, une communication avec des médecins sur des réseaux sécurisés...

Les années à venir, du fait de l'augmentation de l'âge de la population et de l'amélioration des différents algorithmes, verront probablement cet état de faits changer et le développement de ce type d'habitat en France et dans le monde. Cependant, il faudra trouver un prix acceptable pour un service maximum et une contribution minimum de la personne âgée (mais qui restera probablement bien moindre que celle nécessaire pour l'entrée en maison de retraite).

Bibliographie

- [Abowd *et al.*, 2002] ABOWD, G., MYNATT, E. et RODDEN, T. (2002). The human experience [of ubiquitous computing]. *IEEE Pervasive Computing*, 1(1):48–57.
- [Adami *et al.*, 2003] ADAMI, A., HAYES, T. et PAVEL, M. (2003). Unobtrusive monitoring of sleep patterns. *In Proc. 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 2, pages 1360–1363 Vol.2.
- [Adami *et al.*, 2005] ADAMI, A., HAYES, T., PAVEL, M. et SINGER, C. (2005). Detection and classification of movements in bed using load cells. *In Proc. 27th Annual International Conference of the Engineering in Medicine and Biology Society IEEE-EMBS 2005*, pages 589–592.
- [Aizerman *et al.*, 1964] AIZERMAN, M., BRAVERMAN, E. et ROZONOER, L. (1964). Theoretical foundations of the potential function method in pattern recognition learning. *Automation and Remote Control*, 25:821–837.
- [Akbar et Caelen, 1998] AKBAR, M. et CAELEN, J. (1998). Parole et traduction automatique : le module de reconnaissance raphael. *In Proc. COLING-ACL'98, Montréal, Quebec*, pages 36 – 40.
- [Alwan *et al.*, 2006] ALWAN, M., RAJENDRAN, P., KELL, S., MACK, D., DALAL, S., WOLFE, M. et FELDER, R. (2006). A smart and passive floor-vibration based fall detector for elderly. *In Proc. 2nd Information and Communication Technologies ICTTA '06*, volume 1, pages 1003–1007.
- [Anderson *et al.*, 2008] ANDERSON, D., LUKE, R., SKUBIC, M., KELLER, J., RANTZ, M. et AUD, M. (2008). Evaluation of a video based fall recognition system for elders using voxel space. *In International Conference of the International Society for Gerontechnology*, Pisa, Tuscany, Italy.
- [Anderson *et al.*, 1999] ANDERSON, S., LIBERMAN, N., BERNSTEIN, E., FOSTER, S., CATE, E., LEVIN, B. et HUDSON, R. (1999). Recognition of elderly speech and voice-driven document retrieval. *In Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing ICASSP '99*, volume 1, pages 145–148 vol.1.
- [Baba *et al.*, 2004] BABA, A., YOSHIZAWA, S., YAMADA, M., LEE, A. et SHIKANO, K. (2004). Acoustic models of the elderly for large-vocabulary continuous speech recognition. *Electronics and Communications in Japan, Part 2, Vol. 87, No. 7, 2004*, 87(2):49–57.
- [Barralon, 2005] BARRALON, P. (2005). *Classification et fusion de données actimétriques pour la télévigilance médicale*. Thèse de doctorat, Université Joseph Fourier, Grenoble I.
- [Barralon *et al.*, 2005] BARRALON, P., NOURY, N. et VUILLERME, N. (2005). Classification of daily physical activities from a single kinematic sensor. *In Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, pages 2447–2450.
- [Beaudin *et al.*, 2006] BEAUDIN, J., INTILLE, S. et MORRIS, M. (2006). To track or not to track : User reactions to concepts in longitudinal health monitoring. *Journal of Medical Internet Research*, 8(4):29–47.
- [Berenguer *et al.*, 2008] BERENGUER, M., GIORDANI, M., GIRAUD-BY, F. et NOURY, N. (2008). Automatic detection of activities of daily living from detecting and classifying electrical events on the residential power line. *In HealthCom'08 - 10th IEEE Intl. Conf. on e-Health Networking, Applications and Service*.

- [Bien *et al.*, 2002] BIEN, Z. Z., PARK, K.-H., BANG, W.-C. et STEFANOV, D. H. (2002). Lares : An intelligent sweet home for assisting the elderly and the handicapped.
- [Bonhomme *et al.*, 2008] BONHOMME, S., CAMPO, E., ESTÈVE, D. et GUENNEC, J. (2008). Prosafe-extended, a telemedicine platform to contribute to medical diagnosis. *J Telemed Telecare*, 14(3):116–119.
- [Bonnet *et al.*, 2004] BONNET, S., COUTURIER, P., FAVRE-REGUILLON, F. et GUILLEMAUD, R. (2004). Evaluation of postural stability by means of a single inertial sensor. *In Proc. 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEMBS '04*, volume 1, pages 2275–2278.
- [Bonnet et Heliot, 2007] BONNET, S. et HELIOT, R. (2007). A magnetometer-based approach for studying human movements. *IEEE Transactions on Biomedical Engineering*, 54(7): 1353–1355.
- [Boser *et al.*, 1992] BOSER, B. E., GUYON, I. M. et VAPNIK, V. N. (1992). A training algorithm for optimal margin classifiers. *In Fifth Annual Workshop on Computational Learning Theory, ACM*, pages 144–152, Pittsburgh.
- [Bouchard *et al.*, 2006] BOUCHARD, B., BOUZOUANE, A. et GIROUX, S. (2006). A smart home agent for plan recognition of cognitively-impaired patients. *Journal of Computers*, 1(5):53 – 62.
- [Bourke *et al.*, 2007a] BOURKE, A., O'DONOVAN, K. et OLAIGHIN, G. (2007a). Distinguishing falls from normal adl using vertical velocity profiles. *In Proc. 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBS 2007*, pages 3176–3179.
- [Bourke et Lyons, 2008] BOURKE, A. K. et LYONS, G. M. (2008). A threshold-based fall-detection algorithm using a bi-axial gyroscope sensor. *Med Eng Phys*, 30(1):84–90.
- [Bourke *et al.*, 2007b] BOURKE, A. K., O'BRIEN, J. V. et LYONS, G. M. (2007b). Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm. *Gait and Posture*, 26(2): 194–199.
- [Bouten *et al.*, 1997] BOUTEN, C., KOEKKOEK, K., VERDUIN, M., KODDE, R. et JANSSEN, J. (1997). A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. *IEEE Transactions on Biomedical Engineering*, 44(3):136–147.
- [Bouten *et al.*, 1993] BOUTEN, C., WESTERTERP, K., VERDUIN, M. et JANSSE, J. (1993). A triaxial accelerometer for the assessment of daily physical activity in relation to energy expenditure. *In Proc. 15th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 985–986.
- [Brandes *et al.*, 2008] BRANDES, M., SCHOMAKER, R., MÖLLENHOFF, G. et ROSENBAUM, D. (2008). Quantity versus quality of gait and quality of life in patients with osteoarthritis. *Gait Posture*, 28(1):74–79.
- [Brandes *et al.*, 2006] BRANDES, M., ZIJLSTRA, W., HEIKENS, S., van LUMMEL, R. et ROSENBAUM, D. (2006). Accelerometry based assessment of gait parameters in children. *Gait Posture*, 24(4):482–486.
- [Burges, 1998] BURGES, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2):121–167.
- [Busmann *et al.*, 1998] BUSSMANN, J. B., TULEN, J. H., van HEREL, E. C. et STAM, H. J. (1998). Quantification of physical activities by means of ambulatory accelerometry : a validation study. *Psychophysiology*, 35(5):488–496.
- [Busmann *et al.*, 2004] BUSSMANN, J. B. J., van den BERG-EMONS, H. J. G., ANGULO, S. M., STIJNEN, T. et STAM, H. J. (2004). Sensitivity and reproducibility of accelerometry and heart rate in physical strain assessment during prosthetic gait. *Eur J Appl Physiol*, 91(1):71–78.

- [Chambodut, 2004] CHAMBODUT, A. (2004). *Le Champ Magnétique Terrestre : Structure Spatiale et Variation Temporelle Vues par les Ondelettes*. Thèse de doctorat, Institut de Physique du Globe de Paris.
- [Chan *et al.*, 2006] CHAN, H.-L., FANG, S.-C., KO, Y.-L., LIN, M.-A., HUANG, H.-H. et LIN, C.-H. (2006). Heart rate variability characterization in daily physical activities using wavelet analysis and multilayer fuzzy activity clustering. *IEEE Transactions on Biomedical Engineering*, 53(1):133–139.
- [Chapelle, 2004] CHAPELLE, O. (2004). *Support Vector Machines : Principes d'induction, réglages automatiques et connaissance à priori*. Thèse de doctorat, Université Pierre et Marie Curie, Paris VI.
- [Cook *et al.*, 2003] COOK, D., YOUNGBLOOD, M., HEIERMAN, E.O., I., GOPALRATNAM, K., RAO, S., LITVIN, A. et KHAWAJA, F. (2003). Mavhome : an agent-based smart home. *In Proc. First IEEE International Conference on Pervasive Computing and Communications (PerCom 2003)*, pages 521–524.
- [Cortes et Vapnik, 1995] CORTES, C. et VAPNIK, V. (1995). Support vector networks. *Machine Learning*, 20:273–297.
- [Coudin, 2008] COUDIN, E. (2008). Projection de population active à l'horizon 2050 : des actifs en nombre stable pour une population âgée toujours plus nombreuse. *Économie et Statistiques*, (408–409):113–136.
- [Das *et al.*, 2002] DAS, S., COOK, D., BATTACHARYA, A., HEIERMAN, E.O., I. et LIN, T.-Y. (2002). The role of prediction algorithms in the mavhome smart home architecture. *IEEE Wireless Communications Magazine*, 9(6):77–84.
- [Daubechies, 1988] DAUBECHIES, I. (1988). Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 41(7):909–996.
- [Dion et Pigot, 2007] DION, A. et PIGOT, H. (2007). Modeling cognitive errors in the realization of an activity of the everyday life. *In Cognition 2007*.
- [Duchêne, 2004] DUCHÊNE, F. (2004). *Fusion de données multicapteurs pour un système de télésurveillance médicale de personnes à domicile*. Thèse de doctorat, Université Joseph Fourier, Grenoble.
- [Elger et Furugren, 1998] ELGER, G. et FURUGREN, B. (1998). *SmartBo – An ICT and computer-based demonstration home for disabled people*. Conférence TIDE, Helsinki, Finlande.
- [Favre *et al.*, 2006a] FAVRE, J., JOLLES, B. M., SIEGRIST, O. et AMINIAN, K. (2006a). Quaternion-based fusion of gyroscopes and accelerometers to improve 3D angle measurement. *Electronics Letters*, 42(11):612–614.
- [Favre *et al.*, 2006b] FAVRE, J., LUTHI, F., JOLLES, B. M., SIEGRIST, O., NAJAFI, B. et AMINIAN, K. (2006b). A new ambulatory system for comparative evaluation of the three-dimensional knee kinematics, applied to anterior cruciate ligament injuries. *Knee Surgery Sports Traumatology Arthroscopy*, 14(7):592–604.
- [Fezari et Bousbia-Salah, 2007] FEZARI, M. et BOUSBIA-SALAH, M. (2007). Speech and sensor in guiding an electric wheelchair. *Automatic Control and Computer Sciences*, 41(1):39–43.
- [Gauvain *et al.*, 1990] GAUVAIN, J.-L., LAMEL, L.-F. et ESKENAZI, M. (1990). Design considerations and text selection for bref, a large french read-speech corpus. *In ICSLP'90*, pages 1097–1100, Kobe, Japan.
- [Glasson, 2008] GLASSON, H. (2008). *Système autonome d'acquisition temps-réel multivoie de signaux sonores dans un Habitat Intelligent pour la Santé (HIS)*. Mémoire C.N.A.M., Conservatoire National des Arts et Métiers, Grenoble, France, 171p.

- [Hagen *et al.*, 2001] HAGEN, I., HOLTHE, T., DUFF, P., CAHILL, S., GILLIARD, J., ORPWOOD, R., TOPO, P. et BJØRNEBY, S. (2001). Can assistive technology enable people with dementia? *Assistive Technology- Added Value to the Quality of Life*, pages 42 – 47.
- [Hagen *et al.*, 2003] HAGEN, I., HOLTHE, T., DUFF, P., CAHILL, S., JONES, K., TOPO, P. et BJØRNEBY, S. (2003). A general structure for assessment studies of assistive technology. *Assistive Technology - Shaping the Future*, pages 393–398.
- [Haley et Manjunath, 1999] HALEY, G. M. et MANJUNATH, B. S. (1999). Rotation-invariant texture classification using a complete space-frequency model. *IEEE Transactions on Image Processing*, 8(2):255 – 269.
- [Hausdorff *et al.*, 2004] HAUSDORFF, J. M., PENG, C.-K., GOLDBERGER, A. L. et STOLL, A. L. (2004). Gait unsteadiness and fall risk in two affective disorders : a preliminary study. *BMC Psychiatry*, 4:39.
- [Helal *et al.*, 2003] HELAL, S., WINKLER, B., LEE, C., KADDOURA, Y., RAN, L., GIRALDO, C., KUCHIBHOTLA, S. et MANN, W. (2003). Enabling location-aware pervasive computing applications for the elderly. *In Proc. First IEEE International Conference on Pervasive Computing and Communications (PerCom 2003)*, pages 531–536.
- [Hong *et al.*, 2008] HONG, X., NUGENT, C., MULVENNA, M., MCCLEAN, S. et SCOTNEY, B. (2008). Evidential fusion of sensor data for activity recognition in smart homes. *Pervasive and Mobile Computing*, pages 1–17.
- [Héliot, 2007] HÉLIOT, R. (2007). *Restauration fonctionnelle de la posture et de la marche : vers la coordination des membres valides et déficients*. Thèse de doctorat, Institut National Polytechnique de Grenoble.
- [Intille, 2002] INTILLE, S. (2002). Designing a home of the future. *IEEE Pervasive Computing*, 1(2):76–82.
- [Intille, 2004] INTILLE, S. (2004). A new research challenge : persuasive technology to motivate healthy aging. *IEEE Transactions on Information Technology in Biomedicine*, 8(3): 235–237.
- [Istrate *et al.*, 2006a] ISTRATE, D., CASTELLI, E., VACHER, M., BESACIER, L. et SERIGNAT, J.-F. (2006a). Information extraction from sound for medical telemonitoring. *IEEE Transactions on Information Technology in Biomedicine*, 10(2):264–274.
- [Istrate *et al.*, 2006b] ISTRATE, D., VACHER, M. et SERIGNAT, J. F. (2006b). Generic implementation of a distress sound extraction system for elder care. *In Proc. 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBS '06*, pages 3309–3312.
- [Istrate, 2003] ISTRATE, D. M. (2003). *Détection et reconnaissances des sons pour la surveillance médicale*. Thèse de doctorat, Institut National Polytechnique de Grenoble.
- [Jolles *et al.*, 2004] JOLLES, B., AMINIAN, K., DEJNABADI, H., NAJAFI, B., TREVISAN, C., MARINONI, E., DUTOIT, M. et LEYVRAZ, P. (2004). Analyse ambulatoire de la marche après prothèse totale de hanche : évaluation des résultats avec un nouveau paramètre. *Rev Chir Orthop*, 90(6):104.
- [Katz et Akpom, 1976] KATZ, S. et AKPOM, C. (1976). A measure of primary sociobiological functions. *International Journal of Health Services*, 6(3):493–508.
- [Kientz *et al.*, 2007] KIENTZ, J., HAYES, G., WESTEYN, T., STARNER, T. et ABOWD, G. (2007). Pervasive computing and autism : Assisting caregivers of children with special needs. *IEEE Pervasive Computing*, 6(1):28–35.
- [Krumm *et al.*, 2000] KRUMM, J., HARRIS, S., MEYERS, B., BRUMITT, B., HALE, M. et SHAFER, S. (2000). Multi-camera multi-person tracking for easy living. *In Proc. Third IEEE International Workshop on Visual Surveillance*, pages 3–10.

- [Kröse *et al.*, 2008] KRÖSE, B., van KASTEREN, T., GIBSON, C. et van den DOOL, T. (2008). Care : Context awareness in residences for elderly. *In International Conference of the International Society for Gerontechnology*, Pisa, Tuscany, Italy.
- [Kumiko *et al.*, 2004] KUMIKO, O., MITSUHIRO, M., ATSUSHI, E., SHOHEI, S. et REIKO, T. (2004). Input support for elderly people using speech recognition. *IEIC Technical Report (Institute of Electronics, Information and Communication Engineers)*, 104(139):1–6.
- [Lai *et al.*, 1999] LAI, J. H., YUEN, P. C. et C., F. G. (1999). Spectroface : A fourier-based approach for human face recognition. *In Proceedings of ICMI '99*, pages VII15–120.
- [Lawton et Brody, 1969] LAWTON, M. et BRODY, E. (1969). Assessment of older people : Self-maintaining and instrumental activities of daily living. *Gerontologist*, 9:179–186.
- [Le *et al.*, 2007] LE, X. H. B., DI MASCOLO, M., GOUIN, A. et NOURY, N. (2007). Health smart home - towards an assistant tool for automatic assessment of the dependence of elders. *In Proc. 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBS 2007*, pages 3806–3809.
- [LeBellego *et al.*, 2006] LEBELLEGO, G., NOURY, N., VIRONE, G., MOUSSEAU, M. et DEMONGEOT, J. (2006). A model for the measurement of patient activity in a hospital suite. *IEEE Transactions on Information Technology in Biomedicine*, 10(1):92–99.
- [Luinge et Veltink, 2004] LUINGE, H. et VELTINK, P. (2004). Inclination measurement of human movement using a 3-d accelerometer with autocalibration. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 12(1):112–121.
- [Lyons *et al.*, 2005] LYONS, G. M., CULHANE, K. M., HILTON, D., GRACE, P. A. et LYONS, D. (2005). A description of an accelerometer-based mobility monitoring technique. *Med Eng Phys*, 27(6):497–504.
- [Léon, 2008] LÉON, O. (2008). Les projections régionales de population 2005-2030. *Économie et Statistiques*, (408–409):137–152.
- [Marins *et al.*, 2001] MARINS, J., YUN, X., BACHMANN, E., MCGHEE, R. et ZYDA, M. (2001). An extended kalman filter for quaternion-based orientation estimation using marg sensors. *In Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 4, pages 2003–2011.
- [McGhee *et al.*, 2000] MCGHEE, R. B., BACHMANN, E. R., YUN, X. et ZYDA, M. J. (2000). Real-time tracking and display of human limb segment motions using sourceless sensors and a quaternion-based filtering algorithm. Rapport technique, Naval Postgraduate School, Monterey, California.
- [McGhee *et al.*, 2001] MCGHEE, R. B., BACHMANN, E. R., YUN, X. et ZYDA, M. J. (2001). An investigation of alternative algorithms for singularity - free estimation of rigid body orientation from earth gravity and magnetic field measurements. Rapport technique, Naval Postgraduate School, Monterey, California.
- [Medjahed *et al.*, 2008] MEDJAHED, H., ISTRATE, D., BOUDY, J., STEENKESTE, F., BALDINGER, J., BELFEKI, I., MARTINS, V. et B. DORIZZI (2008). A multimodal platform for database recording and elderly people monitoring. *In accepted to BIOSIGNALS 2008*, Funchal-Madeira, Portugal.
- [Mercer, 1909] MERCER, J. (1909). Functions of positive and negative type and their connection with the theory of integral equations. *Philosophical Transactions of the Royal Society of London. Series A*, 209:415–446.
- [Misiti *et al.*, 2003] MISITI, M., MISITI, Y., OPPENHEIM, G. et POGGI, J.-M. (2003). *Les ondelettes et leurs applications*. Lavoisier.

- [Moore et Essa, 2002] MOORE, D. et ESSA, I. (2002). Recognizing multitasked activities from video using stochastic context-free grammar. *In Proceedings of American Association of Artificial Intelligence (AAAI) Conference 2002*, Alberta, Canada.
- [Muscillo et al., 2007] MUSCILLO, R., CONFORTO, S., SCHMID, M., CASELLI, P. et D'ALESSIO, T. (2007). Classification of motor activities through derivative dynamic time warping applied on accelerometer data. *In Proc. 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBS 2007*, pages 4930–4933.
- [Najafi et al., 2002] NAJAFI, B., AMINIAN, K., LOEW, F., BLANC, Y. et ROBERT, P. (2002). Measurement of stand-sit and sit-stand transitions using a miniature gyroscope and its application in fall risk evaluation in the elderly. *Ieee Transactions on Biomedical Engineering*, 49(8):843–851.
- [Nakajima et al., 2008] NAKAJIMA, K., MATSUI, H., MOTOYA, T. et SASAKI, K. (2008). Television-operating-state telemonitoring system in caring for elderly people living alone. *In International Conference of the International Society for Gerontechnology*, Pisa, Tuscany, Italy.
- [Noury, 2002] NOURY, N. (2002). A smart sensor for the remote follow up of activity and fall detection of the elderly. *In Proc. Microtechnologies in Medicine & Biology 2nd Annual International IEEE-EMB Special Topic Conference on*, pages 314–317.
- [Noury et al., 2004] NOURY, N., BARRALON, P., COUTURIER, P., FAVRE-REGUILLON, F., GUILLEMAUD, R., MESTAIS, C., CARITU, Y., DAVID, D., MOINE, S., FRANCO, A., GUIRAUD-BY, F., BERRENGUER, M. et PROVOST, H. (2004). Actidom - a microsystem based on mems for activity monitoring of the frail elderly in their daily life. *In Proc. 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEMBS '04*, volume 2, pages 3305–3308.
- [Noury et al., 2007] NOURY, N., FLEURY, A., RUMEAU, P., BOURKE, A., LAIGHIN, G. O., RIALLE, V. et LUNDY, J. (2007). Fall detection - principles and methods. *In Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 1663–1666.
- [Noury et al., 2006] NOURY, N., VILLEMASET, C., BARRALON, P. et RUMEAU, P. (2006). Ambient multi-perceptive system for residential health monitoring based on electronic mailings experimentation within the ailisa project. *In Proc. 8th International Conference on e-Health Networking, Applications and Services HEALTHCOM 2006*, pages 95–100.
- [Philipose et al., 2004] PHILIPOSE, M., FISHKIN, K. P., PERKOWITZ, M., PATTERSON, D. J., FOX, D., KAUTZ, H. et HAHNEL, D. (2004). Inferring activities from interactions with objects. *IEEE Pervasive Computing*, 3(4):50 – 57.
- [Quagliarella et al., 2008] QUAGLIARELLA, L., SASANELLI, N. et BELGIOVINE, G. (2008). A fall and loss of consciousness wearable detector. *In International Conference of the International Society for Gerontechnology*, Pisa, Tuscany, Italy.
- [Renouard et al., 2003] RENOARD, S., CHARBIT, M. et CHOLLET, G. (2003). Vocal interface with a speech memory for dependent people. *Independent Living for Persons with Disabilities*, pages 15–21.
- [Rialle, 2007] RIALLE, V. (2007). *Technologie et Alzheimer : appréciation de la faisabilité de la mise en place de technologies innovantes pour assister les aidants familiaux et pallier les pathologies de type Alzheimer*. Thèse de doctorat, Université René Descartes-Paris 5.
- [Ricquebourg et al., 2006] RICQUEBOURG, V., MARHIC, B., DELAHOCHÉ, L., MENGA, D. et LOGÉ, C. (2006). L'habitat intelligent au service du maintien à domicile : Application au suivi d'individus. *In HANDICAP'2006*.

- [Robert-Bobée, 2008] ROBERT-BOBÉE, I. (2008). Projections de population 2005-2050 – vieillissement de la population en France métropolitaine. *Économie et Statistiques*, (408–409):95–112.
- [Rumeau *et al.*, 2006] RUMEAU, P., RIALLE, V. et NOURY, N. (2006). A priori evaluation of acceptance of an activity monitoring device for the disabled elderly people using the his as a model. In *4th International Conference on smart homes and Health Telematics, ICOST2006, Belfast-Irelande, 26-28 Juin 2006, "Smart Homes and Beyond"*, editors Chris Nugent and Juan Carlos Augusto, IOS Press, pages 130 – 137.
- [Salarian *et al.*, 2007] SALARIAN, A., RUSSMANN, H., WIDER, C., BURKHARD, P., VINGERHOETS, F. et AMINIAN, K. (2007). Quantification of tremor and bradykinesia in Parkinson's disease using a novel ambulatory monitoring system. *IEEE Transactions on Biomedical Engineering*, 54(2):313–322.
- [Salmeron-Quiraz, 2007] SALMERON-QUIRAZ, B. B. (2007). *Fusion de données multi-capteurs pour la capture de mouvement*. Thèse de doctorat, Université Joseph Fourier.
- [Samama *et al.*, 2005] SAMAMA, N., VERVISCH-PICOIS, A. et FRANÇOIS, M. (2005). La localisation en intérieur à l'aide de répéteurs gps vers un système de positionnement universel? In *Ubimob'05 - Deuxièmes Journées Francophones : Mobilité et Ubiquité 2005*, Grenoble.
- [Scanail et al., 2006] SCANAILL, C. N., AHEARNE, B. et LYONS, G. M. (2006). Long-term telemonitoring of mobility trends of elderly people using sms messaging. *IEEE Trans Inf Technol Biomed*, 10(2):412–413.
- [Schwarz, 1978] SCHWARZ, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6:461–464.
- [Selin *et al.*, 2007] SELIN, A., TURUNEN, J. et TANTTU, J. T. (2007). Wavelets in recognition of bird sounds. *EURASIP Journal on Applied Signal Processing*, 2007:141–150.
- [Serna *et al.*, 2007] SERNA, A., PIGOT, H. et RIALLE, V. (2007). Modeling the progression of alzheimer's disease for cognitive assistance and smart homes. *User Modeling and User-Adapted Interaction*, 17:415–438.
- [Stam *et al.*, 1995] STAM, H. J., EIJSKOOT, F. et BUSSMANN, J. B. (1995). A device for long-term ambulatory monitoring in trans-tibial amputees. *Prosthet Orthot Int*, 19(1):53–55.
- [Steele *et al.*, 2003] STEELE, B. G., BELZA, B., CAIN, K., WARMS, C., COPPERSMITH, J. et HOWARD, J. (2003). Bodies in motion : monitoring daily activity and exercise with motion sensors in people with chronic pulmonary disease. *J Rehabil Res Dev*, 40(5 Suppl 2):45–58.
- [Steele *et al.*, 2000] STEELE, B. G., HOLT, L., BELZA, B., FERRIS, S., LAKSHMINARYAN, S. et BUCHNER, D. M. (2000). Quantitating physical activity in copd using a triaxial accelerometer. *Chest*, 117(5):1359–1367.
- [Steenkeste *et al.*, 2001] STEENKESTE, F., BOCQUET, H. et M. CHAN, E. C. (2001). La mise en place d'une technologie pour observer le comportement nocturne des personnes âgées en institution. *ITBM-RBM*, 22(1):25–30.
- [Tait, 1873] TAIT, P. G. (1873). *An elementary treatise on quaternions, 2d edition*. Cambridge University Press.
- [Takahashi *et al.*, 2003] TAKAHASHI, S.-y., MORIMOTO, T., MAEDA, S. et TSURUTA, N. (2003). Dialogue experiment for elderly people in home health care system. In *Text Speech and Dialogue (TSD) 2003*.
- [Tang *et al.*, 2000] TANG, Y. Y., YANG, L. H., LIU, J. et MA, H. (2000). *Wavelet Theory and its Application to Pattern Recognition*. World Scientific.
- [Tapia *et al.*, 2007] TAPIA, E., INTILLE, S., HASKELL, W., LARSON, K., WRIGHT, J., KING, A. et FRIEDMAN, R. (2007). Real-time recognition of physical activities and their intensities

- using wireless accelerometers and a heart rate monitor. *In Proc. 11th IEEE International Symposium on Wearable Computers*, pages 37–40.
- [Thomesse *et al.*, 2001] THOMESSE, J.-P., BELLOT, D., BOYER, A., CAMPO, E., CHAN, M., CHARPILLET, F., FAYN, J., LESCHI, C., NOURY, N., RIALLE, V., ROMARY, L., RUBEL, P., SELMAOUI, N., STEENKESTE, F. et VIRONE, G. (2001). Integrated information technologies for patients remote follow-up and homecare. *In 3rd International Workshop on Enterprise Networking and Computing in Health Care Industry - HealthCom (l'Aquila, Italie)*.
- [Toreyin *et al.*, 2006] TOREYIN, B., DEDEOGLU, Y. et CETIN, A. (2006). Hmm based falling person detection using both audio and video. *In Proc. IEEE 14th Signal Processing and Communications Applications*, pages 1–4.
- [Tran et Mynatt, 2003] TRAN, Q. T. et MYNATT, E. D. (2003). What was i cooking? towards déjà vu displays of everyday memory. Rapport technique.
- [Tsukamoto *et al.*, 2008] TSUKAMOTO, S., HOSHINO, H. et TAMURA, T. (2008). Study on indoor activity monitoring by using electric field sensor. *In International Conference of the International Society for Gerontechnology*, Pisa, Tuscany, Italy.
- [Vacher *et al.*, 2004] VACHER, M., ISTRATE, D. et SERIGNAT, J. (2004). Sound detection and classification through transient models using wavelet coefficient trees. *In LTD, S., éditeur : Proceedings of the 12th European Signal Processing Conference*, pages 1171–1174, Vienna, Austria.
- [Vacher *et al.*, 2006] VACHER, M., SERIGNAT, J., CHAILLOL, S., ISTRATE, D. et POPESCU, V. (2006). Speech and sound use in a remote monitoring system for health care. *Lecture Notes in Artificial Intelligence*, 4188/2006:711–718.
- [Vapnik et Chervonenkis, 1971] VAPNIK, V. et CHERVONENKIS, A. (1971). On the uniform convergence of relative frequencies of events to their probabilities. *Th. Prob. and its Appl.*, 16(2):264–80.
- [Vaufreydaz *et al.*, 2000] VAUFREYDAZ, D., BERGAMINI, C., SERIGNAT, J.-F., BESACIER, L. et AKBAR, M. (2000). A new methodology for speech corpora and definition from internet documents. *In LREC'2000, 2nd International Conference on Language Ressources and Evaluation*, pages 423–426, Athens, Greece.
- [Veltink *et al.*, 1993] VELTINK, P., BUSSMANN, H., KOELMA, F., FRANKEN, H., MARTENS, W. et van LUMMEL, R. (1993). The feasibility of posture and movement detection by accelerometry. *In Proc. 15th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1230–1231.
- [Villemazet, 2004] VILLEMAZET, C. (2004). *Industrialisation du système d'information de l'Habitat Intelligent pour la Santé*. Rapport de Fin d'Étude d'Ingénieur, Polytech'Grenoble, Université Joseph Fourier, Grenoble I.
- [Virone, 2003] VIRONE, G. (2003). *Architecture et Simulation Locales du Système d'Information Domotique-Santé intégré à Domicile (SID²) pour la Détection de Situations à Risque et l'Aide à la Décision*. Thèse de doctorat, Université Joseph Fourier, Grenoble I.
- [Virone et Istrate, 2007] VIRONE, G. et ISTRATE, D. (2007). Integration of an environmental sound module to an existing in-home activity simulator. *In IEEE EMBC 2007, August 23-26, Lyon, France*, pages 3810–3814.
- [Wang *et al.*, 2008] WANG, J.-C., LEE, H.-P., WANG, J.-F. et LIN, C.-B. (2008). Robust environmental sound recognition for home automation. *IEEE Transactions on Automation Science and Engineering*, 5(1):25–31.
- [Yamazaki, 2007] YAMAZAKI, T. (2007). The ubiquitous home. *International Journal of Smart Homes*, 1(1):17–22.

- [Youngblood *et al.*, 2005] YOUNGBLOOD, G., HOLDER, L. et COOK, D. (2005). Managing adaptive versatile environments. *In Proc. Third IEEE International Conference on Pervasive Computing and Communications PerCom 2005*, pages 351–360.
- [Zouba, 2006] ZOUBA, N. (2006). Analyse multicateurs pour la reconnaissance d'activités humaines. application au projet gerhome : Maintien à domicile des personnes Âgées. Mémoire de D.E.A., Université Paris 8, Vincennes Saint-Denis.

Table des figures

1.1	Évolution de la population française au cours des vingt dernières années. Le trait plus épais représente le rapport de soutien démographique, en pourcent, dont l'échelle se trouve sur la droite	20
1.2	Répartition de la population française par âge et par sexe au 1er janvier 2008	21
1.3	Répartition de la population française (prévisions 2005–2050)	22
1.4	Répartition de la population mondiale (1950–2050)	23
1.5	Pourcentage de la population âgé de plus de 65 ans (1950–2050)	24
2.1	Centrales de mesures. De gauche à droite : Dynaport de Mc Roberts, Physilog du LMAM, Trident du CEA-LETI	32
2.2	Détecteurs de chute, de gauche à droite : Le détecteur de chute du TIMC-IMAG, le téléphone d'Attendance, la montre Vivago.	34
2.3	Exemple de la base Quiatil+ d'Intervox, avec ses trois boutons simples d'utilisation : Urgence, appel d'un proche, appel des services à domicile	38
2.4	Outils de géolocalisation pour les personnes âgées. De gauche à droite : la Locate Box, le bracelet Columba et la balise Aloise	40
4.1	Prototype du circuit ACTIM6D	63
4.2	Positionnement du capteur sur le sujet et trièdre direct créé par la centrale et fixe sur la personne (figure de gauche). Positionnement de ce trièdre sur le capteur embarqué (figure de droite)	65
4.3	Diagramme décrivant le fonctionnement d'ACTIM6D. Le diagramme de gauche représente la procédure principale et le diagramme de droite la procédure d'interruption associée au timer.	66
4.4	Signaux accélérométriques (à gauche) et magnétométriques (à droite), période d'immobilité éloignée de fortes perturbations électromagnétiques	68
4.5	Signaux accélérométriques (figure de gauche) et magnétométriques (figure de droite), période d'immobilité à proximité (20cm) d'un ascenseur	69
4.6	Signaux des accéléromètres pour une séquence de marche lente puis une séquence de marche rapide. À gauche, les signaux bruts et à droite, les signaux filtrés par le filtre de Butterworth	71
4.7	Signaux filtrés des accéléromètres de dix séquences Debout-Allongé-Debout à gauche. Agrandissement sur deux des séquences à droite	72
4.8	Signaux filtrés des accéléromètres de dix séquences Debout-Assis-Debout à gauche. Agrandissement sur deux des séquences à droite	72
4.9	Positionnement du dipôle magnétique équivalent au champ magnétique terrestre	73
4.10	Orientation du champ magnétique Terrestre ([Chambodut, 2004])	73
4.11	Signaux filtrés des magnétomètres. À gauche pour une séquence de dix Debout-Assis-Debout et à droite pour une séquence de dix Debout-Allongé-Debout	74
4.12	Signaux filtrés des magnétomètres pour une séquence de dix Debout-Assis-Debout. À gauche sur une première chaise et à droite sur une seconde placée à 90° de la première	75

4.13 Résultats de l'algorithme de segmentation des signaux, à gauche la segmentation d'une séquence Debout-Assis-Debout répétée dix fois, à droite la segmentation d'une séquence Debout-Allongé-Debout répétée dix fois	78
4.14 Séquence Debout-Assis sur une première chaise. À gauche le modèle utilisé pour la classification ultérieure, à droite une réalisation du mouvement. Les abscisses représentent le numéro d'échantillon et l'ordonnée la valeur en g pour les accéléromètres et en radian pour les angles	82
4.15 Séquence Debout-Assis sur une seconde chaise (à 90°). À gauche le modèle utilisé, à droite une réalisation du mouvement	82
4.16 Construction des ondelettes adaptées pour la séquence Debout-Assis. En rouge le signal d'origine, en bleu l'ondelette correspondante respectant les critères d'admissibilité	82
4.17 Séquence Debout-Assis sur la première chaise. À gauche le modèle utilisé, à droite une réalisation du mouvement.	83
4.18 Séquence Debout-Assis sur la seconde chaise. À gauche le modèle utilisé, à droite une réalisation du mouvement.	83
4.19 Construction des ondelettes adaptées pour la séquence Assis-Debout. En rouge le signal d'origine, en bleu l'ondelette correspondante respectant les critères d'admissibilité	83
4.20 Séquence Debout-Allongé. À gauche le modèle utilisé, à droite une réalisation du mouvement.	84
4.21 Construction des ondelettes adaptées pour la séquence Debout-Allongé. En rouge le signal d'origine, en bleu l'ondelette correspondante respectant les critères d'admissibilité	84
4.22 Séquence Allongé-Debout. À gauche le modèle utilisé, à droite une réalisation du mouvement.	85
4.23 Construction des ondelettes adaptées pour la séquence Allongé-Debout. En rouge le signal d'origine, en bleu l'ondelette correspondante respectant les critères d'admissibilité	86
4.24 Signaux filtrés des accéléromètres à gauche et des magnétomètres à droite correspondant à une séquence de dix aller-retours (distance de 15m) à marche lente puis dix à marche rapide	86
4.25 Résultats de la détection de la marche sur deux séquences. À gauche, une séquence de dix aller-retours d'une distance de 15m environ, d'abord lents, puis rapides. À droite, une séquence de cinq sessions de marche (plus deux sur le début et la fin pour se rendre jusqu'à la chaise), chaque séquence étant composée de 80m de marche suivie d'un transfert Debout-Assis	88
4.26 Résultats de l'algorithme sur un scénario de dix séquences de Debout-Allongé-Debout	89
4.27 Résultats de l'algorithme sur un scénario de dix séquences Debout-Assis-Debout	90
4.28 Résultats de l'algorithme sur un scénario de Up&Go	91
4.29 Transformée en ondelettes avec le chapeau mexicain du signal de marche des accéléromètres	97
5.1 Organisation globale du système de reconnaissance de sons et de la parole . . .	101
5.2 Répartition des microphones dans l'appartement HIS et intégration dans l'environnement	104
5.3 Construction de l'arbre pour l'analyse multi-résolution du signal avec les ondelettes de Daubechies [Vacher <i>et al.</i> , 2004]	107
5.4 Algorithme de détection des sons [Vacher <i>et al.</i> , 2004]	107
5.5 Banc de filtres pour le calcul des coefficients LFCC [Glasson, 2008]	108

5.6	Banc de filtres pour le calcul des coefficients MFCC (échelle Mel) [Glasson, 2008]	108
5.7	Organisation générale de l'application AuditHIS	115
5.8	Capture d'écran de la fenêtre principale de l'application AuditHIS	115
5.9	Capture d'écran de la fenêtre de configuration du logiciel AuditHIS	116
5.10	Résultat de la détection des mots de détresse par locuteur	125
6.1	Positionnement des différents équipements dans l'Habitat Intelligent pour la Santé de Grenoble	129
6.2	Recepteur radio et bus CAN du HIS	132
6.3	Carte SmartCAN 2 gérant la transmission des données des détecteurs infrarouges sur le bus CAN	132
6.4	Interface Homme-Machine du logiciel AILISA de recueil et de traitement des données de l'HIS	134
6.5	Station météorologique utilisée pour l'enregistrement des données environnementales dans le HIS	137
7.1	Représentation sous forme de plusieurs modalités de deux des activités de la vie quotidienne	142
7.2	Cas linéairement séparable. Un ensemble de points blancs et gris du plan. Des exemples de séparation linéaires possibles sont également représentées	145
7.3	Cas linéairement séparable résolu avec les SVM	146
7.4	Cas non linéairement séparable de classification binaire	148
7.5	Méthode Un-Contre-Tous pour les SVM multiclassées	150
7.6	Méthode Un-Contre-Un pour les SVM multiclassées	151
9.1	Taux d'erreur global de l'algorithme de classification en fonction du paramètre σ du noyau	171

Liste des tableaux

3.1	Grille ADL de Katz avec les scores associés à chaque comportement	54
3.2	Échelle des Activités Instrumentales de la Vie Quotidienne de Lawton	56
4.1	Matrice de confusion pour les mouvements Assis-Debout et Debout-Assis . . .	92
4.2	Matrice de confusion pour les mouvements Allongé-Debout et Debout-Allongé	93
5.1	Répartition des éléments de chacune des classes de la base d'apprentissage des sons	111
5.2	Taux d'erreur de détection pour un bruit blanc et un bruit de fond capté dans le HIS, à différents rapports signal sur bruit	113
5.3	Taux d'erreur de segmentation Son/Parole	113
5.4	Erreur de classification entre les 8 classes de sons	113
5.5	Taux d'erreur de reconnaissance de la parole	114
5.6	Performances globales du système de reconnaissance de la parole	120
5.7	Performances globales du système de reconnaissance de sons	121
5.8	Matrice de confusion pour la reconnaissance des sons. Les valeurs en gras correspondent aux sons correctement classifiés.	122
5.9	Matrice de confusion pour la reconnaissance de sons et de la parole. Les valeurs en gras représentent les bonnes classifications.	123
5.10	Taux d'erreur de segmentation/classification pour la reconnaissance des phrases anodine/détresse	124
5.11	Matrice de confusion pour la reconnaissance de sons et de la parole. Les valeurs en gras représentent les bonnes classifications.	126
9.1	Répartition des fenêtres temporelles dans la base d'apprentissage	168
9.2	Matrice de confusion pour un noyau polynômial avec $d = 1$. Pour l'ensemble des matrices de confusion données, les éléments en gras font ressortir les classifications correctes (diagonale), et les chiffres entre parenthèse représentent le nombre d'éléments classés et qui représentent le pourcentage donné	169
9.3	Matrice de confusion pour un noyau polynômial avec $d = 2$	169
9.4	Matrice de confusion pour un noyau gaussien avec $\sigma^2 = 28$	171
9.5	Comparaison des taux de bonne classification pour chacune des classes avec un noyau polynômial et un noyau gaussien	172

Annexes

Grilles AGGIR pour l'admission en établissement de soins de longue durée

BILAN D'AUTONOMIE

A = besoin d'aucune aide, B = aide partielle, C = aide totale

		A	B	C
COHERENCE (conserver et/ou se comporter de façon sensée)				
ORIENTATION (se repérer dans le temps, les moments de la journée et dans les lieux)				C
TOILETTE (concerne l'hygiène corporelle)	Haut			C
	Bas			C
HABILLAGE (s'habiller, se déshabiller)	Haut			C
	Idoyen			C
	Bas			C
ALIMENTATION (manger des aliments préparés)	Se servir			C
	Manger			C
ELIMINATION (assumer l'hygiène de l'élimination urinaire et fécale)	Urinaire			C
	Fécale			C
TRANSFERTS (se lever, se coucher, s'asseoir)		n		
DEPLACEMENT A L'INTERIEUR (avec ou sans canne, déambulateur)		n		
DEPLACEMENT A L'EXTERIEUR (à partir de la porte d'entrée sans moyen de transport)			B	
COMMUNICATION A DISTANCE (utiliser les moyens de communication téléphonique, sonnette, alarme)				C
G.I.R. :				

EVOLUTION PREVISIBLE DE L'ETAT DE SANTE :

Amélioration ☐ Stabilisation ☐ Aggravation ☒

SOINS INFIRMIERS :

	OUI	NON
Pertusion		/
Injection		/
Pansements		/
Gastrostomie ou S.N.G		/
Trachéotomie		/
Colostomie		/
Sonde urinaire		/
Prévention d'escarres		/
Soins d'escarres		/
Aide à la prise de médicaments	/	/

TYPE D'ALIMENTATION :

	OUI	NON
Normal	/	
Mixé		
Régime (à préciser) -----		

RETENTISSEMENT FONCTIONNEL DE SES DEFICIENCES :

	OUI	NON
Chutes		/
Fugues	/	

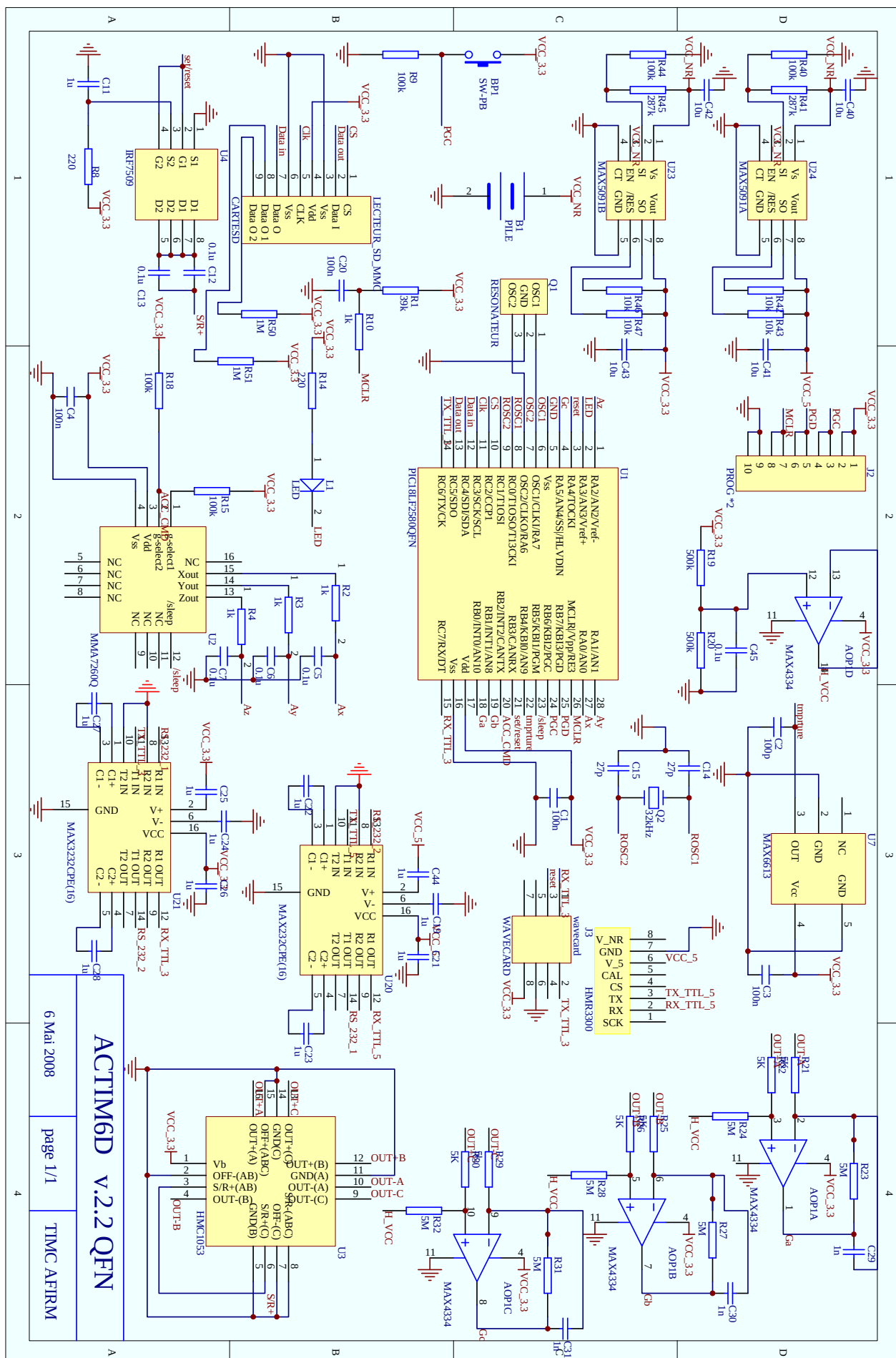
SOINS ET AIDES EXISTANTS :

	OUI	NON
Toilette	/	
Aides techniques infirmiers		/
Aides de Kinésithérapie		/

AIDES TECHNIQUES :

	OUI	NON
Canne		/
Déambulateur		/
Fauteuil roulant		/

Schéma électrique de la carte ACTIM6D

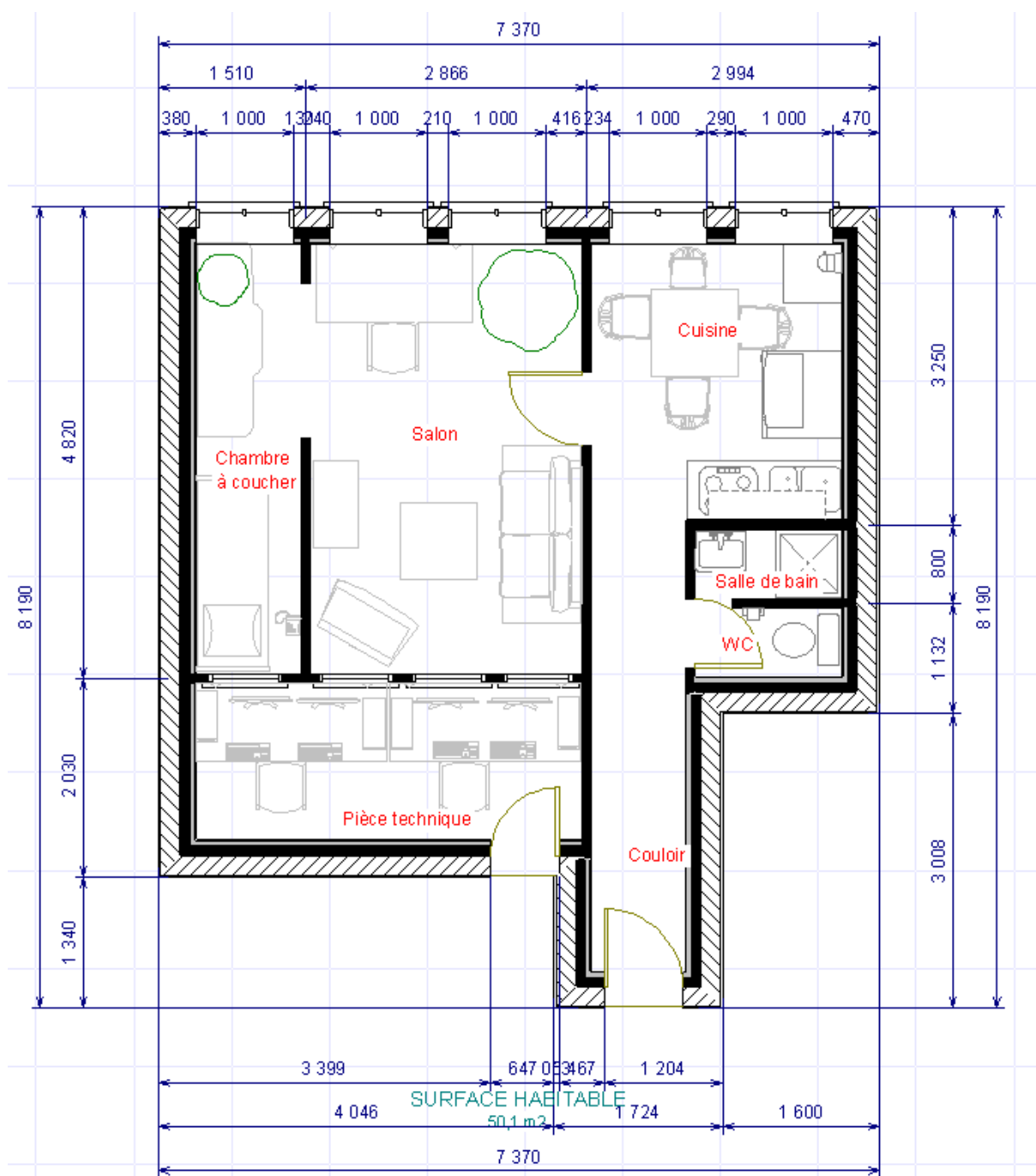


6 Mai 2008

page 1/1

TIMC AFIRM

Plan de l'appartement avec les principales cotes



Expérimentations : protocoles et lettres d'acceptation

Description des études 2007/2008

Anthony Fleury <anthony.fleury@imag.fr> – Laboratoire TIMC-IMAG, Équipe AFIRM

Dernière révision : 21 juillet 2008

Étude 1 : ACTIM6D

But de l'étude

Le but de cette étude est d'évaluer la classification des postures à partir des données accélérométriques et magnétométriques de l'ACTIM6D et de déterminer la spécificité et la sensibilité des algorithmes développés.

Les sessions devront comporter :

- Des séquences de marche
- Des transitions assis/debout
- Des transitions debout/assis
- Des transitions debout/allongé
- Des transitions allongé/debout
- Des changements d'orientation
- Des séquences de la vie quotidienne (refaire son lacet et se baisser pour ramasser un objet)

Les valeurs des magnétomètres varient selon « l'orientation » du sujet par rapport au nord magnétique. Pour vérifier la robustesse des algorithmes lorsque cette position par rapport au nord magnétique varie, les transitions assis/debout et debout/assis seront effectuées sur deux chaises disposées à angle droit l'une par rapport à l'autre.

Demande au sujet

Les sujets seront équipés de l'ACTIM6D et du capteur Freescale (ZStar). Ils auront pour consigne de réaliser le scénario demandé à leur rythme de confort.

La séance démarre et se termine avec un saut servant à la synchronisation des signaux acquis.

Matériel et emplacement

Ce scénario sera effectué dans le HIS, avec comme matériel :

- L'ordinateur servant aux acquisitions.
- ZStar de Freescale avec acquisition des accéléromètres à 100Hz sur 8 bits.
- ACTIM6D (Freescale + HMR 3300) avec acquisition des accéléromètres à 20Hz sur 10 bits et des magnétomètres à 10Hz sur 8 bits.
- Deux chaises : dures, sans accoudoir, hauteur du siège : 45 cm, angle du dossier : 10°.
- Un lit : dur, hauteur : 46 cm.

Déroulement de l'étude

Pour chaque sujet, le scénario effectué sera le suivant (sauf mention contraire, tout mouvement sera fait à une vitesse confortable pour le sujet) :

- 5 Up & Go sur la chaise 1. Chaque Up & Go seront effectués comme dans la littérature. Départ assis, marche jusqu'au mur de la cuisine (4 mètres), demi-tour, retour vers la chaise, demi-tour autour de cette chaise et retour assis avec une pause de quelques secondes. À la fin du dernier Up & Go, le sujet se relève et reste immobile quelques secondes.
- Une série de 5 allers-retours, départ assis de la chaise numéro 2, marche dans le HIS jusqu'à la porte d'entrée, demi-tour, retour vers le salon du HIS puis s'assoit de nouveau sur la chaise numéro 2. Après la cinquième réalisation le sujet se relève et attend quelques secondes.
- Départ ensuite à côté de la commode de la chambre, le sujet va s'allonger, reste dans cette position sans bouger quelques secondes, bouge pendant quelques secondes en restant allongé (passage de allongé sur le dos à allongé sur le bras gauche puis sur le dos puis sur le bras droit puis sur le dos) puis se relève et retourne à son endroit de départ. Ce scénario sera répété également 5 fois.
- Debout au milieu du salon, il est demandé au sujet de refaire ses deux lacets puis de ramasser un objet posé au sol à côté de la chaise.
- Optionnellement, pour les personnes acceptant, deux chutes (une avant, une arrière) sur le matelas installé dans le HIS.

Étude 2 : Validation du système sonore

But de l'étude

Le but de cette étude est de valider, dans les conditions offertes par l'Habitat Intelligent pour la Santé (HIS) installé au bâtiment Jean Roget, du laboratoire TIMC-IMAG, les résultats obtenus par le LIG sur son système de reconnaissance de sons et de paroles.

Pour ceci, nous utiliserons le système d'acquisition sonore du LIG avec son système de reconnaissance de parole. La version utilisée est la dernière installée, version du 19/11/2007 sous Debian 4.0.

Matériel et emplacement

L'étude se déroule dans l'Habitat Intelligent pour la Santé de la faculté de médecine de Grenoble. Les données sont enregistrées par « chipster.imag.fr » (Régie, salle 333). La session expérimentale sera également entièrement enregistrée par 5 caméras (avec leur micro activé pour avoir une trace sonore) disposées dans les pièces du HIS. Le sujet devra être seul dans le HIS. Les consignes lui seront données avant le début de l'enregistrement sonore.

Déroulement de l'étude

Le sujet devra reproduire des sons de la vie quotidienne, sans contrainte d'ordre.

Dans un premier temps il rentrera dans l'habitat intelligent, ouvrira la porte fermée à clé puis la refermera derrière lui. À la fin de l'expérimentation il lui sera demandé de ressortir en refermant la porte à clé.

Il lui sera demandé d'effectuer chacune des actions suivantes 3 fois (non nécessairement consécutives) :

- Faire couler l'eau du robinet de la salle de bain pendant dix secondes puis tirer la chasse d'eau des toilettes.
- Fermer la porte des toilettes/salle de bain du HIS.
- Marcher de la porte d'entrée au lit puis du lit à la porte d'entrée.
- Répondre au téléphone (qui sonnera durant l'expérimentation). Pour effectuer cette étape de l'expérimentation, restez assis quelques secondes sur une chaise de la cuisine et le téléphone sonnera. Une fois répondu vous raccrochez aussitôt.
- Faire tomber un livre disposé sur la table de la cuisine.
- Faire couler l'eau du robinet de la cuisine pendant dix secondes.
- Faire bouger une cuillère dans une tasse durant 5 secondes (la tasse et la cuillère seront sur la table de la cuisine également).
- Crier (Aïe).

À la suite de tout ceci, le sujet devra énoncer, avec son débit et son volume de voix normal, trois phrases de chacun des corpus fournis avant le début de la session (S0, S1, S4 et S5).

Étude 2 : Validation du système sonore v2

But de l'étude

Le but de cette étude est de valider, dans les conditions offertes par l'Habitat Intelligent pour la Santé (HIS) installé au bâtiment Jean Roget, du laboratoire TIMC-IMAG, les résultats obtenus par le LIG sur son système de reconnaissance de sons et de paroles.

Pour ceci, nous utiliserons le système d'acquisition sonore du LIG avec son système de reconnaissance de parole. La version utilisée est la dernière installée, version du 19/11/2007 sous Debian 4.0.

Matériel et emplacement

L'étude se déroule dans l'Habitat Intelligent pour la Santé de la faculté de médecine de Grenoble. Les données sont enregistrées par « chipster.imag.fr » (Régie, salle 333). La session expérimentale sera également entièrement enregistrée par 5 caméras (avec leur micro activé pour avoir une trace sonore) disposées dans les pièces du HIS. Le sujet devra être seul dans le HIS. Les consignes lui seront données avant le début de l'enregistrement sonore.

Déroulement de l'étude

Le sujet devra reproduire des sons de la vie quotidienne, sans contrainte d'ordre.

Dans un premier temps il rentrera dans l'habitat intelligent, ouvrira la porte fermée à clé puis la refermera derrière lui. À la fin de l'expérimentation il lui sera demandé de ressortir en refermant la porte à clé.

Il lui sera demandé d'effectuer chacune des actions suivantes 3 fois (non nécessairement consécutives) :

- Fermer la porte des toilettes/salle de bain du HIS.
- Répondre au téléphone (qui sonnera durant l'expérimentation). Pour effectuer cette étape de l'expérimentation, restez assis quelques secondes sur une chaise du salon et le téléphone sonnera. Une fois répondu vous direz "Allo" puis vous répéterez les phrases dites par votre interlocuteur jusqu'à ce qu'il vous demande de raccrocher.
- Faire tomber un livre disposé sur la table de la cuisine.
- Faire bouger une cuillère dans une tasse durant 5 secondes (la tasse et la cuillère seront sur la table de la cuisine également).
- Crier (Aïe).

À la suite de tout ceci, le sujet devra énoncer, avec son débit et son volume de voix normal, dix phrases de chacun des corpus fournis avant le début de la session (S0, S1, S4 et S5).

Étude 3 : Fusion de données

But de l'étude

Le but de cette étude est de caractériser les différentes activités de la vie quotidienne d'une personne, d'en calculer des modèles en vue de la classification de nouvelles données, et de calculer la sensibilité et la spécificité des algorithmes développés.

Demande au sujet

Le sujet effectuera avant le début de la session une visite complète de l'appartement intelligent du bâtiment Jean Roget, et ce afin de connaître cet appartement pour ne pas fausser l'étude.

Ensuite, il lui sera demandé de réaliser plusieurs activités de la vie quotidienne, sans limitation de temps ni aucune contrainte sur la manière de l'effectuer, l'ordre ou autre.

Matériel et emplacement

L'étude se déroule dans l'Habitat Intelligent pour la Santé de la faculté de médecine de Grenoble. Le matériel mis en place est :

- L'appartement en lui-même, équipé des principales commodités.
- Des capteurs de présence infra-rouge, avec le logiciel d'acquisition dans la régie, sur « tourtisseau.imag.fr ».
- Cinq caméras de surveillance, réservés à l'apprentissage, sur « tourtisseau.imag.fr » et « kaamelot.imag.fr ».
- Des capteurs d'ouverture/fermeture de porte simulés (avec usage de la vidéo).
- ACTIM6D.
- Les microphones et leur logiciel d'acquisition sur « chipster.imag.fr ».
- La station météorologique dont les données sont sur « kaamelot.imag.fr ».

Déroulement de l'étude

Le sujet est enfermé seul dans le HIS et doit effectuer certaines activités de la vie quotidienne :

- Repos. Le sujet s'allonge sur le lit et reste à « dormir » pendant un temps raisonnable (de son choix)
- Détente. Le sujet se met sur le canapé ou sur une chaise dans le salon et effectue l'activité qui lui plaît lorsqu'il se détend (regarder la TV, lire, ne rien faire...)
- Se nourrir. Le sujet se prépare un petit-déjeuner avec le matériel et les ingrédients présents dans les placards et la cuisine du HIS puis le mange (à la table de la cuisine).
- Hygiène. Le sujet se lave les mains, le visage, et feint de se laver les dents dans la salle de bain du HIS (robinet).
- Toilettes. Le sujet feint de se rendre aux toilettes (s'assoit sur la cuvette, tire la chasse d'eau...)
- S'habiller. Le sujet s'habille et se déshabille (avant d'aller dormir et après avoir dormi ou bien après avoir effectué les tâches d'hygiène) en prenant les vêtements présents dans l'armoire à côté du lit.
- Communication. Le sujet effectuera plusieurs conversations téléphoniques dans lesquelles il lira une conversation donnée.

L'ensemble de l'étude dure le temps qu'il faudra au sujet pour « simuler cette journée de la vie » ; le minimum étant d'environ 2h pour avoir une quantité de fenêtre temporelles suffisantes.



Étude de validation d'un système embarqué : ACTIM6D

Laboratoire TIMC-IMAG, Équipe AFIRM

Informations :

Nom : Prénom : ID :
Âge : Poids :
Taille : Date/heure :

Description de l'étude

Cette étude est menée par le laboratoire TIMC-IMAG, équipe AFIRM, par Anthony FLEURY (doctorant), sous la direction de **Norbert NOURY**, responsable d'équipe.

Dans cette étude, nous cherchons à valider un système de mesure composé de plusieurs capteurs et nous permettant de détecter la « posture » (assis, debout ou allongé) d'une personne, et de caractériser cette posture (immobile, dynamique, marche, ...).

Pour cette étude, vous serez équipé(e) d'un tee-shirt comportant deux types de capteurs : deux accéléromètres et un magnétomètre. Les accéléromètres détectent l'accélération de la pesanteur ainsi que la force que vous mettez à effectuer un mouvement. Les magnétomètres détectent votre position par rapport au champ magnétique terrestre (champ fixe donnant le nord). Cet équipement enregistrera toutes ces variables au cours du temps. Il est disposé dans une poche positionnée sous l'aisselle gauche.

L'étude se déroulera à la faculté de médecine de Grenoble, au troisième étage, dans un appartement mis en place par le laboratoire TIMC-IMAG. Dans un but de vérification et d'indexation, cinq caméras enregistreront la séance. Ces enregistrements vidéos, comme toutes les données recueillies lors de l'étude, resteront confidentiels.

Cette étude est effectuée sur des personnes majeures et consentantes, qui auront pris connaissance et compris dans son intégralité cette fiche explicative et l'auront signée en présence de l'expérimentateur. La participation à celle-ci est volontaire et purement bénévole.

Vous serez guidé(e) tout au long de l'étude par l'expérimentateur (Anthony FLEURY). Il vous sera demandé d'effectuer, à votre rythme de confort et sans contrainte temporelle, les scénarios suivants :

- Cinq Up & Go, mouvement qui consiste à partir assis, se lever, marcher puis revenir vers la chaise, en faire le tour et s'asseoir de nouveau sur celle-ci.
- Départ assis sur une seconde chaise, séquence de marche avec changement de direction, puis retour sur la chaise en position assise. Ce scénario sera également répété cinq fois.
- Départ debout, vous irez vous allonger avant de vous relever et de revenir à l'endroit initial. Ce scénario sera également répété cinq fois.
- Deux mouvements de la vie quotidienne : refaire vos deux lacets puis ramasser un objet au sol.
- Si vous l'acceptez, deux chutes sur un matelas en mousse.

Je soussigné(e) Melle/Mme/M., reconnaît avoir pris connaissance de l'intégralité de ces informations sur le protocole expérimental présenté par M. Anthony FLEURY.

Je suis informé(e) que ma participation n'est en aucun cas contrainte et que je suis libre de quitter cette session expérimentale lorsque je le souhaite. Je suis également informé(e) que toutes les données recueillies dans le cadre de cette étude sont confidentielles et ne pourront être utilisées ou consultées que par l'équipe de recherche dans une finalité de recherche à l'exclusion de toute exploitation à des fins commerciales.

Toutes les données de l'étude sont soumises à la loi informatique et libertés.

Je reconnais par la présente donner mon consentement libre et éclairé en présence de l'expérimentateur.

Fait à Grenoble, le :

Signature de l'expérimentateur :

Signature de l'intéressé :



Étude de validation d'un système de reconnaissance de parole et de son

Laboratoire TIMC-IMAG, Équipe AFIRM et Laboratoire LIG, Équipe GETALP

Informations :

Nom :	Prénom :	ID :
Âge :	Poids :	
Taille :	Date/heure :	

Description de l'étude

Cette étude est menée par le laboratoire TIMC-IMAG, équipe AFIRM, par Anthony FLEURY (doctorant), sous la direction de **Norbert NOURY** du laboratoire TIMC-IMAG, responsable d'équipe et de **Michel VACHER**, du laboratoire LIG, équipe GETALP.

Dans cette étude, nous cherchons à valider un système de classification de sons de la vie quotidienne et de reconnaissance de la parole.

Pour cette étude, vous serez dans un appartement situé à la faculté de médecine de Grenoble, dans le bâtiment Jean Roget, équipé d'un ensemble de capteurs dont des caméras de surveillance et des microphones. Le système sonore enregistrera l'ensemble des sons de l'appartement et les stockera sur un ordinateur situé dans la régie, et y stockera également les décisions prises pour chaque événement sonore. Les caméras enregistreront toute la séance dans un but d'indexation et de vérification. Ces enregistrements vidéos et sonores recueillis lors de l'étude resteront confidentiels. Cette étude est effectuée sur des personnes majeures et consentantes, qui auront pris connaissance et compris dans son intégralité cette fiche explicative et l'auront signée en présence de l'expérimentateur. La participation à celle-ci est volontaire et purement bénévole.

Vous partirez de l'extérieur de l'appartement, vous ouvrirez la porte avec la clé fournie puis vous vous enfermerez avec le verrou intérieur. À la fin de l'étude, vous ressortirez du HIS que vous refermerez à clé. Durant cette étude, vous allez devoir effectuer les actions suivantes, dans l'ordre que vous voulez, trois fois chacune (non nécessairement consécutives) :

- Faire couler l'eau du robinet de la salle de bain pendant dix secondes puis tirer la chasse d'eau des toilettes.
- Fermer la porte des toilettes/salle de bain du HIS.
- Marcher de la porte d'entrée au lit puis du lit à la porte d'entrée.
- Répondre au téléphone (qui sonnera durant l'expérimentation). Pour effectuer cette action, restez assis quelques secondes sur une chaise de la cuisine et le téléphone sonnera. Après avoir répondu vous raccrochez aussitôt.
- Faire tomber un livre disposé sur la table de la cuisine.
- Faire couler l'eau du robinet de la cuisine pendant dix secondes.
- Faire bouger une cuillère dans une tasse durant 5 secondes (la tasse et la cuillère seront sur la table de la cuisine également).
- Crier (Ahhh).

Il vous est également fourni 4 feuilles, numérotées S0, S1, S4 et S5. Dans chacune de ces quatre feuilles, vous devrez choisir trois phrases que vous devrez dire une fois chacune.

Je soussigné(e) Melle/Mme/M., reconnaît avoir pris connaissance de l'intégralité de ces informations sur le protocole expérimental présenté par M. Anthony FLEURY.

Je suis informé(e) que ma participation n'est en aucun cas contrainte et que je suis libre de quitter cette session expérimentale lorsque je le souhaite. Je suis également informé(e) que toutes les données recueillies dans le cadre de cette étude sont confidentielles. J'accepte que ces données soient conservées par les équipes de recherche dans les meilleures conditions de sécurité, et que ces données soient utilisés par les deux équipes de recherches dans une finalité de recherche à l'exclusion de toute exploitation à des fins commerciales. Toutes les données de l'étude sont soumises à la loi informatique et libertés.

Je reconnais par la présente donner mon consentement libre et éclairé en présence de l'expérimentateur.

Fait à Grenoble, le :

Signature de l'expérimentateur :

Signature de l'intéressé :



Étude sur la fusion de données dans un HIS.

Laboratoire TIMC-IMAG, Équipe AFIRM

Informations :

Nom :

Âge :

Poids :

Prénom :

Taille :

Date/heure :

Description de l'étude

Cette étude est menée par le laboratoire TIMC-IMAG, équipe AFIRM, par Anthony FLEURY (doctorant), sous la direction de **Norbert NOURY** du laboratoire TIMC-IMAG, responsable d'équipe et de **Michel VACHER**, du laboratoire LIG, équipe GETALP.

Nous cherchons à valider la reconnaissance d'activités de la vie quotidienne dans un appartement équipés de plusieurs types de capteurs. L'étude se déroulera à la faculté de médecine de Grenoble, au troisième étage, dans l'appartement mis en place par le laboratoire TIMC-IMAG. Dans celui-ci, que vous visiterez au préalable pour vous y familiariser et agir ainsi comme si il s'agissait du vôtre, vous trouverez l'ensemble des équipements nécessaires à la vie quotidienne. Plusieurs capteurs y ont également été installés : des capteurs de présence infra-rouge, équivalent à ceux des alarmes, qui détecte un mouvement dans une zone donnée ; des contacteurs de porte sur certains équipements de la cuisine et de la chambre ; une station météorologique dans la salle de bain et vous porterez également un tee-shirt équipé d'un capteur cinématique (responsable d'enregistrer vos mouvements). Dans le but d'utiliser des algorithmes à apprentissage, des caméras vidéos ont été placées pour surveiller vos actions au sein de l'appartement. Ces enregistrements vidéos, comme toutes les données recueillies lors de l'étude resteront confidentielles.

Cette étude est effectuée sur des personnes majeures et consentantes, qui auront pris connaissance et compris dans son intégralité cette fiche explicative et l'auront signée en présence de l'expérimentateur. La participation à celle-ci est purement bénévole.

Vous entrerez dans l'appartement et à partir de ce moment, vous y serez seul(e). Vous ne ressortirez qu'à la fin de la session expérimentale. Vous aurez à effectuer les tâches suivantes de la vie quotidienne, dans l'ordre que vous voulez (en commençant par le réveil étant donné que nous simulons une matinée) :

- Repos. Vous vous allongez sur le lit et restez à « dormir » pendant un temps raisonnable (de votre choix)
- Détente. Vous vous installez sur le canapé ou sur une chaise dans le salon et effectuez l'activité qui vous plaît lorsque vous vous détendez (regarder la TV, lire, ne rien faire...)
- Se nourrir. Vous vous préparez un petit-déjeuner avec le matériel et les ingrédients présents dans les placards et la cuisine du HIS puis le mangez (à la table de la cuisine). Vous laverez ensuite votre vaisselle dans le robinet de la cuisine.
- Hygiène. Vous vous lavez les mains, le visage, et feignez de vous laver les dents dans la salle de bain du HIS (robinet).
- Toilettés. Vous feignez de vous rendre aux toilettes (s'asseoir sur la cuvette, tirer la chasse d'eau...)
- S'habiller. Vous vous habillez et vous déshabillez (avant d'aller dormir et après avoir dormi ou bien après avoir effectué les tâches d'hygiène) en prenant les vêtements présents dans l'armoire à côté du lit.
- Communication. Vous lirez plusieurs conversations téléphoniques données lorsque vous serez appelé.

Il vous sera demandé, dans la mesure du possible, que chaque activité dure au moins **3 minutes consécutives** afin qu'elle soit utilisable.

Je soussigné(e) Melle/Mme/M., reconnaît avoir pris connaissance de l'intégralité de ces informations sur le protocole expérimental présenté par M. Anthony FLEURY.

Je suis informé(e) que ma participation n'est en aucun cas contrainte et que je suis libre de quitter cette session expérimentale lorsque je le souhaite. Je suis également informé(e) que toutes les données recueillies dans le cadre de cette étude sont confidentielles et ne pourront être utilisées ou consultées que par l'expérimentateur ou par une personne habilitée ou ayant reçu mon accord. Toutes les données de l'étude sont soumises à la loi informatique et libertés.

Je reconnais par la présente donner mon consentement libre et éclairé en présence de l'expérimentateur.

Fait à Grenoble, le :

Signature de l'expérimentateur :

Signature de l'intéressé :

Ensemble des corpus de paroles : anodins, détresse et téléphone

Corpus de paroles pour l'étude de validation du système sonore (1/4)

S0 : type détresse

Aidez moi !
Aouh !
Appelez le Samu !
Appelez les pompiers !
Appelez les secours !
Appelez un docteur !
Appelez un toubib !
Appelez une ambulance !
Appelez une infirmière !
Appelez quelqu'un !
Au secours !
Aïe !
Aïe aïe aïe !
Docteur !
Du secours s'il vous plait !
Faîtes vite !
Help !
Infirmière !
J'ai besoin d'aide !
J'ai besoin d'un docteur !

Corpus de paroles pour l'étude de validation du système sonore (2/4)

S1 : type détresse

J'ai besoin d'une infirmière !
J'ai besoin d'un toubib !
J'ai besoin de secours !
J'ai la tête qui tourne !
J'ai mal !
J'ai mal à la tête !
J'ai très mal !
J'ai un malaise.
Je me sens mal.
Je me sens très mal
Je ne me sens pas très bien.
Je ne peux plus bouger.
Je ne suis pas bien.
Je suis blessé.
Je suis tombé.
Ne me laissez pas tout seul !
Ne me laissez pas toute seule !

Corpus de paroles pour l'étude de validation du système sonore (3/4)

S4 : type anodin

J'ai lu le journal.
J'ai ouvert la porte.
J'ai sommeil.
J'allume la lumière.
Je dois prendre mon médicament.
Je me suis réveillé tôt.
Je me suis endormi tout de suite.
Je n'entends rien.
J'ai faim.
J'ai fermé la porte.
J'ai froid.
J'ai soif.
J'écoute la radio.
Je ferme la fenêtre.
Je ferme la porte.
Je me suis endormi tard.
Je n'ai pas faim.
Je n'ai plus faim.
Je n'ai vu personne.
J'ai éteint la lumière.

Corpus de paroles pour l'étude de validation du système sonore (4/4)

S5 : type anodin

J'ouvre la porte.
La chaise est tombée.
La lumière est éteinte.
La porte est fermée.
La porte est ouverte.
Le café est brûlant.
Les pâtes sont cuites.
Le sucre est sur la table.
Où est la boîte de cachets ?
J'ouvre la fenêtre.
Je vais bien.
Je vais fermer la porte.
Où est mon verre.
Où est le sel.
Où est le sucre.
Le médecin est venu hier.
L'infirmière va venir.
Où sont mes lunettes.
Quel temps fait-il dehors ?
Je vais faire la vaisselle.

Téléphone

A ce soir
À demain
A dimanche
A jeudi
Allo
Allo-allo
Allo c'est bien moi
Allo c'est ça
Allo c'est moi
Allo c'est qui
Allo heu
Allo non
Allo oui
Allo qui c'est
A lundi
A mardi
A mercredi
A midi
A vendredi
A samedi
Bonjour
Bonjour Madame
Bonjour mademoiselle
Bonjour Monsieur
Ça va
Ca va bien
Ça va bien
Ça va pas
Ca va pas bien
Ca va pas bien du tout
Ca va pas du tout
Ça va pas du tout
Ca va très bien
C'est bien
C'est ça
C'est d'accord
C'est entendu
C'est pas ça
C'est qui à l'appareil
Comment
Comment allez vous
Comment ça va
C'était hier
D'accord
Je n'entend rien

Je n'entends rien
Non
Non ça va
Non ça va pas
Non c'est pas ça
Non merci
Non merci bien
Oui
Oui ça va
Oui c'est ça
Parlez plus fort
Pour l'instant ça va
Qui c'est
Qui est à l'appareil
Tout va bien

Publications

A Fast Algorithm to Track Changes of Direction of a Person Using Magnetometers

A. Fleury, *Student Member, IEEE*, N. Noury, *Senior Member, IEEE*, N. Vuillerme

Abstract—Technological advances in signal processing and in circuits integration offer numerous perspectives in telemedicine and telemonitoring. Considering the increase of life expectancy, accurate and reliable assessment of modification and/or deterioration in the health status of a person is needed. One possible indicator is the “activity index” of a person. To compute such an index, previous studies have used accelerometers. Although these sensors are appropriate for the detection of postural transitions (e.g. Sit To Stand and Stand To Sit), they do not allow to detect changes of direction of a walking individual insofar as such activity occurs in a constant gravitation field. Within this context, the purpose of the present work is to investigate whether magnetometers can be used to monitor the changes of direction of a person.

Index Terms—Magnetometers, embedded system, yaw angle, quaternion, clustering

I. INTRODUCTION

PERCENTAGE of population over 65 is increasing and this trend will accelerate in a near future. To reduce the time spent in hospital and to allow older adults to stay at home, smart sensors and smart house [1] can be used not only to monitor the activity, but also to detect abnormal situation in the daily life of the person. Most of these systems integrate accelerometers to detect fall [2] or to classify postural transitions (e.g. Sit To Stand and Stand To Sit) [3]. Data relative to postural transitions of a person can further be used in a data fusion engine to detect and characterize daily activity of the person and to compute his/her “activity index”. At this point, however, accelerometers, by definition, do not allow to detect changes of direction of a walking individual insofar as such activity occurs in a constant gravitation field. Other sensors are thus needed. Interestingly, [4], [5] used a central consisting of accelerometers, magnetometers and gyroscopes to track movements of a limb. In these studies, however, magnetometers are used to compensate the integration drift of the gyroscopes, only. Within this context, the purpose of the present work is to investigate whether magnetometers can be used to monitor the changes of direction of a person.

II. MATERIALS AND METHODS

A. Central of sensors – Delivered signals

As illustrated in Fig. 1, the central of sensors comprises 1 one-axis magnetometer (HMC1021Z, Honeywell) and 1 two-

Anthony Fleury, Norbert Noury, Nicolas Vuillerme are with the Laboratoire TIMC-IMAG, UMR CNRS/UJF 5525, équipe AFIRM. Faculté de Médecine de Grenoble, bâtiment Jean Roget, 38706 La Tronche, France (e-mail: {Anthony.Fleury,Norbert.Noury,Nicolas.Vuillerme}@imag.fr).

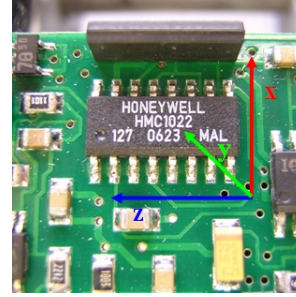


Fig. 1. 3D axis set on the circuit board

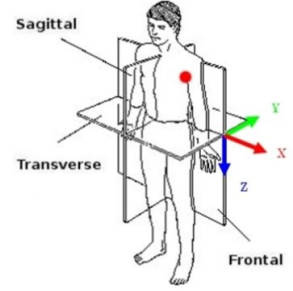


Fig. 2. 3D axis set on the subject (the red point designates location of the sensors)

axis one (HMC1022, Honeywell). This central, maintained immobile on the subject's body (Fig. 2), creates a tri-dimensional direct orthogonal basis, and allows the collection of three distinct signals for x , y and z axes, respectively. An example of collected data from these two magnetometers is given in Fig. 3-A.

B. Signal processing

To reduce the number of computations, and thus to decrease the calculus consumption of the algorithm, raw signals computed from magnetometers along x , y and z axes are first pre-processed. On the whole, the filters used for both walking and non activity time period detections, are applied on a two seconds window, allowing the use of a relatively rapid algorithm that can be executed in real time.

1) *Walk filtering*: Considering walking as a periodic signal in the bandwidth $[0.5Hz; 2Hz]$, the first process is a three-order bandstop digital filter aimed at reducing the unwanted variations in the magnetometers signal. Note that the signal is duplicated and inverted before and after (mirror effect) to remove discontinuities and hence to prevent edge effects. The second process is a low-pass filter with a $4Hz$ cutoff frequency to remove the high frequency noise. Fig. 3-B and Fig. 4-B. illustrate filtered signals computed from data collected during 250s and 40s, respectively. The 40s window, illustrated in Fig. 4-B, corresponds to a temporal frame during which two 180° “rotations” and a walking sequence were executed.

2) *Detection of non-activity periods of time*: Since numerous temporal frames contains very few movements and thus do not present any interest, we use a filter which computes

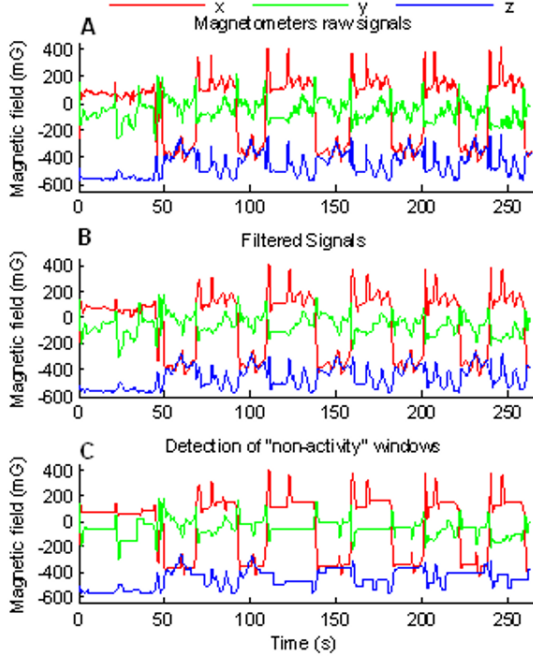


Fig. 3. Successive steps of the filtering process shown on the signal of an experimentation in which the subject walked, sat, stand and made different changes of direction

the running standard deviation of the signal. The standard deviation is defined by the following expression:

$$s_n = \sum_{i=1}^n \frac{(x_i - \bar{x}_n)^2}{n} = \frac{1}{n} \left(\sum_{i=1}^n x_i^2 - 2 \cdot \bar{x}_n \cdot \sum_{i=1}^n x_i \right) + \bar{x}_n^2 \quad (1)$$

where:

$$\bar{x}_n = \frac{1}{n} \cdot \sum_{i=1}^n x_i = \frac{n-1}{n} \cdot \bar{x}_{n-1} + \frac{1}{n} \cdot x_n \quad (2)$$

This procedure allows to reduce the number of windows to be analyzed, and hence the calculus consumption of the algorithm.

On the one hand, to maintain the continuity of information from one sampling window to another and to avoid saving all collected samples, the following seven floating point parameters have to be computed: (1) the number of samples processed (1 variable), (2) the sum of the samples (1 parameter per axis = 3 parameters), (3) the sum of the square of the samples (1 parameter per axis = 3 parameters).

On the other hand, to improve the effectiveness of this filter, before collecting magnetometers data, we initialize these seven variables in a temporal frame during which a half-turn movement is executed. This procedure allows the first windows to be defined as “activity windows” only if they really are.

The standard deviation of each window is compared to the running standard deviation. There are two possibilities : (1)

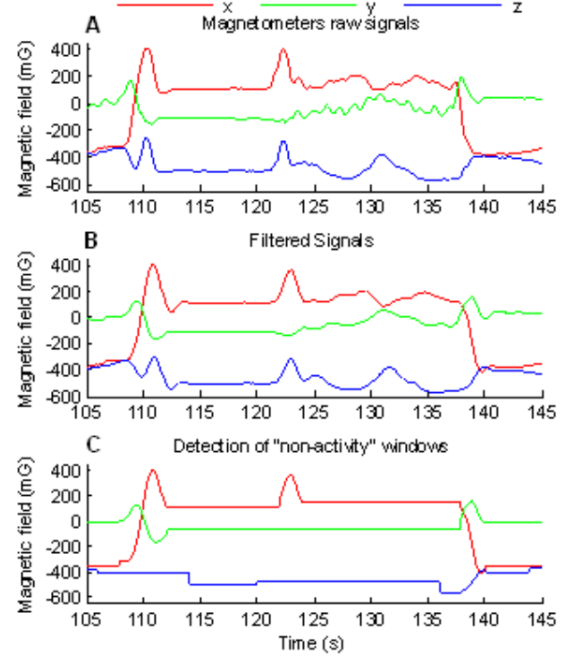


Fig. 4. Successive steps of the filtering process shown on a significant part of the signal of Fig. 3 in which the subject made a 180° rotation, walked and made a second 180° rotation in the opposite direction

if its value is small compared to the one of the signal, the window is declared as “non-activity” and the whole window signal is replaced by its mean (low pass filter that keeps only the constant component); (2) if not, the signal is kept as is (all-pass filter). This filter can be modelled as follows:

$$\forall i \in [1; N_w], x_i = \begin{cases} x_i & \text{if } s_w > \epsilon \cdot s_t \\ \bar{x}_w & \text{else} \end{cases} \quad (3)$$

with:

- N_w The size of the Window
- s_w Standard deviation of the Window
- ϵ a coefficient determined by experimentation
- $\bar{x}_w$ Mean value on the window

Results of the above-mentioned process are illustrated on Fig. 3-C and Fig. 4-C. They show that most of the windows are not processed due to their classification as “non-activity” windows. This could allow to decrease time and calculus consumption, that could be highly relevant if we aim at developing embedded systems for tracking changes of directions of a person. Note that even if this filter creates a discontinuity in the signal (see Fig. 4-C), the processe is not impaired. Indeed, when a “non-activity” window is detected, the starting point of the analysis is automatically set at the first point of the next “activity” window.

C. Movement clustering

As we can see on Fig. 4-C, change of direction of a walking subject yields a significant increase in the euclidian

norm of the difference between the two points in the plan $(\mathbf{x}; \mathbf{y})$. Within this context, the changes in direction can be detected in the signal by a significant increased in the euclidian norm of the difference vector measured during two distinct moments of the experiment. Let's define ϵ_{90} and ϵ_{180} as the two thresholds values corresponding to two rotation of 90° and 180° , respectively. Interestingly, if the movement lasts more that 2 seconds, the analysis can be pursued on the next windows.

A change of direction of a walking person is detected when this euclidian norm is superior to ϵ_{90} and ϵ_{180} values. Furthermore, the direction of this change is indicated by the orientation of the difference vector.

On the whole, this filtering procedure, requiring to only compute euclidian norm of the three $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ signals, is relatively rapid. However, this algorithm necessitates to verify the decision given by this filtering procedure by computing the real yaw angle of the signal. Indeed, one of the drawback concerns the correctness regarding to orientation of the movement. Indeed, outside the interval $[-\frac{\pi}{2}; \frac{\pi}{2}]$, the classes for “plus” or “minus” angles are inverted. Once a false decision is taken, all the following orientations are false.

D. Uncertainty minimisation

1) *Quaternion of the rotation:* To check the last decision independantly of the previous ones, we compute the real yaw angle of the transformation by using the quaternion that describes the composition of the three rotations (around $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$). A preliminary step is to normalize the values delivered by the conditioning circuit positioned behind the magnetometers. Considering that (1) this conditioning circuit delivers linearly transformed values, and (2) each individual has morphological differences, we center the data from the $(\mathbf{x}, \mathbf{y})$ plan in the interval $[-1; 1]$ for the subject, by using a first 360° movement performed at the beginning of the experiment.

The norm of the magnetic field vector is almost constant (the small variations are caused by the acquisition noise) during the experimentation session. We measure the variable projection of a fixed vector on a moving referential (attached to the body of the subject).

To compute the yaw angle of the transformation, we first determine the quaternion of the transformation. Let's consider $\mathbf{B}_1$ and $\mathbf{B}_2$ the magnetic field measured at two moments of the experimentation. We want to determine a rotation of vector $\mathbf{k} = (k_x \ k_y \ k_z)^T$ and of angle θ that transforms $\mathbf{B}_1$ in $\mathbf{B}_2$. Then, the quaternion is:

$$Q = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) k_x \\ \sin\left(\frac{\theta}{2}\right) k_y \\ \sin\left(\frac{\theta}{2}\right) k_z \end{pmatrix} = \begin{pmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{pmatrix} \quad (4)$$

One possible basis in which the rotation can be written as

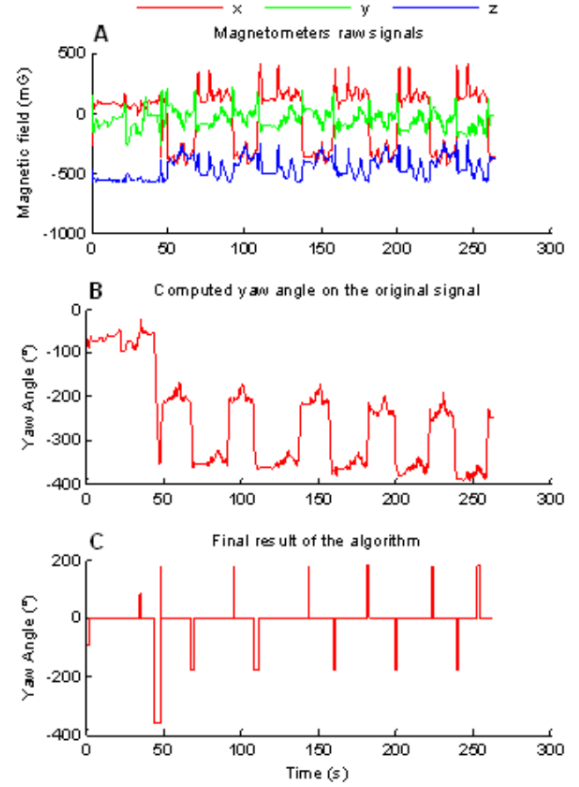


Fig. 5. Results of the execution of the algorithm on real data

a simple rotation of angle θ around one axis $\mathbf{k}$ is given by the set of three vectors:

$$\begin{cases} \mathbf{i} = \frac{\mathbf{B}_1}{\|\mathbf{B}_1\|} \\ \mathbf{j} = \frac{\|\mathbf{B}_1\|^2 \mathbf{B}_2 - (\mathbf{B}_1 \cdot \mathbf{B}_2) \mathbf{B}_1}{\|\|\mathbf{B}_1\|^2 \mathbf{B}_2 - (\mathbf{B}_1 \cdot \mathbf{B}_2) \mathbf{B}_1\|} \\ \mathbf{k} = \frac{\mathbf{B}_1 \wedge \mathbf{B}_2}{\|\mathbf{B}_1 \wedge \mathbf{B}_2\|} \end{cases} \quad (5)$$

Let $\mathbf{b}_1$ be the projection of $\mathbf{B}_1$ in the basis given by Eq. 5 and $\mathbf{b}_2$ the projection of $\mathbf{B}_2$ in the same basis. The angle of the transformation is given by the expression:

$$\theta = \arccos\left(\frac{\mathbf{b}_1 \cdot \mathbf{b}_2}{\|\mathbf{b}_1\| \cdot \|\mathbf{b}_2\|}\right) \quad (6)$$

With Eq. 5 and Eq. 6, we calculated the main vector and the angle of this rotation. We have now a numerical expression of the quaternion Q given by Eq. 4.

From this quaternion, we can write the rotation matrix in the basis $(\mathbf{x}, \mathbf{y}, \mathbf{z})$. The expression of this matrix $M_{r_{xyz}}$ is:

$$\begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(-q_0 q_3 + q_1 q_2) & 2(q_0 q_2 + q_1 q_3) \\ 2(q_0 q_3 + q_2 q_1) & q_0^2 - q_1^2 - q_2^2 - q_3^2 & 2(-q_0 q_1 + q_2 q_3) \\ 2(-q_0 q_2 + q_3 q_1) & 2(q_0 q_1 + q_3 q_2) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{pmatrix} \quad (7)$$

2) *Roll, pitch and yaw angle*: Assuming that our movement is made of three rotations, one around the $\mathbf{z}$ axis (yaw), of angle ψ , the second of angle ϕ around $\mathbf{y}$, called pitch and the third, the roll angle, θ , around $\mathbf{x}$, we compute the yaw angle of the rotation.

Let M_ψ be the matrix of rotation around $\mathbf{z}$, M_ϕ around $\mathbf{y}$ and M_θ around $\mathbf{x}$. To compute the roll, pitch and yaw angles, we can identify the product of these three matrix, $M'_{xyz} = M_\psi M_\phi M_\theta$, with the matrix given by Eq. 7.

We obtain, with the same notations as previously :

$$\begin{cases} \psi = \arctan\left(\frac{M_{xyz}(1,2)}{M_{xyz}(1,1)}\right) \\ \theta = \arcsin\left(M_{xyz}(1,3)\right) \\ \phi = \arctan\left(\frac{M_{xyz}(2,3)}{M_{xyz}(3,3)}\right) \end{cases} \quad (8)$$

Since our application concerns the tracking of changes of direction of a person, only ψ is of interest and will be computed for the period of time that we are currently checking only. To avoid any ambiguity, the computed yaw angle for the whole signal of Fig. 5-A is illustrated in Fig. 5-B.

E. Final output of the algorithm

After the euclidian norm having being computed and the decision having being verified by the quaternion, the executed movement has been detected accurately. This decision allows the calculation of final output of the algorithm, as follows. As illustrated in Fig. 5-C, for the temporal frame during which the movement takes place, the final output of the algorithm is the value corresponding to the angle induced by the executed rotation. this output gives us two main information: (1) the angle of the modification of direction and (2) the time spent to perform this rotation.

III. RESULTS

Eight young healthy adults (mean age : 27 ± 3) voluntarily participated to two experiments. Before these experiments, each subject stood upright and were asked to perform a 360° rotation followed by a 180° rotation to calibrate the algorithm. In Experiment 1, they were asked to walk in a corridor, at a comfortable speed, to make multiples changes of direction (ten 180° changes, five in each direction of rotation) and to sit on a chair between these changes. In Experiment 2, they were asked to walk, to go up and down the stairs and to make four 180° changes (two in each direction) and eight 90° changes (four in each direction). Results showed that our algorithm was able to accurately detect and classify changes of direction executed by each subject during each experiment.

IV. DISCUSSION AND CONCLUSION

We have demonstrated the effectiveness of using a tri-axis magnetometer in accurately tracking the changes of directions of a person. Our algorithm, analysing the magnetometer

signals by 2 seconds windows, is relatively rapid and can be executed in real time.

On the one hand, since the effectiveness of this algorithm mainly stem from the determination of the two values of ϵ_{90} and ϵ_{180} , pursuing investigations to increase their optimality is necessary.

On the other hand, it is important to mention that our experiments were not conducted in “ideal” situations, i.e. without any magnetic disturbances. On the contrary, both experimental sessions required subjects to walk near TV, fridge, ... representing similar conditions to those occurring at home. What is more, we further assessed the electro-magnetic sensitivity. To achieve this goal, we positioned the magnetometer 20 cm away from a TV for 30 minutes and we compared the norm of the magnetic field acquired and the standard deviation of this norm obtained when the TV was on to those obtained when the TV was off. Results showed similar magnetic field in these two conditions (mean measured magnetic field of $528.9 \text{ mG} \pm 1.94$ and of $528.7 \text{ mG} \pm 1.65$, for TV off and TV on, respectively). Similar results were observed with fridge, microwave, fan and CRT display. On the whole, results showed that neither the TV, fridge, microwave, fan nor CRT display have any notable effect on the magnetometer acquisition circuit. Nevertheless, mean measured magnetic field obtained with the magnetometer positioned close to the elevator is significantly different to that obtained with the magnetometer positioned 1 m away from the elevator ($2193.64 \text{ mG} \pm 303.13$ versus $565.19 \text{ mG} \pm 10.21$). This result suggests that being very closed to an elevator, or more largely to a direct current engine, induces a large modification of the acquired magnetic field, that can be detected by our algorithm, but that limits the use of the magnetometers for tracking changes of direction of a person.

ACKNOWLEDGMENT

The authors would like to thanks Damien Flammarion and Jérémy Groussier for their help on the design of the sensors circuit and all the subjects from the TIMC-IMAG laboratory for their time accorded to these experimentations.

REFERENCES

- [1] C. N. Scanail, S. Carew, P. Barralon, N. Noury, D. Lyons, and G. M. Lyons, “A review of approaches to mobility telemonitoring of the elderly in their living environment,” *Annals of Biomedical Engineering*, vol. 34, pp. 547 – 563, April 2006.
- [2] A. K. Bourke, J. V. O’Brien, and G. M. Lyons, “Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm,” *Gait Posture*, vol. 26, no. 2, pp. 194–199, Jul 2007.
- [3] P. Barralon, N. Noury, and N. Vuillerme, “Classification of daily physical activities from a single kinematic sensor,” in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, 2005, pp. 2447–2450.
- [4] R. B. McGhee, E. R. Bachmann, X. Yun, and M. J. Zyda, “An investigation of alternative algorithms for singularity - free estimation of rigid body orientation from earth gravity and magnetic field measurements,” Naval Postgraduate School, Monterey, California, Tech. Rep., 2001.
- [5] R. Zhu and Z. Zhou, “A real-time articulated human motion tracking using tri-axis inertial/magnetic sensors package,” *IEEE Trans. Neural Syst. Rehab. Eng.*, vol. 12, no. 2, pp. 295 – 301, June 2004.

Sound and Speech Detection and Classification in a Health Smart Home

A. Fleury, *Student Member, IEEE*, N. Noury, *Senior Member, IEEE*, M. Vacher, H. Glasson and J.-F. Serignat

Abstract—Improvements in medicine increase life expectancy in the world and create a new bottleneck at the entrance of specialized and equipped institutions. To allow elderly people to stay at home, researchers work on ways to monitor them in their own environment, with non-invasive sensors. To meet this goal, smart homes, equipped with lots of sensors, deliver information on the activities of the person and can help detect distress situations. In this paper, we present a global speech and sound recognition system that can be set-up in a flat. We placed eight microphones in the Health Smart Home of Grenoble (a real living flat of $47m^2$) and we automatically analyze and sort out the different sounds recorded in the flat and the speech uttered (to detect normal or distress french sentences). We introduce the methods for the sound and speech recognition, the post-processing of the data and finally the experimental results obtained in real conditions in the flat.

Index Terms—Sound recognition, Speech recognition, Health Smart Home.

I. INTRODUCTION

DEPENDANCY of elderly persons become an important social problem. Indeed, in France nowadays, 1.3 millions people are over 85 and, in 2015, they will be 2 millions. Loss of autonomy concerns today about 2 millions of people (half are elderly and half are handicapped) and threatens eventually a quarter of the population of elderly people.

For these reasons, geriatrics ask the researchers for tools to automatically detect the decrease in autonomy, so that they can plan the best moment to accept the person in a specialized institution – not too early and not in a hurry. The gain of time before the entrance is one of the solutions to regulate the lack of places in institutions. It is also a chance for the person to live longer in their own environment.

Smart sensors and smart homes have proven their efficiency to give information on the patient. Sensors can deliver information on postures and movements of the person [1], [2] or detect a fall [3], [4]. Smart Homes are used to measure the activity of the person [5], [6], or to help people (with cognitive impairments for instance) in their activities [7]. Moreover, few projects work on sound and

speech recognition systems in smart homes, for automation purpose for instance [8].

This paper describes the implementation and results on the use of microphones for sound classification and speech recognition (in French) in the Health Smart Home of Grenoble. These additional information could be used for two purposes:

- Detect distress situations in the flat by analysing the sounds and also by recognizing a distress sentence,
- By fusing this information with other ones in the flat, we could deduce the activity of daily living actually performed.

II. MATERIALS

A. The Grenoble Health Smart Home

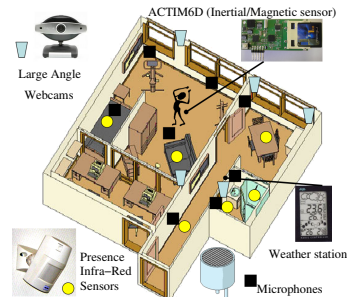


Fig. 1. The Health Smart Home (HIS) at the TIMC-IMAG laboratory, Grenoble

The Health Smart Home set-up by the TIMC-IMAG Laboratory of Grenoble is a real living and equipped flat that measures $47m^2$ (with an equipped kitchen, a bedroom, a living-room, a bathroom) in which the laboratory installed several sensors (see Fig. 1). The sensors used are:

- ▷ Presence infra-red sensors, that give an information on the localization of the patient at a given moment,
- ▷ Open/Close detectors, placed on communication doors and on some other strategic positions such as the door of the fridge or of the cupboard),
- ▷ A weather station delivers hygrometry and temperature,
- ▷ A kinematic sensor, ACTIM6D, placed on the patient, that detects changes in posture (sit-down, lie-down, stand-up, etc.) and provides information on his level of activity,
- ▷ Microphones for sound and speech recognition,
- ▷ Large-angle webcams used only for indexation in learning-based fusion algorithms.

Anthony Fleury and Norbert Noury are with the Laboratory TIMC-IMAG, UMR CNRS/UJF 5525, team AFIRM. Faculté de Médecine de Grenoble, bâtiment Jean Roget, 38706 La Tronche, France (e-mail: {Anthony.Fleury,Norbert.Noury}@imag.fr).

Michel Vacher, Jean-François Serignat and Hubert Glasson are with the LIG Laboratory, UMR UJF/CNRS/INPG 5217, team GETALP. 220 rue de la Chimie, BP 53, 38041 Grenoble Cedex 9. (e-mail: {First-Name.Last-Name}@imag.fr).

All these sensors linked to the four computers that are in the technical room of the HIS and data are stored on the fly by all of them.

B. Microphones Installation



Fig. 2. Integration of the microphones in the environment of the Health Smart Home

Eight omnidirectional Electret microphones (ECM-1) have been integrated in the Health Smart Home. They have been placed in the ceiling around the flat, and have been hidden as much as possible (see Fig. 2). All of them are plugged to the channel inputs on the acquisition board (NI PCI-6034E, National Instrument), of a computer in the technical room of the HIS.

The microphones are more less equally distributed in the flat. For instance, in the living-room, that is $2.9m \times 4.75m$, we placed two microphones, at the axis of the $2.9m$, and about $1m$ from each wall. Two other microphones are placed in the bed-room, one in the kitchen, one in the hall, one in the bathroom and one in the WC.

For each microphones, we have adjusted the gain in the software to the gain of the acquisition chain (that depend mostly of the microphone itself) to reach the best dynamic range (maximum detection without saturation).

III. METHODS

A. Data Acquisition and Processing

The global organisation of the system is shown on Fig. 3. The following sections introduce the different parts of the system.

1) *Sound Detection*: The first stage of the sound and speech analysis is the sound detection. The eight analog input channels are continuously and simultaneously sampled by the system at $16kHz$. The noise level is evaluated and the detection of the beginning and the end of the signal use an adaptive threshold. When the beginning and the end of the signal are evaluated, a sound file (wav format) is created ready to be used by the next thread of the application in charge with the segmentation.

2) *Sound vs Speech Segmentation*: This part has to classify a given sound into speech or sound of daily life. Segmentation is achieved through a GMM classifier trained with the everyday life sound corpus and the normal/distress speech corpus recorded in the LIG laboratory. Acoustical features are LFCC with 16 filter banks and the classifier uses 24 Gaussian models. These features were used because life sounds are better discriminated from speech with constant

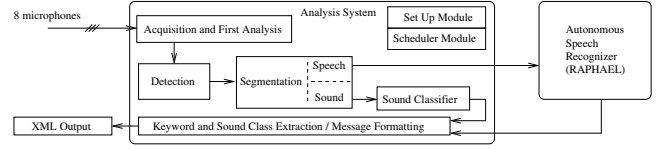


Fig. 3. The global sound and speech recognition system.

bandwidth filters than with MFCC and Mel scale. Frame width is 16 ms with an overlap of 50%.

The validation of this segmentation module was made by mixing the sounds and speech records from the corpora and adding them noise recorded in the HIS at 4 Signal to Noise Ratios (training was performed on pure sounds). In these "laboratory" conditions, we obtained a Segmentation Error Rate of 17.3% for a SNR of 0 dB, 5.1% at 10 dB, 3.8% at 20 dB and finally 3.6% at 40 dB. We can notice that SER remains quite constant with a 5% value above 10 dB.

3) *Sound Classification*: When segmented as sounds, the wav file is then processed by the classification part of the algorithm. Everyday life sounds are classified with a GMM or HMM classifier, the classifier is selected before the beginning of the experiment. They were trained with the eight classes of the everyday life sound corpus using LFCC features (24 filter banks) and 12 Gaussian models.

The every day life sounds are divided into 8 classes corresponding to 2 categories: *normal* sounds related to usual activities of the patient (door clapping, phone ringing, step sounds, dishes sounds, door lock), *abnormal* sounds related to distress situations (breaking glass, fall of an object, screams). This corpus contains some records made at LIG laboratory (61%) using super-cardioids microphones (eW500, Sennheiser), some files coming from a preceding corpus recorded at the time of former studies in the CLIPS laboratory and some files obtained from the Web. The corpus is constituted of 1,985 audio files for a total duration of 35 min 38 s, each file contains one sound.

We also evaluated the performance of this classification, in the same conditions as for segmentation, using different SNR. With the GMM model, 24 LFCC, the Classification Error Rate is 36.6% at 0 dB, 21.3% at 10 dB, 12% at 20 dB and finally 9.3% at 40 dB. We notice again that the error is highly dependant of the SNR.

4) *Speech Recognition*: The autonomous speech recognizer RAPHAEL [9] is running as an independent application and analyzes the speech events resulting from the segmentation module through a file exchange protocol. As soon as the requested file has been analyzed, it is deleted and the 5 best hypothesis are stored in a hypothesis file. This event allows the scheduler to send another file to be analyzed. The language model of this system is a medium vocabulary statistical system (around 11,000 words in French). This model was obtained by extraction of textual information from the Internet and from the French journal "Le Monde" corpora.

In order to train and validate the system we have recorded

two adapted corpora: the *normal/distress speech corpus* in French and the *life sound corpus*. For speech recognition, the training of the acoustic models was made with large corpora in order to insure a good speaker independence. They were recorded by 300 French speakers in the CLIPS laboratory (BRAFI00) and LIMSI laboratory (BREF80 and BREF120). All corpora were recorded using the same 16 kHz sampling rate as the analysis system. We have 66 normal and 60 distress sentences [10].

The speech recognition system was evaluated with the sentences from 5 speakers of our corpus (630 tests). In 6% of the cases, for normal sentences, an unexpected distress keyword is detected by the system and leads to a *False Alarm Sentence*. In 16% of the cases, for distress sentences, the distress keyword is not recognized (missed): this leads to a *Missed Alarm Sentence*. This often occurs in isolated words like "Aïe" (Ouch) or "SOS" or in syntactically incorrect French expressions like "Ça va pas bien" (I don't feel well). The global Speech Recognition Error Rate is then 11%.

B. Acquisition Software

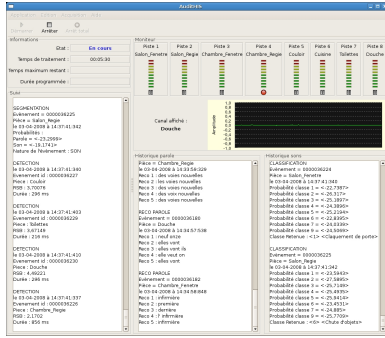


Fig. 4. The interface of the Acquisition software developed at the LIG Laboratory with: 3 text columns containing from left to right (1) the information on the detections, (2) the speech recognition and (3) the sound classification. The level of the eight microphones is represented on the right and the wave of the last detected microphone is continuously drawn.

Fig. 4 presents the application realised under GNU/Linux that implements all the preceding algorithms. This application is a multi-threading application that performs the following tasks: communication with the NI Board, detection of the sounds and creation the Wav files, then segmentation and classification of a sound, or for a speech, communication with the speech recognition system (that is an independent application). For each sounds and speech detected and classified by the software, an XML file containing all the information (date, time, SNR, segmentation and classification or sentences) is created with the associated logs and wav files.

This application allows us both to realize real experimentations and to post-process the data using the created XML files. Moreover, in addition to all that has been described, it has a modifiable threshold so that every sounds under a SNR will be ignored. This allows us to reduce the amount of

data to be processed by ignoring the sounds under 5 dB that would be undoubtedly very badly classified. These files are not segmented.

C. Post-processing

Each sound is recorded by the application and stored on the hard drive of a computer, with the associated XML file containing the information on the file (from detection to classification). Afterwards, these collected data are processed using MatlabTM.

Then the sounds are classified considering and fusing the results of the different microphones using the following algorithm. For a sound that will be done in the flat, we will take the SNR of the best microphone (named x), and keep all the microphones having a SNR greater than $0.8 * x$. We further take the decision from a vote between these different decisions. We apply two rules in case of equality: (1) if a distress speech is detected, we keep this decision and (2) in case of equality with another decision than a distress speech, we keep the decision of the microphone having the best SNR. This classification will give us two pieces of information for each event: the kind of event (sound or speech) and the retained class. We will create (for future use in data fusion) a suite of sound and speech events.

IV. EXPERIMENTAL RESULTS

A. Protocol

To validate the system in unsupervised conditions, we built a scenario in which every subject has to pronounce 45 sentences (20 distress, 10 normal and 3 phone conversations of 5 sentences each). For this experimentation, 13 subjects volunteered, 3 women and 10 men (age: 33 ± 12 years, weight: 64 ± 20 kgs, height: 1.74 ± 0.06 m). The number of sounds collected by this experimentation was 5,417 (2,399 of them were not segmented because their SNR was less than 5 dB), with an SNR of 12.5 ± 5.6 dB. After classification, we kept 1,820 sounds with a mean SNR of 13.6 ± 6.5 dB.

The experimentation took place during the day – so we do not control the environmental conditions of the experimental session (such as all the noises in the neighbourhood). The sentences were uttered in the flat, with the subject sat down or stand up. He was between 1 and 10 meters away from the microphones and had no instructions on his orientation with respect to the microphones (he could choose to turn his back to the microphone).

The protocol was quite simple. The subject was asked to first make a little scenario (close a door, make a noise with a cup and a spoon, let a box fall on the floor and scream "Aïe"). This whole scenario was repeated 3 times. Then, he had to read a succession of 10 normal and 20 distress sentences. After, he received a phone call and had to answer and read the phone conversation given. To realise the five different successions of sentences, we choose 30 representative ones and realised 5 phone conversations, and

TABLE I
CONFUSION MATRIX FOR SOUND AND SPEECH RECOGNITION (BOLD VALUES CORRESPONDS TO THE WELL CLASSIFIED SOUNDS).

		Results									
		Clap	Step	Phone	Dishes	Lock	Break	Falls	Scream	Normal Speech	Distress Speech
Action	Doors Clapping	81.25 %	0 %	0 %	0 %	0 %	0 %	18.75 %	0 %	0 %	0 %
	Phone Ringing	0 %	0 %	100 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %
	Dishes Sound	0 %	0 %	0 %	42.86 %	0 %	0 %	0 %	4.76 %	52.38 %	0 %
	Object Fall	19.05 %	0 %	0 %	4.76 %	0 %	0 %	76.19 %	0 %	0 %	0 %
	Scream	8.7 %	0 %	0 %	8.7 %	0 %	0 %	30.43 %	30.43 %	21.74 %	0.00 %
Normal Speech		0.74 %	0 %	0.37 %	4.1 %	0 %	0 %	3.35 %	4.48 %	83.44 %	3.49 %
Distress Speech		0.74 %	0.37 %	0 %	2.4 %	0.37 %	0 %	3.35 %	0 %	62.92 %	29.85 %

then we scrambled the sentences five times, and we randomly chose 3 of the 5 conversations for a given subject. This leads to a large number of 563 sentences uttered, out of which 268 are distress sentences.

B. Results

The results of this experimentation are summed-up in the confusion matrix of the global system (Table I). The different lines are the action performed, and the columns give the result of the system. The bold values are the correct decisions that were taken by the system.

This table shows us the classes that are close (e.g. object fall and doors clapping or dishes and normal sentences) and difficult to separate. We note 0% between screams and distress sentences, due to the fact that distress sentences could be reduced to a short word uttered by the subject like a scream. It is neither a bad segmentation nor a bad classification to take a scream instead of a distress sentence in this case. Screams are also close to object falls and speech (normal sentences). To complete this table, we could had that the global performances of the system are 89.76% of good segmentation, 72.14% of well-classified sounds and 41% of well-recognized sentences. This leads to 18.1% of false alarms and unfortunately to 70.1% of missed alarms. For the detection part, with our adaptive threshold, each sound is well detected by the system.

V. DISCUSSION AND CONCLUSION

This paper presents a complete sound and speech recognition system, with evaluation results in unsupervised and real conditions, compared to the results obtained in laboratory conditions. For the events tested, we can see that the results for the sound recognition are good and conform with the results obtained in laboratory conditions, if we consider the SNR of the HIS.

As far as speech recognition is concerned, the results are too low, especially for the distress sentence recognition. Even if the corpora was made independent of the speaker, we face difficulties of recognition because each subject pronounces the sentence differently. Moreover, the acquisition line, the microphones and the environment are all imperfect. We could

have a noise added to the sound and disturb the HMM process. The conditions are also uncontrolled because the subject could pronounce the sentence when he decides, and could freely choose his orientation to the microphone. Thus our conditions are the worst possible, far from the laboratory conditions (no noise and the microphone just behind the subject). Nevertheless these real collected sounds will be used to improve the models of language and the results for next experimentations. We are also working on the learning of other classes.

ACKNOWLEDGMENT

The authors would like to thanks all the subject from both laboratories for their time spent doing the experiments.

REFERENCES

- [1] P. Barralon, N. Noury, and N. Vuillerme, "Classification of daily physical activities from a single kinematic sensor," in *27th Annual International Conference of the IEEE-EMBS 2005*, 2005, pp. 2447–2450.
- [2] B. Najafi, K. Aminian, A. Paraschiv-Ionescu, F. Loew, C. J. Büla, and P. Robert, "Ambulatory system for human motion analysis using a kinematic sensor: Monitoring of daily physical activity in the elderly," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 6, pp. 711 – 723, June 2003.
- [3] N. Noury, A. Fleury, P. Rumeau, A. Bourke, G. O. Laighin, V. Rialle, and J. Lundy, "Fall detection - principles and methods," in *29th Annual International Conference of the IEEE-EMBS 2007.*, 22–26 Aug. 2007, pp. 1663–1666.
- [4] A. K. Bourke, J. V. O'brien, and G. M. Lyons, "Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm," *Gait Posture*, vol. 26, no. 2, pp. 194–199, Jul 2007.
- [5] G. LeBellego, N. Noury, G. Virone, M. Mousseau, and J. Demongeot, "A model for the measurement of patient activity in a hospital suite," *IEEE Trans. Inf. Technol. Biomed.*, vol. 10, no. 1, pp. 92–99, Jan. 2006.
- [6] V. Riquembourg, D. Menga, D. Durand, B. Marhic, L. Delahoche, and C. Loge, "The smart home concept : our immediate future," in *E-Learning in Industrial Electronics, 2006 1ST IEEE International Conference on*, 18–20 Dec. 2006, pp. 23–28.
- [7] B. Bouchard, A. Bouzouane, and S. Giroux, "A smart home agent for plan recognition of cognitively-impaired patients," *Journal of Computers*, vol. 1, no. 5, pp. 53 – 62, Aug. 2006.
- [8] J.-C. Wang, H.-P. Lee, J.-F. Wang, and C.-B. Lin, "Robust environmental sound recognition for home automation," *IEEE Trans. Autom. Sci. Eng.*, vol. 5, no. 1, pp. 25–31, Jan. 2008.
- [9] M. Akbar and J. Caelen, "Parole et traduction automatique : le module de reconnaissance raphael," in *Proc. COLING-ACL'98, Montréal, Québec*, Aug. 10–14 1998, pp. 36 – 40.
- [10] M. Vacher, J.-F. Serignat, S. Chaillol, D. Istrate, and V. Popescu, "Speech and sound use in a remote monitoring system for health care," in *Proceedings of the Ninth International Conference on TEXT, SPEECH and DIALOGUE, Brno, Czech Republic, Sep. 11U15*, 2006.



Auteur : Anthony FLEURY

Titre de la thèse : Détection de motifs temporels dans les environnements multi-perceptifs. Application à la classification automatique des Activités de la Vie Quotidienne d'une personne suivie à domicile par télémedecine

Directeurs de thèse : Norbert Noury, Michel Vacher

Date de soutenance : 24 octobre 2008

Résumé :

À l'horizon 2050, environ un tiers de notre population sera âgée de soixante-cinq ans et plus. Les travaux de l'équipe AFIRM du TIMC-IMAG visent à surveiller les personnes âgées à domicile pour détecter une perte d'autonomie le plus précocement possible. Pour ce faire, les travaux de cette thèse tentent d'objectiver les critères ADL ou les grilles de type AGGIR, en classifiant de manière automatique les différentes activités de la vie quotidienne d'une personne par l'intégration de capteurs, créant un Habitat Intelligent pour la Santé (HIS).

L'appartement HIS possède des détecteurs de présence infrarouges (localisation), des contacteurs de porte (utilisation de certaines commodités), un capteur de température et d'hygrométrie dans la salle de bains et des microphones (classification des sons/ reconnaissance de la parole avec l'équipe GETALP du LIG). Un capteur cinématique embarqué détecte les transferts posturaux (reconnaissance de formes avec la transformée en ondelettes) et les périodes de marche (analyse fréquentielle).

La première partie de ce manuscrit présente la réalisation du capteur cinématique et les algorithmes associés puis une première validation sur des sujets jeunes suivi de la mise en place et de la validation des autres capteurs de l'appartement HIS et enfin l'algorithme de classification des sept activités de la vie quotidienne reconnues (hygiène, élimination, préparation et prise de repas, repos, habillage/déshabillage, détente et communication), par l'intermédiaire des séparateurs à vaste marge. La seconde partie décrit le protocole expérimental pour valider cette classification et discute de la généralisation des premiers résultats présentés.

Mots clés :

Classification, Activités de la Vie Quotidienne, Séparateurs à Vaste Marge (SVM), Habitat Intelligent pour la Santé (HIS), accéléromètres, magnétomètres, transformée en ondelettes, création de bases d'ondelettes, quaternions.