Contribution à l'analyse des signaux et à leur classification. Applications aux Technologies pour la Santé, au maintien à domicile de la personne fragile et à l'analyse du comportement humain.

Anthony Fleury

▶ **To cite this version:**

Anthony Fleury. Contribution à l'analyse des signaux et à leur classification. Applications aux Technologies pour la Santé, au maintien à domicile de la personne fragile et à l'analyse du comportement humain.. Traitement du signal et de l'image [eess.SP]. Université de Lille, 2019. tel-02080185

HAL Id: tel-02080185
https://theses.hal.science/tel-02080185

Submitted on 26 Mar 2019

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.

Habilitation à Diriger des Recherches

Spécialité : Génie informatique et traitement du signal

Présentée et soutenue publiquement par

Anthony FLEURY

Le 20 Février 2019

Contribution à l'analyse des signaux et à leur classification. Applications aux Technologies pour la Santé, au maintien à domicile de la personne fragile et à l'analyse du comportement humain.

JURY :

Président	François BRÉMOND	Directeur de Recherche INRIA, Sophia Antipolis
Rapporteurs	François CHARPILLET	Directeur de Recherche INRIA, Nancy
	Bernadette DORIZZI	Professeure, Télécom Sud-Paris, Institut Mines-Télécom
	Edwige PISSALOUX	Professeure des Universités, Université de Rouen
Examinateurs	Régis BEUSCART	Professeur Émérite, Université de Lille
	Frédéric LERASLE	Professeur des Universités, Université Paul Sabatier, Toulouse
Garant	Stéphane LECOEUCHE	Professeur, IMT Lille Douai, Institut Mines-Télécom

Table des matières

Table des figures **ix**

Liste des tableaux **xi**

Remerciements **xv**

1 Introduction **1**
1.1 Contexte des travaux et historique . . . 1
1.2 Parcours professionnel . . . 2
1.3 Le traitement de données et ses applications . . . 3
1.4 Objectifs des travaux . . . 4
1.5 Organisation du manuscrit . . . 4

2 La personne et ses automatismes : autour de sa posture et de sa proprioception **7**
2.1 Introduction . . . 7
2.2 Objectifs de ces travaux et verrous scientifiques à lever . . . 8
2.3 État de l'art sur la mesure de données physiologiques et de données inertielles . . . 8
2.3.1 Les données cardio-respiratoires . . . 8
2.3.2 Les données de posture et de marche . . . 9
2.3.2.1 Le système de contrôle de l'équilibre . . . 9
2.3.2.2 Outils de mesure et données smartphone . . . 10
2.4 Analyse des signaux posturaux avec l'entropie d'échantillon . . . 12
2.4.1 L'entropie d'échantillon . . . 12
2.4.2 Application au contrôle postural . . . 13
2.4.2.1 L'entropie d'échantillon pour caractériser un signal de posture . . . 13
2.4.2.2 Expérimentations et analyse statistique . . . 14
2.4.2.3 Résultats . . . 14
2.4.2.4 Conclusions sur cette étude . . . 15
2.5 Analyse non linéaire des signaux par décomposition : modification de l'entropie spectrale pour la construction d'un indice local de complexité . . . 17
2.5.1 L'entropie spectrale (Wavelet Entropy) . . . 17
2.5.2 La transformée par synchrosqueezing . . . 18
2.5.3 Calcul de l'entropie spectrale modifiée WE_{SQT} . . . 18
2.5.4 Tests sur des signaux posturaux et protocole . . . 18
2.5.5 Résultats et discussion . . . 19
2.6 Les signaux du smartphone : vers l'application iProprio . . . 20
2.6.1 Test de la validité de ces mesures . . . 21
2.6.1.1 Matériel, algorithmes et méthode . . . 21
2.6.2 iBalance . . . 26
2.6.2.1 Architecture du système proposé . . . 26
2.6.2.2 Retour sensoriel . . . 27
2.6.2.3 Expérimentations et résultats . . . 28
2.6.3 iProprio . . . 29
2.6.3.1 Présentation de l'application et du protocole de test . . . 30
2.6.3.2 Résultats . . . 31

2.6.4 Demande attentionnelle . . . 32
2.6.4.1 Système de mesure . . . 32
2.6.4.2 Expérimentation . . . 32
2.6.4.3 Analyse de données et résultats . . . 33
2.7 Conclusion . . . 34

3 La personne et ses actions : extraction de caractéristiques et reconnaissance de formes 35
3.1 Introduction et verrous scientifiques à lever . . . 35
3.2 Identification/ré-identification de personnes dans des réseaux de caméras : de l'extraction d'information à la classification en ligne . . . 36
3.2.1 Acquisition et traitement de données vidéos . . . 36
3.2.2 But et originalité des travaux . . . 37
3.2.3 Base de données de test . . . 38
3.2.4 Extraction des caractéristiques . . . 38
3.2.5 Sélection et résultats . . . 39
3.2.5.1 ACP . . . 40
3.2.5.2 Sélection basée sur la corrélation . . . 40
3.2.5.3 Sélection de type Wrapper . . . 40
3.2.6 Discussion sur les attributs extraits . . . 40
3.2.7 Techniques d'apprentissage pour l'identification et la ré-identification de personnes . . . 41
3.2.7.1 Motivation et origines des travaux . . . 41
3.2.7.2 Apprentissage par SVM Incrémental : MidSVM – ré-identification de personnes . . . 42
3.2.7.3 Self-Adaptive Kernel Machine (SAKM) – Identification et ré-identification de personnes 43
3.2.8 Conclusions des travaux . . . 44
3.3 Reconnaissance d'actions . . . 45
3.3.1 Les cameras RGB-D . . . 45
3.3.2 Extraction d'informations dans les signaux vidéos RGB-D . . . 47
3.3.2.1 Travail sur l'indépendance à l'orientation . . . 47
3.3.2.2 Le descripteur Body-part Directional Velocity (BDV) : direction et vitesse . . . 47
3.3.3 Classification indépendante de l'orientation . . . 48
3.3.3.1 Approche proposée . . . 48
3.3.3.2 Fusion des informations . . . 49
3.3.4 Expérimentations . . . 51
3.3.5 Classification au fil de l'eau et anticipée . . . 53
3.3.5.1 Système proposé : classification avec des HMMs . . . 53
3.3.5.2 Résultats expérimentaux . . . 54
3.3.6 Conclusion de ces travaux . . . 57
3.4 Conclusions . . . 57

4 La personne et ses activités : analyse/reconnaissance dans les habitats intelligents et création de profils 59
4.1 Introduction et verrous scientifiques à lever . . . 59
4.2 État de l'art sur l'extraction d'informations dans les habitats intelligents . . . 61
4.2.1 Segmentation des données . . . 61
4.2.1.1 Segmentation basée sur l'activité . . . 61
4.2.1.2 Segmentation basée sur le temps . . . 62
4.2.1.3 Segmentation basée sur les événements (déclenchement de capteurs) . . . 62
4.2.2 Extraction des caractéristiques par fenêtrage . . . 62
4.2.2.1 Fenêtrage basé sur le temps . . . 62
4.2.2.2 Fenêtrage basé sur les évènements . . . 63

4.3 Ensemble de données utilisés dans ce chapitre . . . 63
4.3.1 HIS . . . 64
4.3.2 Domus . . . 65
4.4 Sélection d'attributs dans les signaux issus des environnements intelligents . . . 67
4.4.1 Descripteurs . . . 67
4.4.2 Sélection des attributs . . . 67
4.4.2.1 Méthode de sélection . . . 67
4.4.2.2 L'ensemble de données Sweet-Home . . . 68
4.4.2.3 L'ensemble de données HIS . . . 69
4.4.3 Temps du fenêtrage et son influence . . . 69
4.4.4 Conclusions . . . 69
4.5 Un aperçu de la classification d'activités dans les habitats intelligents avec des méthodes en ligne . . . 70
4.5.1 Introduction . . . 70
4.5.2 Différentes méthodes de modélisation d'activités . . . 71
4.5.2.1 Modèles de Markov Cachés (HMM) . . . 71
4.5.2.2 Les Conditional Random Fields (CRFs) . . . 72
4.5.2.3 Modélisation par des réseaux logiques de Markov (Markov Logic Network – MLN) . . . 72
4.5.2.4 Modélisation d'activités par des méthodes non séquentielles : SVMs and Random Forests 72
4.5.3 Expérimentations . . . 72
4.5.3.1 Évaluation des performances . . . 73
4.5.3.2 Résultats globaux . . . 73
4.5.3.3 Résultats sur les données Sweet-Home . . . 74
4.5.3.4 Résultats pour l'ensemble des données HIS . . . 74
4.5.3.5 Performance des CRFs pour reconnaître les activités . . . 75
4.5.3.6 Analyse complémentaire . . . 76
4.5.4 Discussion . . . 76
4.5.5 Conclusion sur ces travaux . . . 78
4.6 Utilisation de la reconnaissance d'activités et création de profils dans les habitats . . . 78
4.6.1 Problématique . . . 79
4.6.2 Architecture globale de la solution proposée . . . 79
4.6.3 Plateforme de test Eco-Confort d'IMT Lille Douai et retours utilisateurs . . . 80
4.6.4 Suite des expérimentations . . . 82
4.7 Conclusions . . . 83

5 Conclusions et perspectives des travaux de recherche 85
5.1 Conclusions des travaux présentés . . . 86
5.2 Perspectives de recherche . . . 87
5.3 Développements des recherches à court terme . . . 88
5.3.1 Le projet iBalance/iProprio . . . 89
5.3.2 Profils utilisateurs – jeu sérieux dans les habitats . . . 89
5.3.3 Analyse des signaux smartphone dans la vie courante – Déduction d'indicateurs de qualité de vie? 90
5.4 La place de la santé, de la donnée et des habitats intelligents dans l'IMT et l'IMT Lille Douai . . . 90
5.5 Collaboration et communauté scientifique . . . 91

Bibliographie 93

Annexes – Curriculum Vitæ détaillé **105**

A Curriculum Vitæ **105**
Formation Universitaire . . . 105
Expérience Professionnelle . . . 105
Compétences Techniques . . . 106
Langues . . . 106

B Activités d'enseignement **107**
B.1 Résumé des activités d'enseignement . . . 107
B.2 Implication dans les activités administratives liées à la pédagogie d'IMT Lille Douai . . . 108
B.3 Enseignements réalisés et en cours . . . 108
B.3.1 Tronc commun IMT Lille Douai . . . 108
B.3.2 Cours électifs et cours de mineures . . . 109
B.4 Domaine de la Santé . . . 109
B.5 Nouvelles pratiques et innovations pédagogiques . . . 109
B.5.1 Méthode ALPES . . . 110
B.5.2 Classes inversées . . . 110
B.5.3 Notation automatisée et basée sur le jeu . . . 110
B.6 Initiation à la recherche des élèves ingénieurs . . . 111
B.7 Évolution des activités d'enseignement dans le nouveau programme d'IMT Lille Douai . . . 111

C Activités de recherche académique et appliquée **113**
C.1 Mots clés . . . 113
C.1.1 Recherche . . . 113
C.1.2 Applications . . . 113
C.2 Résumé des activités . . . 114
C.2.1 Formation à la recherche . . . 114
C.2.2 Formation par la recherche . . . 114
C.2.3 Principales participations à des projets de recherche . . . 115
C.3 Jury de thèse . . . 116
C.4 Distinctions . . . 116
C.5 Activités de recherche . . . 116
C.5.1 Verrous scientifiques . . . 117
C.5.2 Extraction et traitement de signaux : application à la médecine personnelle . . . 117
C.5.3 Extraction, sélection d'attributs et classification statique/dynamique . . . 118
C.5.3.1 Appliquée à l'habitat intelligent et à la reconnaissance d'activités . . . 118
C.5.3.2 Applications à la sécurité et à l'usine du futur . . . 118

D Implication dans la communauté scientifique et éducative **121**
D.1 Activités en lien avec la communauté scientifique . . . 121
D.2 Appartenances à des groupements scientifiques . . . 122
D.3 Évaluation de projets . . . 122
D.4 Relectures, comités éditoriaux . . . 122
D.5 Organisations de conférences et de sessions . . . 122
D.6 Participation à des réseaux scientifiques Santé . . . 123
D.7 Collaborations suivies nationales et internationales . . . 123
D.8 Vie de l'équipe et de l'école . . . 123

E Indices de production et liste des publications 125
E.1 Indices de production . . . 125
E.1.1 Résumé des contributions . . . 125
E.1.2 Bibliométrie et citations . . . 125
E.2 Liste des publications . . . 126
Brevets . . . 126
Dépôts Logiciels (APP) . . . 126
Journaux Internationaux (avec IF JCR) . . . 127
Journaux Internationaux . . . 128
Chapitres d'ouvrages . . . 128
Journaux nationaux . . . 129
Conférences Internationales avec Actes et Comité de Lecture . . . 129
Conférences Nationales avec Actes et Comité de Lecture . . . 132
Participations à des journées techniques . . . 132
Séminaires Invités . . . 133
Article de vulgarisation . . . 133
Mémoires . . . 133
Activités d'édition . . . 133
Bases de données publiques . . . 134

Table des figures

1.1 Vue synthétique de mon parcours professionnel . . . 2
1.2 Mise en relation des travaux avec les domaines (relatifs à des sections CNU) . . . 3
1.3 Résumé synthétique des thématiques présentées dans chacun des trois chapitres de ce manuscrit, dans l'ordre croissant du plus intérieur au plus extérieur. . . . 5

2.1 Moyenne et variance de la variabilité (A_{CoP}) (panel A), de la vitesse (V_{CoP}) (panel B), et de la régularité (SEn_{CoP}) (panel C) de la trajectoire du centre des pressions dans les deux conditions, sol ferme (désigné par Firm et représenté en blanc) et mousse (désigné par Foam et représenté en noir). La significativité est prise à deux niveaux (* : $p < 0.05$; *** : $p < 0.001$). . . . 15
2.2 Moyenne et variance de la variabilité (A_{CoP}) (panel A), de la vitesse (V_{CoP}) (panel B), et de la régularité (SEn_{CoP}) (panel C) de la trajectoire du centre des pressions dans les deux conditions, sans orthèse (désigné comme No-AFO, blanc) et avec orthèse (désigné comme AFO, noir). La significativité est prise à deux niveaux (* : $p < 0.05$; ** : $p < 0.01$). . . . 15
2.3 Vue schématique des résultats de l'évolution de l'entropie d'échantillon et des trajectoires et vitesse du centre des pressions en fonction de la condition. . . . 16
2.4 Transformée pour synchrosqueezing et WE_{SQT} correspondante . . . 19
2.5 Bras robotisé Kuka et ses axes (figure fournie par le manuel de spécifications "KR 5 sixx R650, R850"). . . . 21
2.6 Estimation des angles . . . 22
2.7 Positionnement des capteurs sur le robot . . . 23
2.8 Résultats des mesures en statique sur les angles de roulis et tangage . . . 24
2.9 Résultats des mesures en dynamique sur l'angle de roulis (à 20 et 50% de la vitesse maximale) . . . 25
2.10 Résultats des mesures en dynamique sur l'angle de tangage (à 20 et 50% de la vitesse maximale) . . . 25
2.11 Architecture du système iBalance . . . 26
2.12 Déclenchement du retour sensoriel auditif dans le système iBalance . . . 27
2.13 Illustration du positionnement du smartphone et des deux positions de pieds (à gauche pieds parallèles, à droite pieds tandem) . . . 28
2.14 Résultats sur les sujets jeunes. Les valeurs grises sont sans retour sensoriel audio (No-ABF) alors que les valeurs noires sont avec ce retour (ABF). La figure (A) présente le RMS en degré (la quantité de mouvement du tronc sur l'axe ML), la figure (B) présente l'énergie, la figure (C) la Spectral Edge Frequency à 95% et la figure (D) le temps en dehors de la zone morte. . . . 29
2.15 Résultats sur les sujets plus âgés. De gauche à droite : Le RMS des mouvements du tronc en degré, le pourcentage de temps en dehors de la zone morte et enfin la fréquence donnant la moitié de la puissance (Mean Power Frequency) . . . 30
2.16 Interface graphique et types de retours sensoriels . . . 30
2.17 Erreurs absolues et variables de repositionnement pour les sujets jeunes, avec ** : $P < 0.01$ et * : $P < 0.05$ 31
2.18 Erreurs absolues et variables de repositionnement pour les sujets âgés, avec ** : $P < 0.01$ et * : $P < 0.05$. 32
2.19 Architecture du système de mesure de la demande attentionnelle . . . 33
2.20 Résultats sur les trois variables analysées à savoir le RMS, la Fréquence donnant la moitié de la puissance cumulée et le Temps de réaction, pour différentes position, avec *** : $P < 0.001$ et * : $P < 0.05$ . . . 33

3.1 Structure du système de reconnaissance de personnes proposé . . . 37
3.2 Une personne et les six marches différentes dans la base CASIA . . . 38
3.3 Deux exemples ($P19$ et $P15$) de classes de tests qui sont difficilement reconnues. . . . 44
3.4 Exemples de classes avec un fort taux de reconnaissance. . . . 44

3.5 Subdivision du corps humain en cinq parties (B_1, B_2, B_3, B_4 et B_5). 48
3.6 Architecture du système logique flou. Les seuils Q_0 et Q_1 sont deux seuils flous déterminés sur les bases expérimentales et qui vont déterminer les fonctions floues utilisées 49
3.7 Algorithme DFSDA 50
3.8 Trois exemples de cartes de profondeur obtenues pour différents points de vue dans l'ensemble Multi-view3D. 51
3.9 Architecture de l'algorithme DSFDA auquel nous avons ajouté l'algorithme d'agrégation. 52
3.10 Vision générale de l'architecture de reconnaissance proposée 53
3.11 Illustration du processus de classification mis en place 54
3.12 Matrices de confusion sur la base MSRAction3D 56
3.13 Matrice de confusion de notre algorithme sur la base Florence3D 56
3.14 Évolution de la vraisemblance dans le cas d'une action « Side kick » 56
3.15 Boîtes à moustache indiquant la vitesse de convergence de la décision (pour obtenir une classification stable) sur les trois sous-ensembles de MSRAction3D 57

4.1 Exemple d'un habitat intelligent (Laboratoire TIMC-IMAG, Grenoble) avec la vue des caméras. 60
4.2 Présentation succincte des différentes approches de segmentation. Le haut présente la séquence de déclenchement/mesure des capteurs et les activités en cours de réalisation au moment de ces déclenchements (de 1 à 4). (a) : une fenêtre pour une activité, (b) : fenêtre fixée par le temps et (c) : fenêtre avec un nombre fixe de déclenchements/mesures de capteurs. 61
4.3 Les deux plateformes d'expérimentation. À gauche la plateforme DOMUS du LIG, à droite la plateforme HIS du laboratoire TIMC-IMAG 64
4.4 Méthodologie générale pour le système de reconnaissance d'activités humaines. 70
4.5 Précision vs. pourcentage de classe *« Autre »* dans l'ensemble des données pour CRF et RF sur la base de données HIS. 77
4.6 Architecture globale du système proposé 80
4.7 Local Eco-Confort, IMT Lille Douai. 81
4.8 Illustration de l'installation. 82
4.9 Table interactive et application web permettant de gérer l'habitat. 82

5.1 Résumé synthétique des thématiques présentées dans chacun des trois chapitres de ce manuscrit, dans l'ordre croissant du plus intérieur au plus extérieur. 86
5.2 Prévision d'évolution de démographie (source INSEE) en France 87
5.3 Les déserts médicaux et leur répartition dans le pays 87

E.1 Nombre de citations pour les différentes années, Google Scholar 126
E.2 Nombre de publications référencées (gauche) et nombre de citations dans des publications référencées pour les différentes années (droite) sur ISI Web Of Knowledge 126

Liste des tableaux

2.1 Effet de la condition visuelle sur le déplacement du centre des pressions. Significativité * : $p < 0,05$ . . . 19

3.1 Ensemble initial d'attributs sur la couleur et la texture . . . 39
3.2 Attributs sélectionnés par les différentes méthodes. . . . 41
3.3 Résultats de classification avec MidSVM, en considérant les ensembles CASIA-Wholeset, CASIA-PCA, CASIA-CFS et CASIA-Wrapper décrits précédemment . . . 42
3.4 Résultats de classification avec SAKM et la découverte de nouveautés (3 classes apprises et 7 testées). . . 43
3.5 Classification en ligne avec SAKM. 5 Expérimentations sur l'ensemble CFS, avec trois classes en apprentissage et un nombre croissant de classes ajoutées en test. . . . 43
3.6 Classification pour les 20 personnes avec l'ensemble CFS. L'apprentissage s'est fait avec 3, 5, 8 ou 10 personnes selon les cas. . . . 45
3.7 Comparaison entre les différentes approches . . . 52
3.8 Performances temporelles . . . 53
3.9 Subdivision de MSRAction3D en trois sous ensembles AS1, AS2 et AS3 . . . 55
3.10 Résultats (précision) de notre algorithme sur MSRAction3D en comparaison à différentes méthodes de la littérature . . . 55
3.11 Résultats de différents algorithmes sur la base Florence3D . . . 55

4.1 Distribution des fenêtres temporelles dans le corpus HIS ($T = 60s$). . . . 65
4.2 Distribution des fenêtres temporelles pour le corpus Sweet-Home ($T = 60s$). . . . 66
4.3 Attributs sélectionnés dans le corpus Sweet-Home. Les vingt meilleurs (avec comme critère la valeur de leur IGR) sont mis en gras . . . 68
4.4 Attributs sélectionnés dans le corpus HIS (ici nous ne précisons pas les 20 meilleurs car les écarts sont plus faibles et le nombre de dimensions éliminées moins important) . . . 69
4.5 Précision globale (%) sur les deux ensembles de données, avec et sans la classe *« Autre »* (la classe *« Autre »* comprenant, pour rappel, toute fenêtre qui n'est pas étiquetée dans la base de données) . . . 73
4.6 Résultats de classification pour l'ensemble des données Sweet-Home sans classe *« Autre »* (Exp 1) : forme résultats (moyenne par participant $\pm EcType$) . . . 74
4.7 Résultats de classification pour l'ensemble des données Sweet-Home avec classe *« Autre »* (Exp 2) : forme résultats (moyenne par participant $\pm EcType$). . . . 74
4.8 Résultats de classification pour l'ensemble des données HIS sans classe *« Autre »* (Exp 3) : forme résultats (moyenne par participant $\pm EcType$) . . . 75
4.9 Résultats de classification pour l'ensemble des données HIS avec classe *« Autre »* (Exp 4) : forme résultats (moyenne par participant $\pm EcType$) . . . 75
4.10 Matrice de confusion pour CRF avec l'ensemble des données Sweet-Home . . . 76
4.11 Matrice de confusion pour CRF avec l'ensemble de données HIS . . . 76

À Katerine, Enaël et Lyana
À Séverine
À Brigitte et Gérard.

Remerciements

En tant que chercheurs et enseignants-chercheurs, nous sommes tous très occupés. Je remercie donc tous les membres du jury ayant pris sur leur temps (y compris lors des congés de fin d'année) pour relire ces travaux et venir assister à la présentation. Je remercie ainsi François Charpillet, qui dix ans après avoir examiné mon rapport de thèse verra l'évolution avec la soutenance de HDR, Edwige Pissaloux qui avait déjà participé au jury de Quentin, l'un des doctorants, et qui a maintenant une vision plus large de mes travaux, et enfin Bernadette Dorizzi qui a également accepté de remettre un rapport sur ce manuscrit. Je remercie également Bernadette pour tout le temps qu'elle passe, avec Jérémie Pourchez, à animer la thématique ingénierie pour la santé à l'IMT et pour m'avoir demandé d'en faire partie du bureau.
En tant qu'examinateurs, je remercie tout d'abord Régis Beuscart d'avoir accepté de participer à ce jury et d'examiner mes travaux. J'en profite pour le remercier également pour tous ses efforts pour animer une thématique ingénierie pour la santé en région en ayant créé la SFR TSM. Je remercie également Frédéric Lerasle et François Brémond, qui prennent tous les deux le temps de venir de nouveau dans nos locaux pour juger ces travaux, qui vont s'investir dans la lecture et l'examen du manuscrit et de la soutenance qui sera associée. Enfin je remercie Stéphane pour son implication dans cette habilitation, ses conseils de présentation qui font qu'elle est plus clair et ses efforts pour m'aider à mener ce travail à bien.
La thèse est un travail personnel (qui se fait tout de même avec un encadrant) sur trois ans. L'habilitation à diriger des recherches, elle, est un travail qui est présenté par une personne mais qui reflète une collaboration avec de nombreuses personnes, que ce soit des stagiaires, des doctorants, des post-doctorants ou encore des collègues. Ces remerciements vont donc à toutes les personnes qui ont contribué à ces travaux.
Je commence ici par les étudiants et post-doctorants. Je voudrais adresser mes remerciements à Manon Lecomte et Olivia Carlton (deux étudiantes de l'université de Lille), Mounir Hammouche et Sid Ahmed Walid Talha que j'ai co-encadré en Master (Sid Ahmed qui continue en thèse), Céline Franco, Yanyun Lu et Quentin Mourcou, que j'ai tous les trois co-encadré en thèse, Pédro Chahuara, Nawel Yala, M'Hamed et Antoine Ghorra avec qui j'ai travaillé durant leur thèse ou partie de thèse, Houari Toubakh que j'ai co-encadré en post-doctorat et Jonathan Villain avec qui je travaille encore pour son second post-doctorat. Enfin, je voudrais remercier Abir Karami pour son travail en tant que Maitre Assistante Associée chez nous et bientôt en post-doctorat (en attendant vite mieux).
Je vais forcément en oublier mais je voudrais également remercier tous les collègues de l'école, cités ici dans le désordre, avec qui nous avons une bonne ambiance, qu'elle soit de travail ou de détente sur les pauses du midi (tarot, étoiles, etc.). Alors je remercie ici Mathieu, Gaëlle, Jérémie, Sébastien, Carole, Fabienne, Christine, Aurore, Muriel, Arnaud, Cécile, Noury, Eric, Luc, Sanda, Marin, Guillaume, Lucien, Moamar, Rémy, Isabelle, Frédéric, Florine, Patrick, Philippe, Annie, Etienne, Catherine, Jérôme, Hélène, Julie, Boulbaba, Hassen, Hazem, Jannick, Laurent, etc. etc. etc. et tous ceux qui sont passés par chez nous et avec qui j'ai pu avoir de bons contacts.
Je n'oublie pas non plus les collègues qui ne sont pas d'IMT Lille Douai et avec qui je peux travailler, encore dans le désordre Norbert, Marc, Eric, Damien, Michel, François, Nicolas, Belkacem, Virginie, Christophe, Eric, Jérôme, Nesma, Mossaab, Dan, Jacques, tous les collègues Sud-Américains avec qui nous montons en ce moment des collaborations et qui nous accueillent aussi chaleureusement (Carla, Rodrigo, Paulo, Windson, etc.).
Enfin, j'associe ce travail à toute ma famille qui m'apportent chaque jour leur soutien et font mon bonheur au niveau personnel (Lyana, Enaël, Séverine, Brigitte, Gérard, Frédéric, Sophie, Gabriel, Mathis, ma grand-mère, Marie-Ange, etc. Merci à tous, votre présence fait chaque jour que je suis celui qui est en mesure aujourd'hui de présenter ce manuscrit.

CHAPITRE 1

Introduction

1.1 Contexte des travaux et historique

BIG data, Deep Learning, Intelligence Artificielle, etc., dans notre monde actuel, ces mots ont commencé à rentrer dans le langage courant, même si ils ne sont pas tout le temps utilisés à bon escient.

En effet, dans ces dix dernières années, nous avons vu l'essor (différencié au départ) de deux phénomènes différents, mais qui ont modifié notre manière de vivre : l'arrivée du premier « Smartphone » facile à utiliser et grand public ; et l'arrivée des réseaux sociaux. Mélangés, ces deux phénomènes ont fait que nous sommes tous devenus dépendants de notre connectivité et surtout générateurs d'une quantité de données importante. Ces données ne s'arrêtent pas à nos plus belles histoires partagées sur les réseaux sociaux, mais contiennent aussi des données que nous générons pour mesurer notre bien-être ou nos habitudes de vie (nos données de sport, d'activités, nos rythmes de sommeil avec les applications dédiées, etc.). Le fait de posséder toutes ces données a une valeur inestimable.

À cela, nous pouvons ajouter le fait que l'attrait pour la donnée se développe partout. Nous avons tous maintenant des cartes de fidélité qui enregistrent nos habitudes de consommation, nous commandons régulièrement sur des sites Internet qui retracent, en plus de nos achats, tout notre parcours et nos déambulations sur le site). Nous ne savons pas estimer notre « empreinte en termes de quantité de données » que nous générons chaque jour, mais ce chiffre est important et pose un problème de sociologie intéressant, notamment le fait de pouvoir quantifier son évolution, par l'analyse des usages actuels et à venir.

Ainsi, nous le voyons, les entreprises ont compris l'intérêt d'avoir des données sur les personnes même si elles ne savent pas à tout instant ce qu'elles pourraient en faire. Stocker celles-ci est toujours un bien et ne coûte que peu pour l'instant par rapport à ce que les entreprises pensent que cela pourrait leur rapporter. Dans le domaine de l'industrie et des services, stocker ces données permet de mieux connaître les clients, mais surtout et malheureusement pour le consommateur/le client, cela permet aussi aux entreprises de revendre les données entre-elles. Dans le domaine du bien-être de la personne, qui est un domaine qui se développe de plus en plus car il permet d'éviter le problème de la donnée de santé, le cas est le même. De plus en plus de capteurs ou d'applications stockent et collectent des données, et ceci leur permet de connaître de mieux en mieux les populations visées. Ces modèles de population ensuite peuvent se revendre et se négocier. La donnée, de nos jours, est donc au cœur de notre vie et des différentes applications de l'industrie et des services.

Dans ce contexte, actuellement, se développe de plus en plus l'analyse de données, allant de l'extraction d'information à leur compréhension pour en sortir des informations de plus haut niveau. Les travaux de recherche présentés dans cette habilitation se font sur cette thématique générale de la donnée et de la modélisation de phénomènes ou de comportement à partir de celle-ci. Les applications de ces travaux sont multiples, mais nous nous sommes concentrés sur principalement deux activités : l'application à la santé et aux bien-être de la personne, l'application à la sécurité. L'analyse de la donnée appliquée au bien-être présentée ici se place dans deux contextes applicatifs différents. Le premier est l'analyse de données issues du smartphone pour compléter les mesures faites par un clinicien par exemple et permettre à la personne de continuer des exercices physiques en toute autonomie chez elle. Le second est celui des habitats intelligents, habitats dans lesquels il est possible de récupérer les données de plusieurs capteurs afin de mieux gérer les automatismes aidant la personne dans l'habitation ou de donner des informations sur son état de forme au cours du temps.

Il est à noter que tous ces travaux s'inscrivent également dans les cadres légaux actuels. La Réglementation Générale de la Protection des Données (RGPD) en place en Europe depuis mai 2018, a modifié l'utilisation qui peut être faite des données. Dans un contexte de recherche nous sommes moins contraints sur l'utilisation de ces données, mais toutes les réflexions faites dans ces travaux doivent l'être dans ce cadre afin de rendre possible une industrialisation.

1.2 Parcours professionnel

L'habilitation à Diriger des Recherches permet de faire un point d'avancement sur les travaux effectués au cours du début de carrière professionnelle. La figure 1.1 présente une vue synthétique de ce début de carrière de chercheurs en mettant l'accent sur deux points : les postes occupées et les accompagnements de thèses et de post-doctorats effectués.

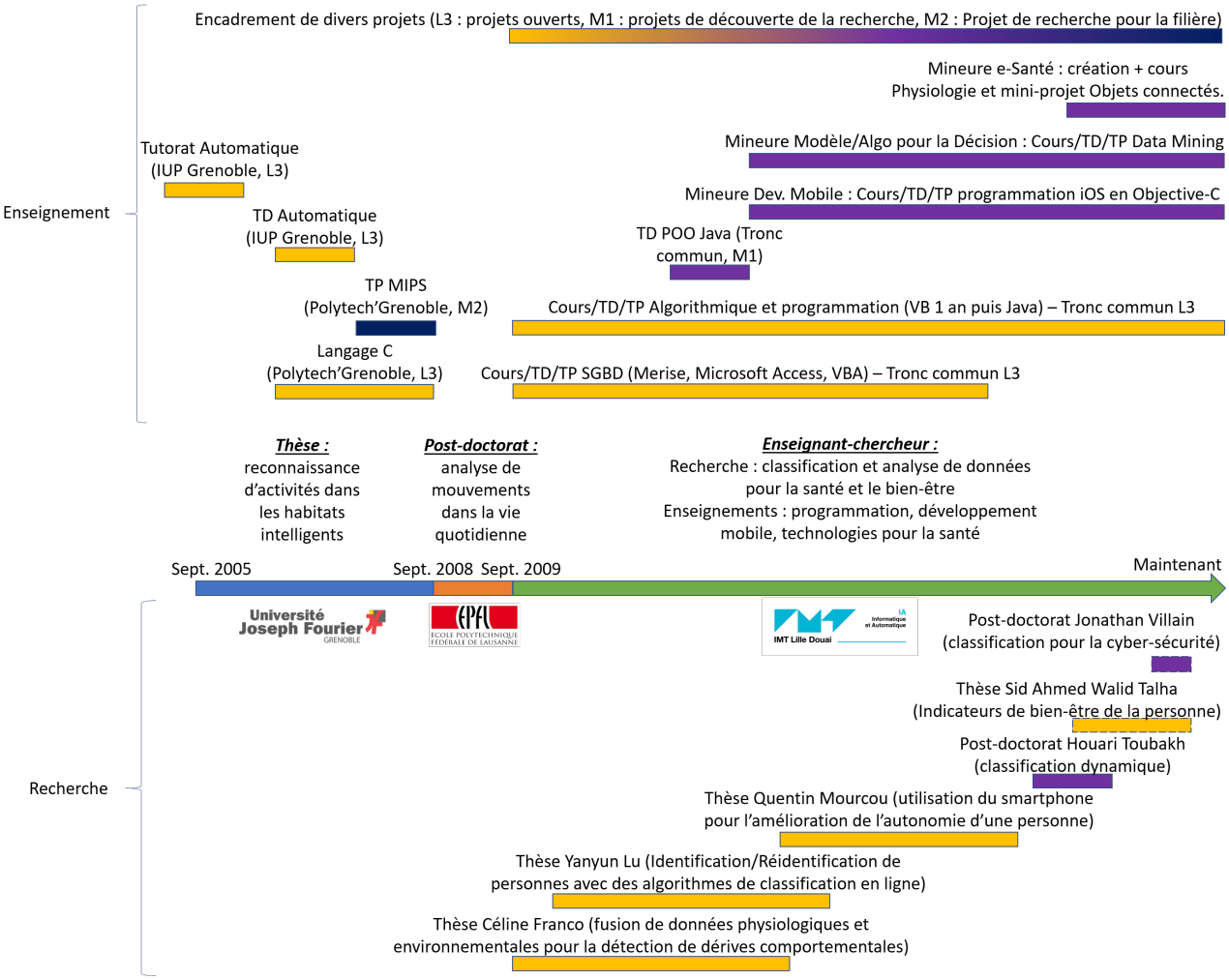


FIGURE 1.1 – Vue synthétique de mon parcours professionnel

À partir de cette expérience, ce manuscrit d'Habilitation à Diriger des Recherches a pour but de présenter les contributions faites après la thèse de doctorat et, comme expliqué par la suite dans l'organisation, il a été pris le parti d'analyser ces contributions à travers deux problématiques associées au processus de reconnaissance de formes, à savoir le traitement du signal pour l'extraction d'indicateurs et la classification de données. Ces deux problématiques, appliquées à des données relatives au comportement humain, seront mises en valeur dans trois chapitres en allant de la personne en elle-même avec ses automatismes internes jusqu'à l'analyse de ses activités dans son environnement afin d'en améliorer les interactions. Ces trois chapitres font intervenir des recherches dans deux champs disciplinaires très proches, à savoir le traitement du signal (plus attaché à la section CNU 61) et l'apprentissage automatique (CNU 27). L'implication de ces deux thématiques dans les différents travaux est résumée par la figure 1.2.

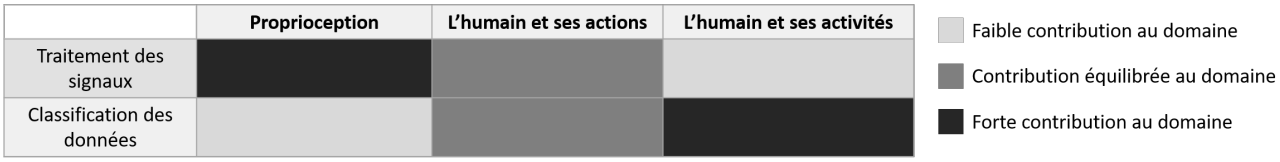


FIGURE 1.2 – Mise en relation des travaux avec les domaines (relatifs à des sections CNU)

1.3 Le traitement de données et ses applications

L'avènement des grands systèmes d'information et la performance sans cesse accrue des calculateurs font que l'analyse et le traitement des données s'appliquent maintenant à tous les domaines. De nouveaux métiers tels que « Data Scientist » ou « Architecte Big Data » ont commencé à se répandre et représentent maintenant des profils très recherchés. Mes travaux de recherche s'inscrivent dans cette dynamique qui m'a permis d'enseigner, de diffuser et de développer de nouveaux concepts autour du traitement de la donnée (de l'extraction à la reconnaissance de forme en passant par l'analyse). Ces travaux ont trouvé des applications dans plusieurs domaines portés par différents projets :

— Les objets connectés pour la santé. Dans ce domaine d'application, il importe d'obtenir des données fiables et maîtrisées dans toutes les conditions (par exemple mesurer un électrocardiogramme lorsque la personne fait sa vaisselle n'est pas la façon de faire la plus naturelle). Ces traitements de données prennent plusieurs formes. Le premier est la mise en forme correcte des données issues de capteurs embarqués avec un minimum de consommation d'énergie, afin d'en tirer une information complémentaire. Cela peut être par exemple la manière de réaliser un mouvement afin d'étudier les fonctions motrices d'une personne. La seconde forme est un traitement des données sur le long terme avec une extraction d'informations dans un signal important (sur des jours et des mois) et l'évaluation de son évolution afin de détecter, le plus précocement possible, l'apparition de problèmes potentiels qui pourraient être liés à un symptôme d'une maladie évolutive et/ou chronique. Cela nous rapproche de la tendance actuelle de la médecine 4P définie par Leroy Hood en 2013 et qui se veut prédictive, personnalisée, préventive et participative (FLORES et al., 2013). En fondant en plus la médecine de demain sur les preuves de l'apport au patient, nous arrivons même à une médecine 5P.

— L'analyse de données issues de réseaux de caméras de surveillance. Dans ces travaux, il est important de créer des algorithmes efficaces et capables de prendre un compte une grande masse de données, tout en en tirant une information pertinente, simple et fiable pour l'utilisateur. Cela peut aller de l'analyse de situations visuelles à la reconnaissance ou la ré-identification de personnes. Ces deux thématiques font donc intervenir autant la biométrie que la reconnaissance d'activités ou d'actions d'une personne seule ou dans un environnement avec plusieurs personnes. L'un des grands défis de ces travaux est de les mener dans un cadre respectant la vie privée des personnes et les différentes lois et normes en matière d'éthique.

— La sécurité, avec des travaux sur la détection de comportement ou d'utilisation anormale de ressources. Cela peut être des réseaux ou de la donnée par exemple et des algorithmes d'apprentissage capables d'identifier les comportements « normaux » et les déviations par rapport à ces comportements peuvent être d'une grande aide dans ce cas. Nous sommes confrontés ici à des données relativement importantes encore (lorsque ce sont par exemple des données spectrales) et dans lesquelles il va falloir trouver des utilisations anormales sans lever d'alarme de manière intempestive. Dans ce cas, des algorithmes incrémentaux semblent plus qu'adaptés, mais n'ont encore, dans la littérature, pas été utilisés pour résoudre ce type de problématique.

— L'analyse de données issues d'environnements ubiquitaires tels que les espaces/villes/habitats intelligents, qui inclut notamment des travaux portant sur l'analyse de comportement des personnes dans ces habitats (avec comme finalité, par exemple, le commerce ou la surveillance à visée médicale). Dans cette thématique, en plus d'avoir une masse de données importante, celle-ci est vraiment multimodale, avec des types de données souvent incompatibles en termes de période de mesures (lorsqu'ils en ont une et ne se déclenchent pas sur événements), de type de valeurs, de plage de variation, etc. La difficulté sera de mettre en forme correctement toutes ces données pour en faire un vecteur d'attribut qui soit informatif et qui réagisse bien pour la résolution du problème

posé.

— Des données relatives à l'amélioration des processus de production (usine du futur, industrie 4.0, etc.), dans lesquels il est possible d'analyser les données issues par exemple des chaînes de production ou bien les données issues de la vente pour améliorer les processus et la tenue de l'usine. Dans ces données, le plus gros défi est leur mise en forme alors que nous nous retrouvons dans un cas pratique où la mesure n'est qu'une chose accessoire pour l'usine et non leur but de base. Nous avons donc des données fortement bruitées, car contenant de gros problèmes de mesures, des problèmes de mise en relation des différentes données, des capteurs qui changent sans que ce ne soit plus documenté, etc. La réalisation d'algorithmes pour résoudre les problèmes de l'usine du futur et améliorer le processus de production est donc complexifiée par ceci et par l'évolution des gammes de production pour coller au mieux aux attentes changeantes des consommateurs.

Toutes ces thématiques sont donc abordées dans des projets divers afin de continuer à adresser des verrous scientifiques en lien avec les problèmes qui se posent actuellement. Les verrous et un résumé des objectifs que porte cette habilitation seront présentés dans la section suivante.

1.4 Objectifs des travaux

Les travaux que j'ai effectués dans un cadre de recherche et d'innovation visent à améliorer les connaissances spécifiques à différents métiers ou domaines par l'analyse de données non stationnaires [1], bruitées et/ou incomplètes. Le but est donc de récupérer la donnée et de la mettre en forme le plus efficacement possible (au niveau qualité comme au niveau temps de traitement) afin d'en ressortir une information la plus complète, la plus précise et la plus répétable possible. Ces traitements sont ensuite utilisés dans des applications relatives à l'analyse de comportement humain afin d'améliorer les connaissances sur ce domaine.
Les verrous scientifiques à lever sont les suivants :

1. Comment peut-on, dans certains signaux, extraire une information pertinente et calculer un indice qui prenne en compte la non-stationnarité des données et les caractères multi-modal et multi-échelles de celles-ci ?
2. Que peut-on proposer comme méthode d'extraction pour avoir des attributs performants et simples à obtenir ?
3. Comment peut-on utiliser au mieux les données présentes et les attributs extraits pour construire des modèles les plus performants possibles et qui généralisent au mieux les comportements que nous recherchons.
4. Comment créer au mieux des algorithmes de classification afin qu'ils soient les plus efficaces possible pour modéliser des phénomènes non stationnaires (et donc avoir des classifieurs évolutifs).
5. Comment créer une intelligence à partir des données de classification et d'autres entrées afin d'adapter un habitat (voir tout système intelligent) à son utilisateur (et non que l'utilisateur s'habitue au comportement de l'habitat/du système intelligent).

Ces cinq points seront abordés par les trois chapitres qui vont suivre.

1.5 Organisation du manuscrit

Ce manuscrit va se présenter en trois chapitres, dont un résumé pourrait être présenté par la figure 1.3 et qui contiendront les travaux suivants, tous relatifs à l'analyse de la personne ou de son comportement :

— Un premier chapitre (Chapitre 2, page 7) qui va se concentrer sur les automatismes « internes » de la personne. Pour ce faire, des travaux ont été réalisés sur l'analyse des capacités posturales et proprioceptives de la personne, avec : (1) des plateformes de force et une analyse entropique du signal permettant de déterminer les capacités intrinsèques d'une personne à gérer les automatismes du contrôle de la posture ; et (2) un smartphone pour mesurer les capacités proprioceptives soit en contrôle de la posture soit en reproduction de mouvements donnés et donnant lieu à la création d'indicateurs pour d'entraîner, par retour sensoriel, les capacités de la personne en toute autonomie à son domicile.

1. Comme ce sera par la suite rappelé, un processus est dit non-stationnaire si ses statistiques dépendent du temps

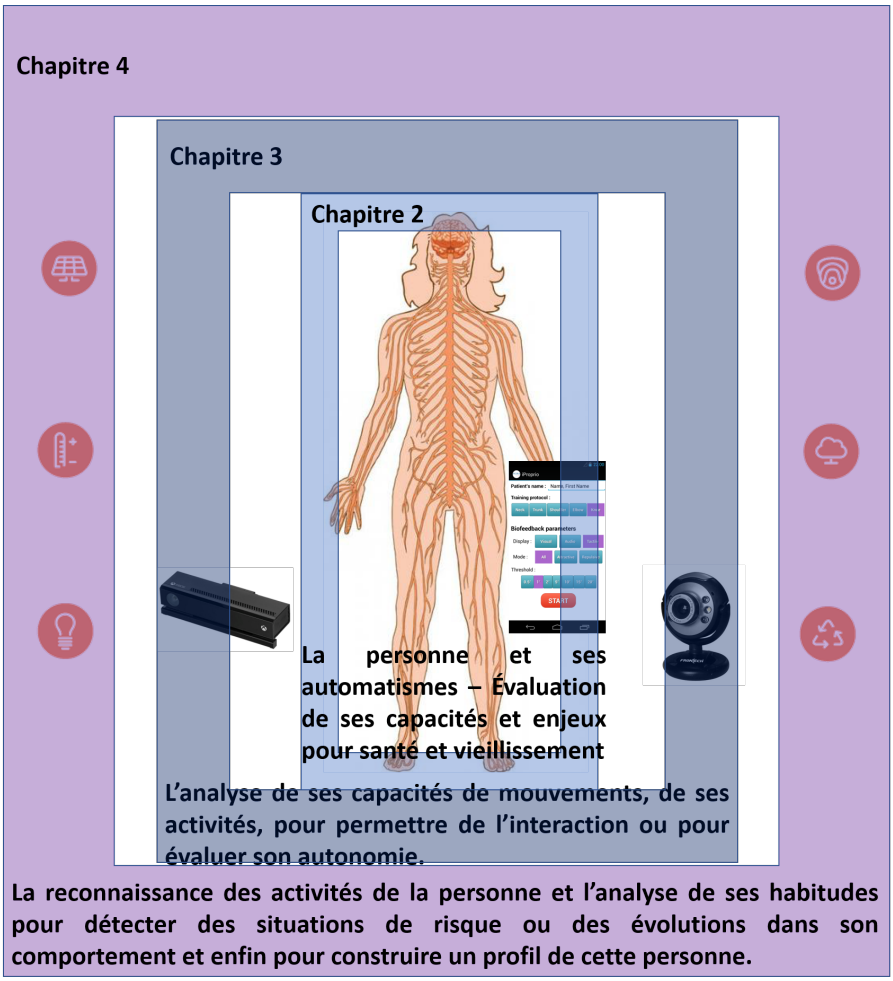


FIGURE 1.3 – Résumé synthétique des thématiques présentées dans chacun des trois chapitres de ce manuscrit, dans l'ordre croissant du plus intérieur au plus extérieur.

— Un second chapitre (Chapitre 3, page 35) qui va lui investiguer l'analyse de la personne dans son ensemble et de ses actions, tout ceci en vue d'inférer des informations sur son comportement. Ainsi, de premiers travaux ont porté sur la reconnaissance de personnes dans des réseaux de caméras avec des algorithmes de classification en ligne et les travaux suivants ont porté sur la reconnaissance d'actions d'une personne avec des caméras RGB-D en vue de permettre une interaction entre l'humain et le robot. Pour cela, ces derniers travaux se sont concentrés sur des algorithmes permettant une reconnaissance avec une orientation de la personne différente et une reconnaissance plus rapide (au cours de l'action).

— Un troisième chapitre (Chapitre 4, page 59) se concentrera sur la reconnaissance d'activités plus complexes, avec des capteurs supplémentaires par exemple dans un habitat intelligent. Des travaux ont porté d'abord sur la classification en elle-même puis sur une utilisation de cette classification afin de créer des profils utilisateurs. Ce chapitre reprendra donc ces deux parties qui suivent mes travaux de thèse sur le sujet et les ont approfondi par l'intermédiaire de plusieurs collaborations.

Le dernier chapitre (Chapitre 5, page 85), présentera les perspectives de ces travaux de recherche et conclura ce manuscrit. Ce chapitre contient ainsi des perspectives à court et long terme sur la thématique de l'analyse de comportements humains et l'utilisation qui pourra être fait par la suite de ces travaux.
Enfin, les annexes contiennent un résumé de ma carrière dans l'enseignement et la recherche sous forme d'un CV détaillé. Celui-ci contient les parties suivantes :

— Un curriculum Vitæ en Annexe A, page 105.

— Une présentation des activités d'enseignement, de mon implication dans ceux-ci ainsi que d'innovations pédagogiques en Annexe B, page 107.

— Un résumé de mes activités de recherche avec les principaux projets auxquels j'ai participé ou les projets en cours en Annexe C, page 113.

— L'implication dans la communauté scientifique et éducative en Annexe D, page 121.

— Une vision synthétique de mes indices de production et une liste complète de mes publications en Annexe E, page 125.

CHAPITRE 2

La personne et ses automatismes : autour de sa posture et de sa proprioception

Sommaire

1.1 Contexte des travaux et historique 1
1.2 Parcours professionnel 2
1.3 Le traitement de données et ses applications 3
1.4 Objectifs des travaux 4
1.5 Organisation du manuscrit 4

2.1 Introduction

La prolifération des capteurs, qui envahissent notre quotidien, leur baisse de prix, leur variété d'usage et les capacités de stockage qui augmentent (avec également l'augmentation des activités de collecte et de stockage en temps réel des données d'objets connectés sur des serveurs distants) font que nous sommes en mesure de récupérer un très grand nombre de données sur des temps longs et avec des fréquences suffisantes. Du fait de l'apparition des smartphones [1], « l'épidémie » de l'auto-mesure (Quantified Self) s'est vite développée (SWAN, 2013 ; SWAN, 2012 ; NAFUS et al., 2014). Ainsi, les personnes étant volontaires pour fournir les données relatives à leurs activités et même certaines données physiologiques, de grandes masses de données sont accumulées en vue de les traiter et d'en extraire de la connaissance, des indicateurs sur les personnes ou sur leurs habitudes de vie (DIAZ et al., 2012 ; TSAI et al., 2014 ; GUBBI et al., 2013 ; ZHU et al., 2015). Parmi les habitudes de vie, de plus en plus, les assureurs, entre autres, souhaitent mieux comprendre le comportement de leurs sociétaires. Ainsi, des applications dans ce domaine sont apparues avec par exemple une analyse, grâce aux signaux inertiels par exemple, du comportement des assurés afin d'évaluer le risque de les garder et ainsi possiblement de baisser leur prime si leur comportement est bon (HUSNJAK et al., 2015 ; HÄNDEL et al., 2014). La même chose est faite dans le cas d'assurances pour la santé en évaluant le taux d'activité de la personne (exemple (GIDDENS et al., 2016) sur des membres d'une grande banque). Des indications à visées commerciales sont également recherchées (WINTER, 2014).

L'amélioration des techniques de traitement des signaux a également permis de mettre en place des protocoles permettant la mesure ambulatoire, soit en auto-mesure, soit en mesure contrôlée et installée par le praticien puis emportée à la maison ensuite par la personne pour une plus longue durée.

Cette partie va porter sur cette thématique de la mesure et du traitement des données afin d'en extraire des informations pertinentes. Les types de données qui sont abordés sont les données physiologiques et inertielles. Ce chapitre commencera par un état de l'art des travaux puis une description de ce qui a été accompli lors des différents masters et des différentes thèses que j'ai pu encadrer sur le sujet.

1. Données de marché Smartphone de l'institut GfK de 2016 http://www.gfk.com/en-gb/insights/press-release/smartphone-market-up-66-percent-year-on-year-in-2016-1/ consultées en juin 2017.

2.2 Objectifs de ces travaux et verrous scientifiques à lever

Nous nous intéressons ici à des signaux extraits de mesures sur la personne, dans un but de donner des évaluations objectives sur ses aptitudes en termes de mobilité ou encore son autonomie. D'une manière générale, l'objectif ici est d'extraire un maximum de connaissance et un maximum d'information d'un signal reflétant les capacités d'équilibre d'une personne (plateforme de force ou centrales inertielles) et ce pour en déduire des informations pertinentes sur les capacités de la personne (et l'évolution de ces capacités), ses automatismes etc. Par exemple, nous savons, qu'avec l'âge notamment, certaines activités vont nécessiter plus de ressources cognitives pour être effectuées et ainsi la personne aura plus de mal à gérer des tâches concurrentes dans sa vie courante. Le but de ces travaux est de trouver des mesures objectives permettant de quantifier le vieillissement ou la dégradation sensorielle d'une personne par des exercices simples qui peuvent être fait de manière autonomes et à partir de ceci de proposer des exercices permettant de palier à ces dégradations (adapter la personne à ses capacités actuelles).
Les verrous scientifiques adressés ici sont les suivants :

— Comment peut-on, dans certains signaux, extraire une information pertinente et calculer un indice qui prenne en compte la non-stationnarité des données et les caractères multi-modal et multi-échelle de celles-ci ? Cette partie sera abordée avec l'entropie d'échantillon et l'entropie spectrale. Nous avons, lors de la thèse de Céline Franco, chercher à caractériser des mesures à l'aide de ces traitements non-linéaires sur les signaux et es indicateurs extraits.

— Quelles informations peut-on obtenir de la mesure inertielle d'une personne dans un certain contexte d'activité ou de posture statique, quelle information cela nous donne-t-il sur son bien-être ou sa santé et peut-on lui proposer des exercices pour améliorer son quotidien ? La thèse de Quentin Mourcou s'est intéressée aux informations que nous pouvons extraire en temps-réel sur ces signaux afin de guider la personne dans des exercices.

Avant la présentation de ces travaux, une partie état de l'art présentera les différentes données et l'état d'avancement des recherches précédant ou s'effectuant en même temps que nos travaux.
Les travaux seront ensuite présentés en introduisant tout d'abord le concept d'entropie d'échantillon puis en montrant son application aux données issues de plateformes de force pour en déduire la demande attentionnelle requise par le maintien de la posture. Ensuite, l'entropie spectrale ainsi que son application pour des indices locaux de complexité sur des données posturales seront introduits. Enfin, les travaux sur les données smartphone avec notamment les projets iBalance et iProprio seront présentés avec leurs résultats.

2.3 État de l'art sur la mesure de données physiologiques et de données inertielles

Dans cette première partie du chapitre, nous allons résumer rapidement l'état de l'art sur certaines mesures cardio-respiratoires et sur les mesures inertielles, ainsi que sur les outils de mesures disponibles. Cela nous montrera les outils technologiques mais aussi méthodologiques disponibles.

2.3.1 Les données cardio-respiratoires

Deux fonctions vitales régissent notre vie, à savoir la circulation du sang et son oxygénation. Ainsi, la respiration et le rythme cardiaque sont des notions vitales et leurs interactions sont fondamentales d'un point de vue de la physiologie clinique. Du fait d'un accroissement de la prise en compte du bien-être du patient, ainsi que du fait des avancées en termes d'outils de mesure portable et embarquable sur la personne, le développement de techniques de mesure non invasives s'accentue. C'est le cas de la « pléthysmographie respiratoire par inductance » (JALLON et al., 2013).
Le signal cardio-respiratoire (CR) mesuré par la pléthysmographie par inductance montre les interactions entre les systèmes vasculaires et respiratoires. Cela peut être modélisé par la superposition de deux composantes à savoir les activités des oscillateurs cardiaques et respiratoires centraux. Cela mène à une non-stationnarité typique du signal caractérisée par une hiérarchie spectrale de modes d'oscillations avec une faible modulation en amplitude. Se pose

alors la question d'arriver à trouver une représentation capable de distinguer ces deux composantes embarquées dans le signal. Ainsi, des méthodes récentes de traitement du signal, qui se caractérisent par leur caractère adaptatif et local, se démarquent pour de telles applications. Parmi celles-ci, la Décomposition Empirique Modale (Empirical Mode Decomposition – EMD dans la suite de ce manuscrit) (HUANG et al., 1998) est intéressante, car elle est uniquement portée par les données et ne fait aucune hypothèse sur une quelconque fonction de base ou sur des plages fréquentielles ou des échelles temporelles. Plus précisément, l'EMD regarde les valeurs extrêmes du signal pour le décomposer en différents modes. Même si cette méthode a été appliquée avec succès sur de nombreux signaux physiologiques par le passé (AL-ANI et al., 2011; ANAS et al., 2011; PACHORI et al., 2011; KARAGIANNIS et al., 2011), le fait qu'elle ne repose pas sur des bases analytiques solides limite sa compréhension et peut la rendre non prédictible.
Une seconde méthode, le synchrosqueezing (SQ) (DAUBECHIES et MAES, 1996) combine l'analyse en ondelettes et la réallocation. La finalité de cette méthode est d'affiner la représentation temps-échelle d'un signal en réallouant les coefficients de la transformée en ondelettes continue en fonction de comportements locaux de la représentation. Cette méthode est différente des méthodes de réallocation classiques, car elle permet une reconstruction exacte du signal à partir des composantes extraites.
Quelle que soit la représentation choisie, la connaissance de la hiérarchie spectrale imposée par la physiologie humaine du signal cardio-respiratoire se révèle être un réel avantage pour estimer l'importance, dans le signal, des composantes trouvées par ces représentations. Pour utiliser au mieux ces connaissances, la représentation qui nous intéresse le plus doit être cohérente avec l'organisation spectrale du signal et doit aussi avoir comme résultat l'extraction de peu de modes afin de pouvoir interpréter leur « sens physiologique ».
Que ce soit l'EMD ou le SQ, ces deux méthodes ont montré leur capacité à extraire des caractéristiques de signaux bimodaux (RILLING et al., 2008; WU et al., 2011). Elles ont également été comparées ce qui a montré que le SQ obtenait des résultats meilleurs et plus prédictibles. Par ailleurs, les résultats sur ces signaux dépendent du ratio d'amplitude et de fréquences entre les composantes. Dans une étude faite en marge de sa thèse (FRANCO, FONTECAVE-JALLON et al., 2011), Céline Franco a étudié la capacité de l'EMD à extraire des composantes de plus faibles amplitudes dans le signal cardio-respiratoire (simulé), à savoir le signal cardiaque.
En considérant tous ces travaux, le but poursuivi dans la thèse de Céline Franco était d'étudier ces méthodes sur le signal cardio-respiratoire puis sur le signal postural afin d'améliorer la compréhension que nous en avions. Cette petite introduction a eu pour but de montrer comment les recherches se sont tournées vers des méthodes de synchrosqueezing. Cependant, par souci de clarté et de cohérence, ces travaux seront présentés dans le cadre du signal postural et non du signal cardio-respiratoire par la suite. La publication (FRANCO, GUMÉRY et al., 2012) [2] résume cette partie de contribution.

2.3.2 Les données de posture et de marche

Dans cette partie de l'état de l'art, nous allons revenir sur les systèmes de mesures permettant d'évaluer l'équilibre ou la posture d'une personne, puis sur ceux permettant d'évaluer les paramètres de la marche afin d'en déduire des informations sur la personne.

2.3.2.1 Le système de contrôle de l'équilibre

Le signal cardio-respiratoire, est compliqué, mais reste limité à deux modes qui sont la superposition des deux phénomènes (dépendants) que sont l'activité cardiaque et l'activité respiratoire de la personne. Le signal postural, quant-à-lui, va dépendre de plus de composantes. Le système postural ainsi que toutes ses composantes se chargent de contrôler la posture dans différentes positions, qu'elle soit statique ou lors du mouvement par exemple. Ce système postural a trois composantes : le système sensoriel (donnant au corps « l'état » dans lequel il se trouve), le système nerveux central et le système musculo-squelettique (BOUISSET et al., 1995; KEJONEN, 2004).
Pour assurer la stabilité posturale, ce système va répondre au plus vite aux perturbations afin de maintenir la position voulue (les perturbations peuvent être faibles (dues à l'activité du système cardio-respiratoire) ou plus importantes

2. Cette présentation dans une couleur différente des autres citations sera conservée, dans cette première partie, pour les articles pour lesquels je suis l'un des co-auteurs.

(dues à des forces extérieures, du fait de la nature ou de l'environnement de la personne). Ainsi, des oscillations plus ou moins faibles autour du point d'équilibre voulu pourront être observées. La stabilité fait intervenir les capteurs sensoriels amenant l'information (par des voies afférentes) au système nerveux central, qui lui même intègre ces données pour piloter les effecteurs que sont les muscles. L'apprentissage de la posture debout, de la posture assise ou de la marche par un enfant va donc correspondre au réglage de ces systèmes et à un apprentissage des pondérations des différentes informations reçues.

En prenant en compte la complexité de ce système et le fait qu'il possède de nombreuses composantes, nous devons ainsi adapter les méthodes de traitement. L'entropie est le moyen le plus répandu pour quantifier le désordre et donc la complexité. Les approches statistiques utilisées pour les signaux physiologiques, par exemple l'entropie (et plus particulièrement ici la l'entropie d'échantillon – SEn), montrent quand même un manque de reproductibilité et de fiabilité (PENG et al., 2009). Cela pose la question de la prise en compte de phénomènes se présentant sur différentes échelles. Des méthodes ont été proposées pour intégrer ces multiples échelles dans le calcul de l'entropie. Dans le cas de signaux physiologiques caractérisés par une distribution spectrale contenant plusieurs modes, tels que la locomotion et la posture pour le maintien de l'équilibre, une méthode appelée « Intrisic Mode Entropy » (IME) a été proposée. Elle combine la décomposition empirique modale avec la l'entropie d'échantillon (AMOUD et al., 2007). Cette approche amène par contre les contreparties de l'utilisation de l'EMD avec ses désavantages (phénomène de mélange ou de moyennage de modes avec des contenus fréquentiels différents, etc.) et ceux de l'utilisation de l'entropie d'échantillon (réglage des hyper-paramètres).

D'un point de vue des systèmes dynamiques, les processus complexes sont associés à des contenus fréquentiels larges, offrant une richesse d'information (contrairement aux phénomènes périodiques). L'hypothèse de Goldberger est que le vieillissement et la maladie sont caractérisés par une perte de richesse spectrale. La rééducation dépend, elle, de la capacité du système à retrouver ce contenu fréquentiel large (GLEICK, 2011 ; GOLDBERGER, 1996). Estimer la complexité de ce système dynamique pourrait consister en une quantification de cette richesse spectrale. Une partie des travaux présentés dans la section 2.4 portera sur l'analyse des signaux du centre de pression (CoP) acquis sur une plateforme de force, et pour lesquels, en se basant sur une approche nouvelle à l'entropie par ondelettes (Wavelet Entropy) (ROSSO et al., 2001), qui s'appuie sur le synchrosqueezing, nous proposons un indice original rendant compte de la variabilité dans le spectre du signal postural.

Cependant, sur ces dernières années, nous ne nous sommes pas limités au signal issu de plateformes de force pour caractériser l'équilibre de la personne. De nombreuses sociétés proposent des Unités de Mesure Inertielles (IMU pour Inertial Motion Unit) permettant d'acquérir et de faire certains traitements sur les données des accéléromètres, magnétomètres et gyroscopes tels que XSens, Gaitup, etc. (ROETENBERG et al., 2009 ; DADASHI et al., 2013). Cependant, ces unités sont en général coûteuses (y compris en termes de temps de développement) et difficiles à installer sur d'importantes cohortes (ou pour prévoir des utilisations à grande échelle). Dans cette optique, nous nous sommes tournés vers un outil du quotidien, possédant le même type de capteurs avec les mêmes précisions, connecté, et sur lequel nous avons toute latitude pour programmer des algorithmes complexes du fait de leur puissance de calcul : le smartphone.

2.3.2.2 Outils de mesure et données smartphone

Les smartphones sont devenus des outils incontournables du quotidien dans nos populations développées, et sont devenus une part importante de notre vie. Les ventes continuent d'augmenter et bien que le marché dans certains pays ne soit pas loin d'être saturé, le renouvellement se fait régulièrement [3]. Notre smartphone provoque également, d'après des résultats de recherche récents, une charge attentionnelle supplémentaire du fait uniquement de savoir sa présence et le fait qu'il puisse nous notifier d'une activité dans les prochains temps (WARD et al., 2017).

Ces outils sont de plus en plus accessibles tout en étant composés de matériels puissants et permettant une connectivité avancée et une rapidité de traitement. Au niveau des capteurs, nous pourrions être intéressés par trois d'entre eux à savoir, comme dit précédemment, l'accéléromètre, le magnétomètre et le gyroscope. L'accéléromètre est présent dans

3. Données de marché Smartphone de l'institut GfK de 2016 http://www.gfk.com/en-gb/insights/press-release/smartphone-market-up-66-percent-year-on-year-in-2016-1/ consultées en juin 2017.

tous les modèles. Le gyroscope se retrouve lui dans la plupart des modèles (sauf premiers prix). L'IMU complète est tout de même réservée aux modèles de gamme moyenne minimum.

L'avantage du smartphone est aussi de contenir des moyens d'interagir avec la personne tels que l'écran, le système audio ou un dispositif de vibration contrôlable pour en faire un système haptique. La combinaison de tous ces systèmes de captation et d'interaction fait que nous pouvons intégrer, à ce smartphone, des applications développées spécifiquement pour nos besoins.

Depuis quelques années, les smartphones sont devenus des outils utilisés en recherche, et notamment dans le domaine médical. Par exemple, le smartphone a déjà été utilisé pour surveiller le rythme cardiaque dans le cas de réhabilitation suite à des crises cardiaques (VARNFIELD et al., 2014) ou pour le suivi d'exercices (WORRINGHAM et al., 2011). Certains algorithmes se basant sur l'analyse d'images (et utilisation du flash) ont même montré la faisabilité de la détection de certaines anomalies dans le rythme cardiaque (LEE et al., 2013). En lien avec nos travaux sur la posture et le mouvement, le smartphone peut être utilisé pour la mesure de l'amplitude de mouvements et la mesure de l'angle d'une articulation pour le contrôle de la posture ou de la marche (KOSSE et al., 2015; MILANI et al., 2014). Ces applications sont conçues pour fournir une mesure précise et reproductible de l'amplitude de mouvement en comparaison à un outil standard. Les applications smartphones les plus courantes utilisent principalement des algorithmes fournis par le développeur du système d'exploitation, qui procède à une fusion optimale des trois capteurs pour calculer l'angle. Bien que ces mesures aient été comparées à des outils standard, la mesure d'angle par le smartphone dans un contexte tant statique que dynamique n'avait pas encore été évaluée. Une partie des travaux de thèse de Quentin Mourcou a validé cette mesure en comparaison avec le robot Kuka, permettant de reproduire des mouvements de référence, avec différentes vitesses et différentes amplitudes, en statique et en dynamique. Ces contributions seront présentés dans la section 2.6.

Les applications basées sur la mesure apportent des innovations sur la compréhension des activités et des actions de la personne afin d'ajouter une partie intelligence ambiante au téléphone pour permettre d'apporter des aides spécifiques aux patients (CAMPBELL et al., 2012) et d'enrichir l'environnement avec des technologies intelligentes (D. J. COOK et al., 2012). Les systèmes (ubiquitaires) qui se basent sur ces concepts d'intelligence ambiante doivent être sensibles, avoir un temps de réponse court, être adaptatifs et transparents pour la personne. Le smartphone joue un rôle primordial dans ce type de système du fait de son intégration naturelle dans notre vie quotidienne. De nos jours, la fusion d'informations avec des algorithmes intégrés au smartphone peut être utilisée pour mesurer les mouvements humains lors de ses activités de la vie quotidienne (ANGUITA et al., 2013), et ce, pour capter les amplitudes de mouvements, ou pour évaluer le contrôle de la posture et de la marche (KOSSE et al., 2015), la goniométrie (MILANI et al., 2014) et même pour la détection de chute (HABIB et al., 2014; GUVENSAN et al., 2017). L'accéléromètre est utilisé par tous les systèmes de détection de chute basés sur des smartphones, mais, dans les années précédentes, sa plage de mesure pouvait se révéler insuffisante (ABBATE et al., 2012). La qualité des capteurs intégrés est cruciale en utilisation clinique. Certains smartphones qui ne possèdent encore que l'accéléromètre (typiquement les téléphones Android premier prix) permettent de faire certaines mesures d'angle en statique, mais les mesures d'angle en dynamique sont finalement peu fiables. Pour obtenir des mesures fiables, WILLIAMSON et al. (2001) ajoutent le gyroscope qui est insensible à la gravité (mais possède le défaut d'avoir une dérive due à l'intégration du signal). Pour mesurer la composante horizontale de l'orientation, des équipes ont combiné l'accélération et la mesure du champ magnétique terrestre pour suivre la position du corps (MITCHELL et al., 2014; JENNY, 2013; VOHRALIK et al., 2015).

Dans le cadre de la thèse de Quentin Mourcou, une revue de l'état de l'art complète sur les travaux concernant les smartphones et la mesure d'angle à visée médicale a été faite (MOURCOU, FLEURY, DIOT, FRANCO et al., 2015). Dans cette revue de littérature, nous résumons vingt sept différentes expérimentations qui ont testé la mesure avec cet outil dans le cas de la proprioception du genou, de l'épaule, de la cheville, etc. Cela montre l'intérêt pour les chercheurs et les industriels pour ce récent outils ubiquitaire de la vie quotidienne.

Dans la suite de ce chapitre, nous allons nous intéresser au calcul de l'entropie d'échantillon sur les données de plateforme de force afin d'en déduire des informations sur la demande attentionnelle du maintien de la posture pour la personne.

2.4 Analyse des signaux posturaux avec l'entropie d'échantillon

2.4.1 L'entropie d'échantillon

Le terme entropie a été pour la première fois utilisé en 1865 par Rudolf Clausius. Ce terme d'entropie caractérise, pour un système donné, le degré de désorganisation (en physique par exemple) ou d'imprédictibilité du contenu en information (définition se rapprochant de celle de Shannon) de celui-ci.
À l'origine, la mesure de la régularité d'un signal se faisait par un calcul exact qui se basait sur diverses mesures d'entropie (exemple celle de Kolmogorov- Sinai). Cependant, le calcul exact de l'entropie requièrt un grand nombre de données pour avoir une mesure valable (il se base sur des distributions statistiques qui doivent être correctement représentées dans le signal) et le résultat sera de plus sensible au bruit. Sur des données expérimentales, il est donc compliqué en pratique d'appliquer ces calculs. L'« Entropie Approchée », proposée par PINCUS (1991), tente de contourner ce problème en proposant une mesure différente de la régularité. Cette approche est devenue populaire notamment dans le cadre de mesures de données médicales comme le rythme cardiaque. Par la suite, Richamn et Moorman ont proposé une amélioration de cette approche qu'ils ont nommé Entropie d'Échantillon qui corrige le biais introduit par la prise en compte de l'auto-correspondance des modèles (la probabilité de contribution des modèles devant être non nulle) et qui est aussi moins dépendant de la taille des données (RICHMAN et al., 2000). Depuis, l'entropie d'échantillon a été appliquée sur de nombreuses séries temporelles physiologiques (variabilité du rythme cardiaque (GOVINDAN et al., 2007), posture (LAMOTH et al., 2011; MURRAY et al., 2017), variabilité de la force et déplacement du centre de pression pendant le contrôle de la posture (BORG et al., 2010; DONKER et al., 2007; MADELEINE et al., 2011; STINS et al., 2009; ISABLEU et al., 2017; FREEMAN, 2017). Plusieurs études portant sur le contrôle de la posture lors de changement de charge cognitive pendant l'exécution de la tâche (protocole incluant une double tâche avec une seconde très demandeuse cognitivement) suggèrent que l'entropie d'échantillon peut donner une indication sur l'investissement attentionnel (DONKER et al., 2007; ROERDINK et al., 2010; SCHMIT et al., 2005; ISABLEU et al., 2017) : moins la (ou les) tâche(s) requier(en)t d'attention, plus l'entropie est grande (en effet, dans ce cas, la tâche sera gérée de manière plus « automatique » ce qui donnera un comportement imparfait et avec un contenu fréquentiel plus important, augmentant ainsi l'entropie car diminuant la prédictibilité). De ce fait, nous pourrions nous attendre à ce qu'un individu en pleine santé ou un expert du domaine obtienne une valeur d'entropie plus grande, synonyme d'automaticité et d'adaptabilité. De plus faibles valeurs d'entropie, qui rendent compte alors d'une régularité, sont associées à des mouvements contenant des modèles répétés sans possibilité d'adaptation de la personne aux nouvelles requêtes dues aux changements de conditions, et s'observent chez des personnes avec des déficiences ou chez des personnes âgées. Il peut cependant arriver que l'interprétation de l'entropie d'échantillon ne soit pas si simple et doive être complétée par des mesures plus traditionnelles du mouvement (BORG et al., 2010). Ainsi, l'entropie d'échantillon du centre des pressions peut être complétée par la vitesse et l'accélération du centre des pressions.
Le calcul de l'entropie d'échantillon se fait de la manière suivante :

1. La série temporelle est normalisée (moyenne nulle et variance unitaire) pour rendre la mesure indépendante de la taille et de l'échelle.
2. Des vecteurs de taille m et $(m+1)$ avec un retard temporel de 1 sont construits avec des points consécutifs : $z_k = (x_k, x_{k+1}, ..., x_{k+m-1})$ pour $1 \leq k \leq N-m$ et $w_k = (x_k, x_{k+1}, ..., x_{k+m-1}, x_{k+m})$ pour $1 \leq k \leq N-m-1$ (x étant le signal considéré).
3. Pour $1 \leq i \leq N-m$, la probabilité $B_r^m(i)$ qu'un autre vecteur de taille m soit égal à z_i avec une tolérance de r% de la variance de la série temporelle est calculée avec sa moyenne $B_r^m = \frac{1}{N-m}\sum_{i=1}^{N-m} B_r^m(i)$: $B_r^m(i) = \frac{1}{N-m+1}\sum_{j=1,\ j\neq i}^{N-m} \Omega(r.SD - \max_{0\leq k\leq m-1}|x_{j+k} - x_{i+k}|)$ où Ω est la fonction de Heaviside. En d'autres termes, B_r^m est la probabilité que deux différentes séquences de m points correspondent. De la même manière, les $A_r^m(i)$ sont calculés pour les vecteurs de dimension $(m+1)$ w_i. A_r^m est la probabilité que deux séquences différentes de taille $(m+1)$ correspondent.
4. Finalement, pour N suffisamment grand, $SEn_{CoP}(m, r, N) = -\log(CP)$ avec $CP = A_r^m / B_r^m$. C'est la négation

du logarithme naturel de la probabilité conditionnelle qu'une séquence de point de taille N s'étant répétée avec un intervalle de tolérance de r pour m points se répétera également pour $m+1$ points, et ce sans auto-correspondance (RICHMAN et al., 2000). Les choix des hyper-paramètres m ($m=3$) et r ($r=0.05$) sont basés sur une procédure d'optimisation décrite par Ramdani (RAMDANI et al., 2009). l'entropie d'échantillon est calculée pour plusieurs couples m,r avec m variant de 1 à 6 et r de 0,05 à 0,5. Un sous ensemble de valeurs de m a été testé pour observer le comportement avec un critère de pseudo-convergence de l'évolution de la valeur médiane de l'entropie d'échantillon en fonction de la valeur de r (pour un m donné). Ensuite, le r minimisant l'erreur relative maximale pour l'estimation de l'entropie et du CP, que l'on nomme $Q(m,r)$, est choisi. Avec $Q(m,r) = \max(\frac{\sigma_{CP}(m,r)}{CP(m,r)}, \frac{\sigma_{CP}(m,r)}{-\log(CP(m,r))CP(m,r)})$ où σ_{CP} est une estimation de la variabilité de la probabilité conditionnelle CP.

2.4.2 Application au contrôle postural

2.4.2.1 L'entropie d'échantillon pour caractériser un signal de posture

Le contrôle postural et le maintien de la posture nécessitent la mise en œuvre de plusieurs systèmes sensoriels : le système vestibulaire, le système somatosensoriel et la proprioception. Toutes ces informations sont recueillies et intégrées par le système nerveux central et ceci résulte en une régulation automatique de l'équilibre à travers plusieurs boucles de rétroaction. Pour les personnes jeunes et en bonne santé, maintenir une posture droite est une tâche aisée qui requiert une demande attentionnelle faible (du fait de l'automaticité et des « bonnes valeurs de lecture des capteurs »). Cependant, dans des conditions plus difficiles telles que lorsqu'il y a des tâches simultanées demandant une forte implication cognitive, et/ou avec l'absence de certaines entrées sensorielles (sol non ferme, aucune information visuelle, etc.), le contrôle postural demande une plus grande implication pour éviter la chute. Le premier cas, connu sous le nom de paradigme de la double tâche, est très courant dans la vie de tous les jours et a été investigué à de nombreuses reprises. Dans le second cas, une personne jeune et en pleine forme devrait être capable de s'adapter et de compenser la perte de l'un des capteurs en repondérant la prise en compte des différentes afférences sensorielles inconsciemment, alors qu'une personne plus âgée ou avec une déficience devra mettre en œuvre plus de ressources sensorielles aux dépens de l'automaticité.

Les mécanismes sensorimoteurs et cognitifs impliqués dans le contrôle de l'équilibre ont été étudiés, jusqu'à récemment, par le calcul de mesures conventionnelles se résumant à la statistique du déplacement du centre des pressions (surface de l'ellipse à 95%, longueur du « parcours », moyenne et variance de l'amplitude, vitesse, etc.). Ces mesures fournissent une vision globale du tangage de la personne autour de sa position d'équilibre et donnent une indication sur la stratégie musculaire mise en œuvre (hanche/genou/cheville). Cependant, elle ne donne pas d'indication ni sur la demande attentionnelle allouée à ce contrôle ni sur la stratégie de contrôle mise en place. Approfondir la compréhension du mode de contrôle et des mécanismes sous-jacents a poussé les chercheurs à faire appel à des outils d'analyse plus performants et moins génériques. Les méthodes non linéaires ont montré leur efficacité dans l'analyse des systèmes dynamiques et du déplacement du centre des pressions pendant un maintien de la posture debout (RQA (SCHMIT et al., 2005 ; RILEY et al., 2005), DFA, Lyapunov exponent (DONKER et al., 2007), etc.). Parmi ces paramètres, l'entropie est intéressante pour quantifier la régularité du signal.

Le but des travaux de Céline Franco dans cette partie de sa thèse était d'étudier la pertinence de l'entropie calculée sur le centre des pressions d'une personne afin d'en déduire des informations sur le coût attentionnel de la tâche de contrôle postural effectué. Pour cela, une expérimentation avec des personnes jeunes et en pleine santé sera présentée. Ces personnes ont effectué une tâche de posture statique (se tenir droit et aussi immobile que possible avec les yeux fermés). Du fait de l'importance des sensations somatiques dans le maintien de la posture bipédique, cela a été effectué dans trois conditions. Une condition normale, une condition où le système somato-sensoriel était altéré et une condition où il a été amélioré (au niveau de la cheville et du pied).

Nous voulons vérifier les hypothèses suivantes :

— Le fait d'altérer les sensations somatiques au niveau du pied et de la cheville vont augmenter les variations du centre des pressions (vitesse et déplacements) et diminuer son entropie,

— Le fait de l'améliorer va au contraire diminuer les variations et également augmenter l'entropie.

Pour rappel, une plateforme de force donne comme indication le déplacement du centre de gravité de la personne mesuré au niveau de ses pieds. À partir de ce déplacement, nous pouvons analyser la répartition spatiale des positions, mais aussi regarder la vitesse, les fréquences en jeu etc.
Dans la littérature, l'utilisation de mousse pour dégrader le contrôle bipédique de la posture a déjà été étudié pour vérifier les effets sur la régularité et la prédictibilité du centre des pressions, cependant, l'utilisation de l'entropie d'échantillon sur ce paramètre (SCHMIT et al., 2005; BORG et al., 2010; STINS et al., 2009; STRANG et al., 2010) dans des conditions de dégradation et d'amélioration des sensations somatiques et l'effet de ces changements sur la charge attentionnelle n'ont eux pas été étudiés. C'est le but des résultats qui sont présentés dans la section suivante.

2.4.2.2 Expérimentations et analyse statistique

Vingt étudiants (10 sur chaque expérimentation) ont participé à deux campagnes de mesure différentes. Comme pour toutes les expérimentations, ils étaient d'une part d'accord avec les conditions de mesures qui leur ont été présentées et n'étaient également pas au courant de la finalité de la mesure ni de ce que nous voulions exactement extraire comme information afin de ne pas biaiser les résultats. Ils étaient également « sains » dans le sens où ils n'avaient aucun trouble moteur ou neurologique connu.
Les yeux fermés, les personnes devaient se tenir debout sur une plateforme de force dans une position standardisée (les pieds sortant de 30°et les chevilles séparées de trois centimètres, avec les bras pendants le long du corps). Nous leur demandions de rester aussi immobile que possible.
La première expérimentation consistait à exécuter cette tâche sur deux surfaces différentes (ferme et mousse). La plateforme de force a été utilisée comme plan ferme. Dans la condition avec la mousse, un support de deux centimètres d'épaisseur en mousse a été ajouté sur celle-ci. Cela a pour effet de diminuer la qualité des informations somato-sensorielles de la plante des pieds et de la cheville. Dans la seconde expérimentation, la personne devait porter une orthèse pied-cheville afin d'améliorer sa stabilité. Il y avait donc une condition avec cette orthèse et une condition sans cette orthèse afin de vérifier l'utilité de celle-ci.
Pour ces deux expérimentations, l'ordre de présentation des conditions était aléatoire et pour chacune des conditions trois essais de 32 secondes chacun étaient réalisés. La plateforme de force servant à faire la mesure était une plateforme (Equi+, modèle PF01) qui est un plateau triangulaire de 80 centimètres de côté en aluminium posé sur trois capteurs de charge uniaxiaux (mesure entre 0 et 50 daN). Le signal de la plateforme de force est échantillonné à $64Hz$. Les données ne sont pas filtrées dans cette étude (RAMDANI et al., 2009) (l'impact réel du filtrage est compliqué à estimer sur ces traitements). Les séries temporelles sur le plan antéro-postérieur ainsi que sur le plan médial-latéral ont été centrées avant le calcul de la série temporelle de la résultante (RD), qui est le vecteur distance au centre du posturogramme.
Pour estimer les performances posturales, nous avons tout d'abord calculé le RMS du déplacement du centre des pressions (A_{CoP} en mm) ainsi que le RMS de la vitesse du déplacement (V_{CoP} en mm/s) : $A_{CoP} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} RD_i^2}$ and $V_{CoP} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}(\frac{RD_{i+1}-RD_i}{\delta t})^2}$ avec $\delta t = 0.016s$ et N le nombre d'échantillons.
Ensuite, pour nous donner des indications sur la structure dynamique et plus particulièrement sur la régularité et la prédictibilité des trajectoires, nous avons calculé également l'entropie d'échantillon comme présenté précédemment.
La moyenne des trois essais effectués dans les deux conditions a été utilisée pour faire des analyses statistiques. Un test de Kolmogorov-Smirnov d'égalité des variances a rejeté l'hypothèse de non-normalité des distributions ($p_s > 0.05$).
Pour comparer ensuite les données de chaque test entre les deux conditions, nous avons utilisé un t-test bilatéral avec un niveau de significativité à $0,05$.

2.4.2.3 Résultats

2.4.2.3.1 Expérimentation 1

En analysant la variabilité du signal, A_{CoP}, nous remarquons bien une plus petite valeur sur sol ferme que sur mousse ($t = -5.07$, $p < 0.001$, Fig.2.1a). Idem pour la vitesse ($t = -6.54$, $p < 0.001$, Fig.2.1b). Pour l'entropie, comme attendu, nous avons bien le comportement inverse ($t = 3.18$, $p < 0.05$, Fig.2.1c).

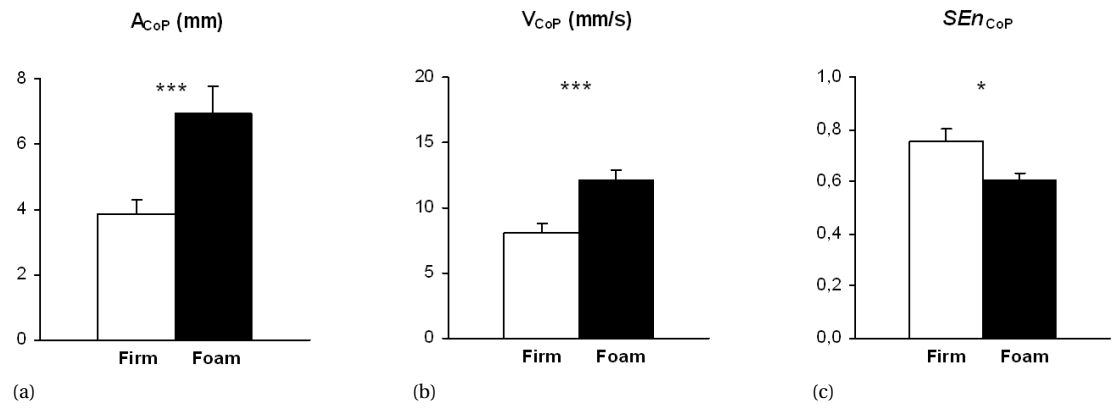


FIGURE 2.1 – Moyenne et variance de la variabilité (A_{CoP}) (panel A), de la vitesse (V_{CoP}) (panel B), et de la régularité (SEn_{CoP}) (panel C) de la trajectoire du centre des pressions dans les deux conditions, sol ferme (désigné par Firm et représenté en blanc) et mousse (désigné par Foam et représenté en noir). La significativité est prise à deux niveaux (* : $p < 0.05$; *** : $p < 0.001$).

2.4.2.3.2 Expérimentation 2

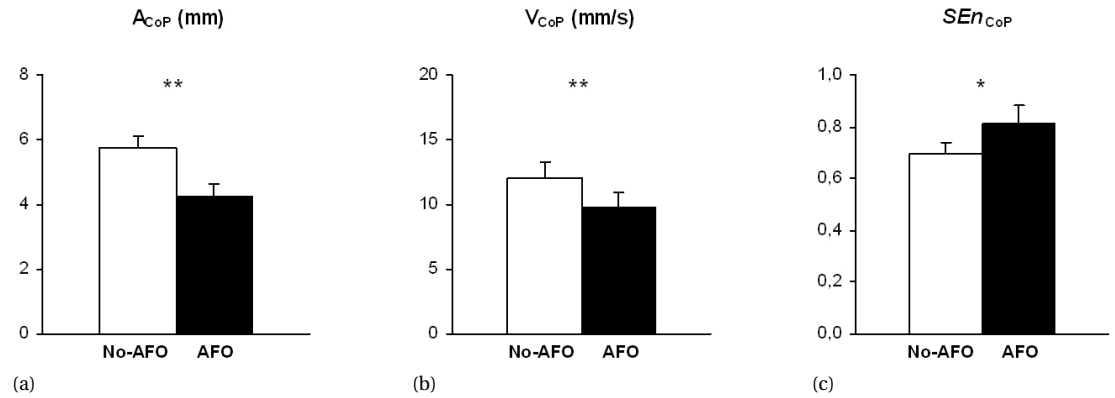


FIGURE 2.2 – Moyenne et variance de la variabilité (A_{CoP}) (panel A), de la vitesse (V_{CoP}) (panel B), et de la régularité (SEn_{CoP}) (panel C) de la trajectoire du centre des pressions dans les deux conditions, sans orthèse (désigné comme No-AFO, blanc) et avec orthèse (désigné comme AFO, noir). La significativité est prise à deux niveaux (* : $p < 0.05$; ** : $p < 0.01$).

Ici encore nous retrouvons le comportement que nous attendions, avec une valeur plus importante de la variabilité et de la vitesse dans la condition sans orthèse par rapport à la condition où le maintien est artificiellement amélioré, avec respectivement $t = 4.70$, $p < 0.01$ pour A_{CoP}, Fig.2.2a et $t = 3.79$, $p < 0.01$ pour V_{CoP}, Fig.2.2b. De la même manière et comme notre hypothèse le laissait à penser, pour la régularité, SEn_{CoP}, nous avons une valeur plus petite lorsque l'orthèse est présente ($t = -3.13$, $p < 0.05$, Fig.2.2c).

2.4.2.4 Conclusions sur cette étude

Le maintien de la posture debout est quelque chose que tout humain a bien expérimenté dès son plus jeune âge. Dans la première année de la vie lorsque le jeune enfant découvre les sensations sous ses pieds et la position de son corps debout dans l'espace, cela devient de plus en plus automatique et c'est ensuite une tâche qui ne demande quasiment aucune réflexion ou action. Cependant, lorsque l'un des sens est diminué ou que l'environnement extérieur en amoindrit, cette tâche peut devenir ensuite plus complexe et plus demandeuse. Ces conditions de contrôle postural dégradées sont notamment observées chez les personnes âgées et sont la première cause de chute pour ces populations (entraînant ainsi des blessures typiques nécessitant des prises en charge). Dans ce contexte, mieux comprendre le contrôle postural et essayer de déterminer des conditions qui pourraient le simplifier est important. L'objectif de cette partie des travaux est d'évaluer la pertinence de la mesure de l'entropie sur ce signal du centre des pressions pour en déduire l'investissement cognitif nécessaire au maintien de la posture.

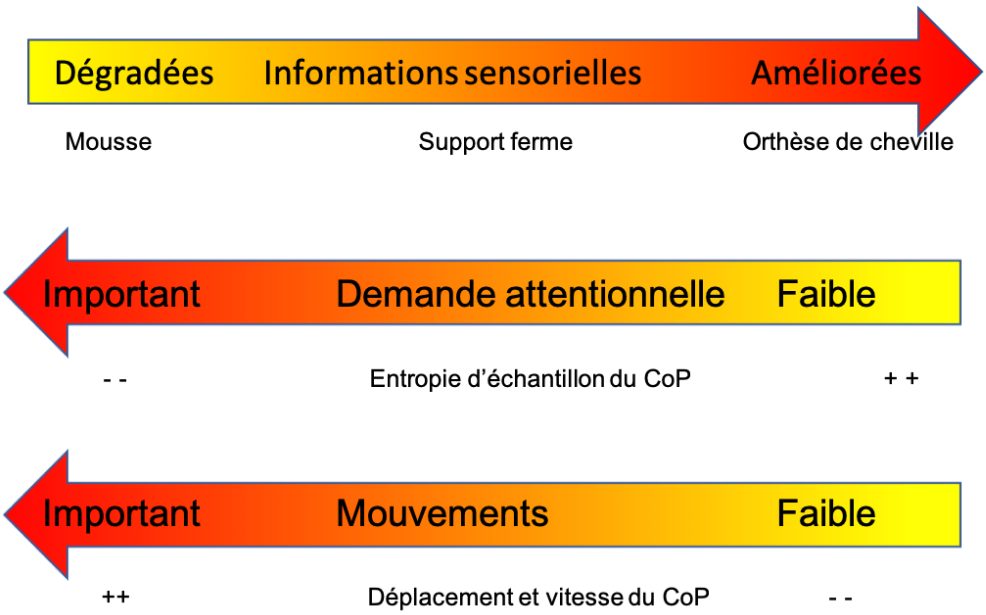


FIGURE 2.3 – Vue schématique des résultats de l'évolution de l'entropie d'échantillon et des trajectoires et vitesse du centre des pressions en fonction de la condition.

Comme précédemment, lorsque les sensations somatiques sont dégradées, les résultats montrent un accroissement de A_{CoP} et V_{CoP} et une décroissance de SEn_{CoP}. Ces deux résultats montrent que les mouvements sont plus amples (mais aussi avec un contenu fréquentiel moindre) dans ce cas du fait d'un moins bon contrôle et qu'un investissement attentionnel plus fort est mis en jeu dans le contrôle de la posture par rapport aux conditions normales. Au contraire, lorsque les sensations somatiques sont améliorées, avec l'utilisation d'une orthèse, A_{CoP} et V_{CoP} décroissent et SEn_{CoP} augmentent, suggérant un meilleur contrôle postural et un plus faible investissement dans ce contrôle (et donc des mouvements « incontrôlés » plus rapides et sporadiques).

En ce qui concerne l'investissement attentionnel pour le contrôle postural, SCHMIT et al. (2005) et RILEY et al. (2005) ont travaillé sur le contrôle bipédal de la posture avec un outil appelé Recurrence Quantification Analysis (RQA) sur le signal du centre des pressions et ont trouvé des résultats opposés aux nôtres (pour ce qui est de l'entropie). Cependant, cette mesure est différente de l'entropie d'échantillon. La variable explicative la plus proche de l'entropie est le pourcentage de récurrence, cependant dans les deux études de Schmidt et Riley le comportement de cette variable est différent pour des mêmes conditions.

L'entropie d'échantillon a été utilisée pour quantifier l'effet du support sur le contrôle de la posture et STINS et al. (2009) et STRANG et al. (2010) ont trouvé le même effet que dans notre cas pour la condition mousse alors que BORG et al. (2010) ont trouvé une différence non significative. Les différences entre tous ces résultats obtenus peuvent être dues aux caractéristiques de la mousse utilisée et principalement son épaisseur et sa densité, que celles-ci soient suffisamment différentes d'un sol ferme pour entraîner un trouble de la perception (PATEL et al., 2008). Dans notre cas, les résultats montrent une baisse de l'entropie comme pour STINS et al. (2009) et STRANG et al. (2010), et dans le cas de l'utilisation d'une orthèse, augmente par rapport au groupe contrôle. Cela signifie que la dynamique du système influant le déplacement du centre des pressions est plus irrégulière dans le cas de l'orthèse et plus régulière dans le cas de la mousse. En d'autres termes, les trajectoires du centre des pressions sont plus riches et moins prédictibles dans le cas de sensations somatiques améliorées alors que des répétitions sont présentes dans le signal dans le cas de conditions dégradées. Pour ce qui est du contrôle, cela signifie que la condition orthèse est moins demandeuse en attention et plus automatique (Fig.2.3). Ces résultats sur l'investissement attentionnel et sur l'automaticité sont en accord avec ceux de ROERDINK et al. (2010) et également avec l'étude de BORG et al. (2010) dans laquelle ils discutent sur l'interprétation de l'entropie en fonction des incohérences de résultats retrouvés dans la littérature.

L'interprétation de l'entropie n'est pas toujours triviale du fait que cela dépend de l'investissement attentionnel, de l'automaticité, mais aussi des bruits de mesure. Cependant, cette difficulté peut être surmontée en combinant

l'interprétation de l'entropie d'échantillon avec les résultats de mesures plus conventionnelles du déplacement du centre des pressions, pour la levée de doute. Dans ce cas, de plus fortes valeurs d'entropie associées à de plus faibles mouvements de balancier du corps seraient interprétées par un réglage plus complexe et précis du contrôle de la posture (Fig. 2.3). Ce comportement est celui qui est attendu dans le cas des personnes jeunes et en bonne santé.

Au final, un changement de mode dans le contrôle de la posture apparaît avec le changement des sensations somatiques et l'entropie d'échantillon est capable de le visualiser et d'en quantifier l'importance. Pris ensemble, les résultats des expérimentations 1 et 2 nous montrent que l'entropie d'échantillon est un outil capable de bien représenter le degré d'automaticité dans le contrôle de l'équilibre. Ces résultats encourageants montrent qu'une mesure régulière de ces paramètres et de leur évolution pourrait être un outil supplémentaire simple et pertinent afin de quantifier l'évolution de l'état d'une personne et de prédire de possibles risques de chutes. Cela pourrait être fait dans le cadre d'un habitat intelligent avec l'intégration de cette mesure dans des outils du quotidien tels que des balances connectées (DUCHÊNE et al., 2011) ou des jeux sérieux (YOUNG et al., 2011).

Afin d'encore mieux comprendre le système de contrôle de la posture, nous avons utilisé une autre méthode de traitement non linéaire, en modifiant l'algorithme habituellement utilisé pour l'entropie basée sur des ondelettes.

2.5 Analyse non linéaire des signaux par décomposition : modification de l'entropie spectrale pour la construction d'un indice local de complexité

Comme vu précédemment, la notion d'entropie peut être intéressante et importante pour évaluer la dynamique d'un système, et donc dans le cas de l'humain, et du fait que le vieillissement est supposé changer le contenu spectral, nous devrions retrouver une richesse moins importante et donc des valeurs d'entropie différentes. Nous le retrouvons déjà dans le cas précédent avec l'entropie d'échantillon, dans lequel une modification des conditions d'expérimentation pour détériorer les sensations de la personne produit un effet notable.

Cependant, l'entropie telle que calculée précédemment est une notion statistique et sa fiabilité sur des données physiologiques sur des temps importants peut être mise en cause, surtout si les conditions changent et que le « système » devient non stationnaire. Le sens que l'on donne au résultat peut alors être faussé.

Céline Franco, dans ses travaux sur le système cardio-respiratoire, a travaillé sur des techniques telles que le synchrosqueezing pour identifier des modes à l'intérieur d'un signal non stationnaire. Ces travaux ont été étendus ensuite au signal postural (et nous nous en tiendrons à ce signal pour la présentation qui est faite ici).

2.5.1 L'entropie spectrale (Wavelet Entropy)

Le but de ces travaux est donc de développer un indice de complexité qui est cohérent dans le cas de signaux oscillatoires multimodaux. L'approche proposée devrait avoir les caractéristiques suivantes : (1) adaptative (on attend le moins de paramétrisation possible), (2) prenant en compte les variations locales et la non-stationnarité (3) lisible, c'est-à-dire que l'on peut voir dans les résultats l'importance des composantes.

Pour cela, une nouvelle approche de l'entropie spectrale (Wavelet Entropy, WE) est proposée. L'entropie spectrale a été développé en premier lieu pour répondre à des soucis de cohérence des résultats sur les signaux de type EEG (ROSSO et al., 2001). Originellement, l'entropie spectrale est basée sur la transformée en ondelettes discrètes orthogonales. Du fait des composantes orthogonales de la base de décomposition, le calcul d'énergie ainsi que le lien aux coefficients temps-fréquence sont relativement aisés. Le revers majeur de cette méthode est qu'elle implique un filtrage d'une sous-bande dont les limites et la largeur sont fixées par la fréquence d'échantillonnage. Ainsi, la résolution est fixée et deux modes qui se trouvent dans la même sous-bande sont indissociables. Dans notre approche, nous remplaçons la transformée en ondelettes orthogonales par une transformée par synchrosqueezing (DAUBECHIES, J. LU et al., 2011). Cette seconde transformée offre une représentation des modes qui est cohérente avec le contenu spectral pour les signaux physiologiques multi-composantes, et ce, en réassignant les coefficients de la transformée en ondelette continue en fréquence.

2.5.2 La transformée par synchrosqueezing

Le choix d'une représentation temps-fréquence dépend de l'analyse que nous souhaitons réaliser et sera un compromis. Ce compromis résulte d'un principe bien connu qui est le principe d'Heisenberg-Gabor, selon lequel une bonne localisation en temps ne peut être atteinte qu'au détriment d'une moins bonne localisation en fréquence et vice-versa. Ce principe (connu des physiciens pour la position/vitesse des particules) énonce en effet que le produit de la dispersion d'énergie en temps par la dispersion d'énergie en fréquence a une borne minimum.
La réallocation part de ce constat en combinant les informations de deux distributions complémentaires afin de corriger la moins bonne localisation. Originellement, cela consiste à déplacer les valeurs d'une distribution présentant des défauts de localisation (par exemple la transformée de Fourier à court terme) à partir d'une distribution bien localisée, mais au détriment d'autres aspects (exemple, la transformée de Wigner-Ville).
Dans le cas de la transformée par synchrosqueezing, les coefficients du scalogramme sont déplacés en fonction du comportement fréquentiel local du signal. Contrairement aux autres méthodes de réallocation, celle-ci ne porte que sur l'une des directions, la direction fréquentielle. Cette spécificité permet de reconstruire le signal de manière exacte.
Le calcul de la transformée par synchrosqueezing se fait de la manière suivante :

— Calcul de la transformée en ondelettes continue (notée $W_s(a,b)$)

— Estimation de la fréquence FM démodulée :

$$w_s(a,b) = \begin{cases} \frac{1}{2i\pi W_s(a,b)} \frac{\delta W_s(a,b)}{\delta b}, \text{ Si } |W_s(a,b)| > 0 \\ \infty, \text{ sinon} \end{cases}$$

— La partie de compression ou encore partie squeezing où l'on réalloue les coefficients : $(b,a) \mapsto (b, w_s(a,b))$. Finalement, nous obtenons pour la transformée : $T_s(w_l,b) = \int_{\{a:|w_s(a,b)-w_l|<\Delta w\}} W_s(a,b) \cdot a^{-\frac{3}{2}} \cdot da$

2.5.3 Calcul de l'entropie spectrale modifiée WE_{SQT}

Si l'on considère un signal s, sa transformée par synchrosqueezing $C_j(k)$, le support du cône d'influence $\text{Supp}_{\text{Cone(k,j)}}$ à l'échantillon temporel k et à l'échelle j (avec $1 \leq j \leq a_{max}$ avec $a_{max} = f_0/f_{min}$, tel que f_0 est la fréquence centrale de l'ondelette mère et f_{min} la borne inférieure de l'ensemble des fréquences, a_{max} est alors la dilatation maximale de l'ondelette dans la transformée), l'entropie spectrale modifiée se calcule ainsi :

$$E(j)_i = \sum_{\text{Supp}_{\text{Cone(k,j)}}} \frac{|b-b_0| \cdot |C_j(k)|^2}{a}, \text{ énergie à l'échelle j,}$$

$$E_{tot}(i) = \sum_{j=1}^{a_{max}} E_j,$$

$$p_j(i) = \frac{E_j(i)}{E_{tot}(i)}, \text{ énergie relative à l'échelle j,}$$

$$WE_{SQT}(i) = -\sum_{j=1}^{a_{max}} p_j(i) \cdot ln(p(j)),$$

pour un instant i donné.
La figure 2.4 illustre l'évolution de cet indice pour un signal donné (et la notion de cône d'influence).

2.5.4 Tests sur des signaux posturaux et protocole

De premiers tests sur des signaux simulés ont d'abord été effectués pour vérifier le comportement correct de l'entropie spectrale modifiée et sa bonne variation comme attendu. Nous avons ensuite testé cela sur des données réelles de centre de pressions, dans une expérimentation particulière.
Comme déjà évoqué, le maintien d'un bon équilibre debout demande l'intégration d'informations visuelles, vestibulaires et somato-sensorielles. Pour évaluer notre indice sur un cas précis, nous avons choisi de nous concentrer sur

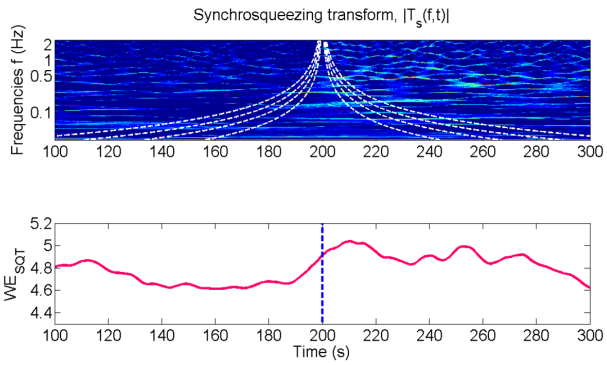


FIGURE 2.4 – Transformée pour synchrosqueezing et WE_{SQT} correspondante

le signal du centre des pressions comme précédemment, et de mettre un changement de conditions sensorielles au milieu de l'expérimentation. Ainsi, notre indice ne requérant pas que le signal soit stationnaire pourra montrer ses avantages par rapport aux autres. Nous voulons également évaluer ici la capacité d'une personne à s'adapter à de tels changements de conditions. La modification du contexte sensoriel va engendrer une repondération des informations reçues. Cette repondération se fait en fonction de la disponibilité de l'information et de sa fiabilité ressentie. En termes de signal, il est évident que nous nous attendons à une non-stationnarité de celui-ci et donc à un avantage de notre indice sur d'autres.

Ainsi, nous avons concentré notre protocole sur une tâche simple de maintien de l'équilibre en incluant un changement abrupt de conditions sensorielles. Ainsi, la personne était debout sur la plateforme avec une information supplémentaire, car elle devait fixer l'image d'une croix représentée sur le mur en face. Cette croix lui donnait donc un point de repère supplémentaire. Au bout de 200 secondes d'expérimentation, il lui était demandé de fermer les yeux et de continuer de maintenir cette position 200 secondes supplémentaires. Ceci a été testé sur 8 personnes jeunes et en bonne santé (âge : $34,63 \pm 11$ ans, IMC : $21,41 \pm 1,57$). Le déplacement du centre des pressions était enregistré par une plateforme de force de marque Satel à une fréquence de $40Hz$. Pour éviter les effets de bords et pour se focaliser sur les données qui ont vraiment un intérêt pour nous, nous concentrons l'analyse sur les 200 secondes autour de la perturbation sensorielle. Du fait de la taille de notre échantillons (peu de personnes), une analyse non paramétrique est utilisé pour la comparaison des moyennes (test signé de rang de Wilcoxon). En plus de ceci, les valeurs obtenues pour la WE_{SQT} sont calculées sur les données yeux ouverts et fermés indépendamment et comparés avec un T-test sur les échantillons deux à deux. Chaque personne a réalisé six essais.

2.5.5 Résultats et discussion

La stabilité posturale est évaluée par des mesures classiques : le RMS (Root Mean Square) de l'amplitude (des mouvements du centre des pressions mesurés par la plateforme de force), en mm, celui de la vitesse et enfin l'aire de l'ellipse de confiance à 95%. Le tableau 2.1 résume les résultats de cette expérimentation.

Mesure	Yeux Ouverts	Yeux Fermés
RMS amplitude	$3,94 \pm 2,09$	$5,83 \pm 3,07$*
RMS vitesse	$9,29 \pm 2,87$	$17,17 \pm 9,67$*
Surface ellipse à 95%	$24,17 \pm 8,28$	$40,19 \pm 24,13$*
Sample Entropy	$0,76 \pm 0,23$	$0,83 \pm 0,42$
WE_SQT	$4,17 \pm 0,48$	$4,39 \pm 0,38$*

TABLE 2.1 – Effet de la condition visuelle sur le déplacement du centre des pressions. Significativité * : $p < 0,05$

Ils indiquent que toutes les valeurs spatio-temporelles augmentent avec le changement de condition, indiquant une moins bonne stabilité avec la privation visuelle ($p < 0,05$, $0,01$, $0,01$ respectivement). Le contrôle postural est estimé par les deux entropies calculées (pour comparaison). Aucune différence significative n'est observée pour l'entropie d'échantillon entre les deux conditions alors que la WE_{SQT} augmente, indiquant un accroissement dans la

complexité et dans l'utilisation des ressources. Une comparaison des valeurs pour chacun des individus montre ce comportement sur 7 personnes sur les 8, confirmant ainsi notre hypothèse de changement de la complexité avec la privation d'information visuelle.
Les travaux sur le contrôle de la posture sont basés sur l'évaluation de la complexité. Cependant, la plupart des approches travaillent sur l'entropie d'un point de vue temporel et évaluent la prédictibilité d'une série temporelle sans prendre en compte les multiples échelles qui entrent en compte ainsi que les processus transitoires. L'entropie multi-échelle tente d'intégrer de telles considérations avant d'appliquer l'entropie d'échantillon, mais les représentations multi-échelles sous-jacentes et la difficulté à paramétrer celle-ci donnent des indices avec des comportements variables et des résultats difficiles à interpréter. L'approche proposée ici considère la complexité d'un point de vue spectral, en cohérence avec le modèle spectral du signal de déplacement du centre des pressions dont le comportement dépend des informations sensorielles.
De plus, la décomposition par des méthodes de synchrosqueezing améliore l'identification des différentes composantes du signal conduisant à une meilleure interprétation de la représentation temps-fréquence.
Nous avons vérifié dans ces travaux que, lorsque nous modifions les informations sensorielles de la personne, le comportement de l'indice basé sur la WE_{SQT} était modifié en cohérence avec ce changement brut. Nous avons vérifié aussi qu'un autre indice basé sur l'entropie ne voyait pas de modification significative étant donné que le signal devenait non-stationnaire et que ce premier calcul moyennait les deux comportements.
Ces travaux et les précédents ont porté sur l'utilisation d'une plateforme de force pour obtenir le signal du centre des pressions. Pour éviter d'utiliser un matériel supplémentaire, nous avons tenté d'utiliser un objet du quotidien connu, le smartphone, sur lequel nous avons concentré nos recherches. Cela fait l'objet de la partie suivante.

2.6 Les signaux du smartphone : vers l'application iProprio

La précédente partie, pour les données posturales, se base sur l'utilisation des données de plateformes de force. Ces plateformes ont cependant comme inconvénient de devoir être posées de manière précise sur un sol stable, et ne peuvent être déplacées lors de l'expérimentation. Elles doivent également être connectées à un PC.
Nous devons donc nous tourner vers des IMU (Inertial Motion Unit, Centrale de Mesure Inertielle). Or, depuis 2008 et l'apparition du premier iPhone, les smartphones embarquent ce genre de centrale, et permettent également une forte connectivité, simple à mettre en œuvre. Nous nous sommes donc tournés vers ces smartphones dans le but d'intégrer nos travaux dans un outil du quotidien et ainsi ne pas dépendre de capteurs extérieurs dont la diffusion et la coût pourraient être un frein au déploiement de telles solutions.
Les travaux présentés par la suite ont contribué à l'état de l'art de manière significatives par les aspects suivants, qui n'avaient pas encore été traités :

— Avant de débuter ces travaux, nous avons évalué la précision des centrales inertielles des smartphones en fonction des différents filtres ou des différents algorithmes de fusion d'information permettant la reconstruction des angles. Ce travail n'avait alors pas encore été effectué.

— Nous avons également intégré toute la chaine de création et de réalisation d'exercices de proprioception au smartphone afin de réaliser des systèmes de mesure et de rééducation à bas coût.

— Enfin nous avons déterminé et évalué l'influence de plusieurs paramètres sur les retours sensoriels et l'efficacité de ce retour sensoriel dans ce système tout intégré qu'est le smartphone.

Les travaux qui seront présentés dans cette section ont donné lieu à de nombreuses publications et deux brevets. Le premier brevet concerne le système iBalance/iProprio (rééducation de la posture et du mouvement) (Fleury, Vuillerme, Franco et al., 2016) et le second brevet intègre l'analyse de la demande attentionnelle (Fleury, Vuillerme, Diot et al., 2017). iBalance a donné lieu à une publication journal (Franco, Fleury, Guméry et al., 2013), ce qui est également le cas d'iProprio (Mourcou, Fleury, Diot, Franco et al., 2015) et de l'analyse des performances angulaires du smartphone (Mourcou, Fleury, Franco et al., 2015). Des résultats expérimentaux ont également été publiés en conférences (Mourcou, Fleury, Diot et Vuillerme, 2016; Fleury, Mourcou, Franco, Diot et al., 2013; Fleury, Mourcou, Franco et Vuillerme, 2018).

2.6.1 Test de la validité de ces mesures

2.6.1.1 Matériel, algorithmes et méthode

2.6.1.1.1 Matériel

Le marché des smartphones, quand il s'agit de s'attaquer à la précision d'une mesure avec celui-ci, est complexe du fait de la présence de plusieurs acteurs tant au niveau matériel que logiciel. En effet, plusieurs vendeurs se partagent le marché (Apple, Samsung, Sony, etc.) et plusieurs systèmes d'exploitation sont présents sur les téléphones (iOS et ses différentes versions, Android et ses différentes évolutions). Afin de couvrir une large partie de ce qui se trouve sur le marché, nous avons travaillé avec deux vendeurs (Apple et Samsung) qui représentent sur les deux systèmes d'exploitation dominants (iOS et Android), avec des capteurs différents.

Le premier smartphone qui a été testé est l'iPhone 4 qui contient : (1) un accéléromètre 3D (ST-Microelectronics, LIS331DLH), (2) un gyroscope 3D (ST-Microelectronics, L3G200D) et (3) un magnétomètre 3D (Asahi Kasei Microdevices, AKM8975). Pour ces tests, le téléphone était équipé d'iOS 7 et le logiciel de recueil de données avait été créé pour les besoins de l'expérimentation. Ce logiciel est une simple interface graphique permettant de collecter les données en utilisant différents algorithmes (programmés ou donnés par le système d'exploitation), en utilisant une fréquence donnée et choisie par la personne et des conditions de mesures choisies également. Toutes ces données sont ensuite stockées dans un fichier CSV. Le second smartphone de test est l'iPhone 5S, avec le même système d'exploitation mais avec une nouvelle génération de capteurs. Il contient donc : (1) un accéléromètre 3D (Bosch Sensortec BMA220), (2) un gyroscope 3D (ST-Microelectronics, L3G4200DH) et (3) un magnétomètre 3D (AKM, AK8963). Nous l'avons équipé de la même application que pour l'iPhone 4. Enfin, le troisième téléphone testé est le Samsung Galaxy Nexus, qui contient : (1) un accéléromètre 3D (Bosch Sensortec BMA220), (2) un gyroscope 3D (InvenSens, MPU-3050) et (3) un magnétomètre 3D (Yamaha, YAS530). Le téléphone est équipé d'Android 4.3 Jelly Bean et d'une application développée pour l'occasion similaire, mais programmée pour la plateforme Android.

À des fins de comparaison, nous avons également pris une centrale inertielle très utilisée en capture de mouvement, de la marque XSens, et servant de référence dans de nombreux travaux scientifiques (CUTTI et al., 2008; HAMACHER et al., 2012). Le tracker de mouvement MTx est un ensemble de capteurs de type MEMS (accéléromètre, gyroscope et magnétomètre 3D) dont les données sont retraitées par le logiciel de la société. La précision statique annoncée est de 0,5° pour le roulis et le tangage et de 1° pour le lacet. Les données sont collectées à l'aide du logiciel proposé par la solution.

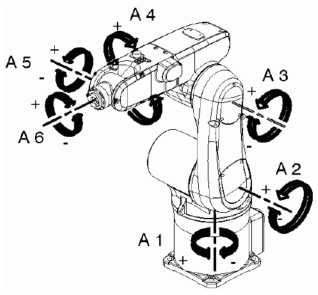


FIGURE 2.5 – Bras robotisé Kuka et ses axes (figure fournie par le manuel de spécifications "KR 5 sixx R650, R850").

Enfin, nous avons pris comme référence le bras robotisé Kuka KR5-SIXX-R650 (Figure 2.5). Ce bras robotisé est un bras 6 axes sur lequel nous avons utilisé l'axe A5 pour le tangage, l'axe A4 pour le roulis et l'axe A6 pour le lacet. L'axe A5 a une plage de mouvements de ±120°, l'axe A4 de ±190° et l'axe A6 de ±358°. La vitesse de mouvement avec une charge de 5 Kgs est de maximum ±410°/s pour A4 et A5 et de ±660°/s pour A6. La précision de répétabilité est de $0,02mm$ d'après le constructeur [4].

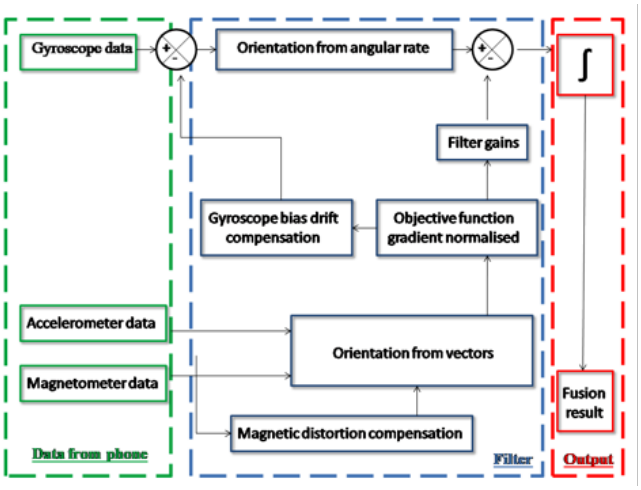


FIGURE 2.6 – Estimation des angles

2.6.1.1.2 Algorithmes de calcul des angles

Les signaux bruts des accéléromètres, magnétomètres et gyroscopes nous donnent respectivement l'accélération (linéaire) de la personne, le champ magnétique (terrestre et le bruit ambiant), et la vitesse angulaire. Cependant, pour calculer l'orientation courante de la personne et l'ensemble de son mouvement, aucun de ces capteurs n'apporte une information parfaite. Cependant, avec la fusion de ces trois capteurs, nous pouvons créer un système AHRS (Attitude and Heading Reference System) dans lequel chaque capteur va compenser le biais induit par les autres. Ainsi, les projections des mesures gravitationnelles et magnétiques vont être fusionnées pour calculer une estimation des angles d'orientation des capteurs (et donc du téléphone). Cette solution est obtenue par l'utilisation de filtres d'orientation parmi lesquels nous trouvons le filtre complémentaire (EUSTON et al., 2008), le filtre de Kalman (WELSH et al., 1995), le filtre de Mahony (MAHONY et al., 2008) et le filtre de Madgwick (MADGWICK, 2010). L'estimation de l'orientation revient à évaluer les équations de la cinématique pour la rotation du téléphone. La tâche du filtre est de calculer, à partir des mesures et pour une orientation donnée, une meilleure estimation de l'orientation. Ainsi, le filtre estime la position en fonction de la mesure de la vitesse angulaire sur chacun des axes et de l'ancienne position. Ensuite, l'erreur dans cette mesure est soustraite en utilisant les signaux de l'accéléromètre et du magnétomètre pour avoir une meilleure estimation (MADGWICK, 2010). Afin que l'estimation soit la meilleure possible, le bruit magnétique ainsi que le biais d'intégration du gyroscope doivent être compensés. Trois des quatre filtres pré-cités (à savoir Kalman, Mahony et Madgwick) ont des manières différentes de résoudre l'estimation de ces biais. Mahony utilise un régulateur proportionnel-intégral pour le biais d'intégration du gyroscope alors que Madgwick se contente d'un proportionnel uniquement. Un diagramme de bloc est présenté à la figure 2.6.

Dans tous ces filtres, la représentation utilisée est les quaternions. Dans cet hyperespace complexe à quatre dimensions, il est plus simple de représenter les orientations (et surtout leurs évolutions) et les calculs sont plus efficaces. Cependant, ces angles sont moins parlants pour une personne que les angles d'Euler utilisés en cinématique ou bien dans le domaine clinique. Les angles d'Euler sont sujets à deux problèmes : des ambiguïtés dans la représentation et un phénomène appelé « Gimbal Lock » (en français blocage de cardan, phénomène présent dans les rotations en 3D lorsque deux des trois plans nécessaires pour appliquer des rotations se retrouvent portés par la même direction). Ces deux problèmes sont pris en compte et résolus dans les protocoles proposés.

Dans les mesures effectuées, nous avons comparé le comportement, sur les données smartphone, des filtres de Mahony et de Madgwick ainsi que des filtres proposés par les systèmes d'exploitation.

2.6.1.1.3 Méthode de validation

Pour évaluer la précision sur smartphone de la mesure angulaire, nous avons reproduit, à l'aide du bras robotisé, des mouvements sur les deux axes de roulis et tangage. Pour cela, nous avons fixé les smartphones sur le bras robotisé 2.7. Ce bras a une précision de l'ordre de 10^{-2}° en position, bien inférieure à ce que l'on pourrait obtenir avec des

4. http://www.kuka-robotics.com/res/sps/e6c77545-9030-49b1-93f5-4d17c92173aa_Spez_KR_5_sixx_CR_en.pdf

FIGURE 2.7 – Positionnement des capteurs sur le robot

capteurs inertiels. Généralement, ce sont des systèmes optiques qui servent de références (MADGWICK, 2010 ; CUTTI et al., 2008). Cependant, le bras robotisé Kuka est bien plus précis que des systèmes optiques type Optotrak ou Vicon 612, les deux systèmes utilisés dans les publications citées en références. Nous avons réalisé deux types de protocoles afin d'évaluer la précision dans différents cas. Un protocole se concentrant sur la répétabilité des mesures et sur l'importance de la manière de fixer le téléphone et un second protocole tentant d'évaluer la précision de la mesure angulaire par rapport à la mesure de référence :

— Le protocole évaluant la répétabilité des mesures : Dans ce protocole, le bras robotisé effectue un mouvement allant de 0 à 180° par pas de cinq degrés avec une pause de dix secondes à chaque position. Ce protocole est répété six fois pour des mouvements sur les deux angles de roulis et tangage. Les mesures d'angle sur les téléphones ne sont faites qu'avec l'angle donné par le système d'exploitation. Tout ceci est exécuté dans quatre conditions différentes : (1) le smartphone est centré sur le robot, (2) deux smartphones, mis dos à dos et attachés à l'aide d'un élastique, sont centrés sur le robot, (3) le smartphone est positionné sur le robot à l'endroit où ses capteurs se trouvent physiquement et (4) le smartphone est centré sur l'opposé de la position précédente.

— Performances par rapport au gold standard : le but est de comparer la précision des smartphones et du XSens en même temps à l'aide de tous les filtres AHRS précédemment cités. Ce protocole se subdivise en deux sous catégories. Les mesures statiques qui reprennent le protocole précédent et les mesures dynamiques qui consistent à faire un mouvement de 0 à 180° par pas de 45 degrés avec des vitesses de 20% et 50% de la vitesse maximum du robot. Entre chaque position, une pause de dix secondes est effectuée et la mesure est répétée six fois.

Pour évaluer la précision de la mesure, pour chaque filtre et chaque objet, le RMS (Root Mean Square) de la différence entre l'angle estimé et l'angle du robot est calculé. Dans notre mesure, le gold standard est l'ensemble théorique alors que le smartphone ou le XSens sont les ensembles variables. Ainsi, le RMS de la différence entre ces deux ensembles permet de mesurer à quel point en moyenne l'erreur est proche de zéro.

Lors du protocole 1, un test de Kruskal-Wallis a été effectué pour évaluer l'effet de la position en déterminant si les valeurs RMS dans les quatre conditions sont similaires. Ce test non paramétrique est très utilisé pour tester si k échantillons indépendants (ici nous cherchons à savoir si les quatre positions de prises de mesures donnent des valeurs similaires) sont extraites d'une même distribution statistique.

Lors du protocole 2 pour chaque objet et chaque filtre, nous avons calculé les angles et le RMS de la différence avec le gold standard comme précédemment. Un test de rang signé de Wilcoxon a ensuite été utilisé pour évaluer l'impact de la vitesse sur la précision de la mesure. Un test de Kruskal-Wallis a été utilisé lui pour comparer les filtres entre eux pour un même objet. Le meilleur des filtres a ensuite été sélectionné pour évaluer la précision des capteurs XSens à l'aide toujours d'un test de Kruskal-Wallis.

2.6.1.1.4 Résultats

En ce qui concerne le tout premier test qui s'attelait à tester la reproductibilité des mesures et le positionnement, le RMS a été calculé pour 27 angles (évitant les positions proches de 90° pour le Gimbal Lock). Pour toutes les conditions, aucune différence significative n'a été observée sur l'axe de roulis. Pour l'axe de tangage, une différence significative ($p < 0.01$) n'a été trouvée que pour le téléphone ventousé sur le robot à l'endroit où les capteurs sont physiquement dans celui-ci.

En ce qui concerne la reproductibilité des mesures (intra-conditions donc), toutes les mesures appartiennent aux mêmes distributions lorsqu'il s'agit d'un même objet avec le même filtre.

Nous pouvons donc conclure de tout ceci que les mesures avec le smartphone sont reproductibles. Le RMS moyen dans cette condition est inférieur à $0,12°$ ce qui est très faible. Cependant, pour éviter ce biais, la suite des expérimentations considérera le smartphone en son centre. Nous pouvons cependant considérer que, dans le cadre de mesures cliniques, notre test est reproductible et indépendant de la position du smartphone.

		Roll		Pitch	
		RMS	**Variance**	**RMS**	**Variance**
Nexus	Manufacturer filter	0.16 (0.05–0.42)	0.36 (0.33–0.39)	0.21 (0.07–0.35)	0.42 (0.38–0.47)
	Madgwick	0.19 (0.07–0.42)	0 (0–0)	0.21 (0.06–0.62)	0.01 (0.01–0.01)
	Mahony	0.16 (0.05–0.38)	0 (0–0)	0.25 (0.05–0.50)	0.01 (0.01–0.01)
iPhone 5S	Manufacturer filter	0.15 (0.02–0.48)	0 (0–0)	0.13 (0.02–0.24)	0 (0–0.01)
	Madgwick	0.14 (0.03–0.47)	0.01 (0–0.01)	0.29 (0.05–0.55)	0 (0–0.02)
	Mahony	0.13 (0–0.50)	0.02 (0–0.09)	0.17 (0.03–0.29)	0 (0.01–0.01)
iPhone 4	Manufacturer filter	0.07 (0.01–0.18)	0.02 (0–0.17)	**0.08 (0.01–0.16)**	0.36 (0–1.12)
	Madgwick	0.10 (0.01–0.17)	0.55 (0–1.57)	**0.13 (0.02–0.63)**	0.08 (0–0.39)
	Mahony	0.12 (0.01–0.61)	0.09 (0–0.42)	0.09 (0.01–0.16)	0.56 (0–1.58)
Xsens	Manufacturer filter	0.22 (0.08–0.36)	0 (0–0.01)	0.22 (0.11–0.28)	0 (0–0.01)
	Madgwick	0.57 (0.02–3.44)	0.05 (0–0.62)	0.16 (0.05–0.29)	0 (0–0.03)
	Mahony	0.68 (0.02–5.45)	0.08 (0–1.26)	0.10 (0.03–0.18)	0 (0–0.02)

FIGURE 2.8 – Résultats des mesures en statique sur les angles de roulis et tangage

2.6.1.1.4.1 Protocole statique La figure 2.8 reprend les résultats obtenus pour les deux angles de roulis et tangage pour le protocole statique. D'une manière générale, les trois smartphones donnent des résultats excellents. Pour l'angle de roulis, le RMS est en-dessous de $0,2°$ pour tous les filtres et tous les smartphones et il n'y a aucune différence significative entre ceux-ci. L'iPhone 4 est meilleur que les deux autres (il a été sélectionné pour ce test, car il inclut l'ancienne génération de capteurs inertiels). Le capteur XSens est, quant à lui et en utilisant le filtre du revendeur, au-dessus des $0,2°$. Pour ces résultats, sur l'angle de roulis, XSens est même au-dessus de $0,5°$ avec les deux filtres que nous proposons. Nous ne pouvons expliquer ce fait sauf à ce que le filtre du revendeur corrige en interne un biais dans la mesure sur cet axe de roulis que nous ne corrigerions pas (surtout que ceci n'est pas présent sur l'autre angle). Pour l'angle de tangage, les résultats sont les mêmes avec des RMS inférieurs tous à $0,3°$ et toujours une supériorité de l'iPhone 4 en termes de résultats.

En ce qui concerne l'influence du filtre, le test de Kruskal-Wallis a été utilisé pour comparer, pour un même objet, les différences entre les filtres. La seule différence significative que nous trouvons est dans le cas de l'iPhone 4 entre le filtre de Madgwick et le filtre du système d'exploitation. Aucune autre différence significative n'est à noter pour chaque objet avec chaque filtre, comme documenté ci-après.

Une comparaison entre les RMS pour le tangage et le roulis entre les smartphones et XSens n'a pas donné non plus de différence significative.

Enfin, lorsque nous regardons les signaux en eux-mêmes, nous avons aussi noté que le filtre proposé par Android donnait une variance plus élevée du signal que les autres filtres, mais la plus faible valeur RMS moyenne.

2.6.1.1.4.2 Protocole dynamique Les tableaux 2.9 et 2.10 donnent les résultats pour les 5 objets aux deux vitesses testées sur les angles respectifs de roulis et de tangage. Les valeurs moyennes ainsi que leurs variances sont présentées. Les valeurs RMS sont supérieures à celles obtenues dans le protocole statique, surtout pour les filtres de Madgwick et

		20%		50%	
		RMS	Variance	RMS	Variance
Nexus	Manufacturer filter	1.55 (0.64–3.51)	0.34 (0.31–0.38)	1.57 (0.81–3.43)	2.15 (0.59–4.37)
	Madgwick	3.36 (0.39–7.95)	0.02 (0–0.06)	2.84 (0.58–6.23)	0.90 (0.12–2.65)
	Mahony	3.56 (0.69–8.24)	0.03 (0–0.08)	3.44 (0.67–7.73)	0.99 (0.13–3.21)
iPhone 5S	Manufacturer filter	0.75 (0.32–1.33)	0 (0–0)	0.78 (0.32–1.29)	0 (0–0.02)
	Madgwick	8.05 (3.09–15.96)	0.73 (0.02–2.11)	8.70 (2.77–17.27)	0.99 (0.04–2.75)
	Mahony	2.42 (0.29–4.54)	3.60 (0–14.11)	4.16 (0.71–7.57)	6.09 (0.02–23.94)
iPhone 4	Manufacturer filter	3.57 (0.55–11.03)	0.55 (0.02–1.27)	3.52 (0.66–10.80)	0.48 (0.02–1.06)
	Madgwick	6.99 (0.91–17.92)	1.89 (0–7.56)	8.16 (0.95–20.35)	1.05 (0.01–3.66)
	Mahony	7.02 (0.95–17.97)	2.94 (0–11.74)	8.76 (1.02–21.35)	3.99 (0.01–15.90)
Xsens	Manufacturer filter	2.21 (0.84–4.29)	0 (0–0)	2.55 (0.84–4.99)	0.01 (0–0.03)
	Madgwick	10.39 (1.05–21.04)	1.28 (0–4.78)	10.51 (0.62–21.99)	2.24 (0–6.48)
	Mahony	3.93 (1.40–8.00)	7.49 (0–29.96)	6.94 (1.33–14.20)	26.79 (0–107.16)

FIGURE 2.9 – Résultats des mesures en dynamique sur l'angle de roulis (à 20 et 50% de la vitesse maximale)

		20%		50%	
		RMS	Variance	RMS	Variance
Nexus	Manufacturer filter	2.29 (1.25–3.77)	3.74 (1.01–6.93)	2.00 (0.71–3.48)	1.42 (0.64–3.10)
	Madgwick	2.88 (1.61–4.78)	2.66 (0.79–5.41)	2.65 (0.77–4.15)	1.68 (0.03–3.53)
	Mahony	5.64 (0.66–15.25)	3.06 (2.03–5.19)	2.97 (1.07–5.17)	2.02 (0.09–3.02)
iPhone 5S	Manufacturer filter	1.94 (0.85–7.41)	0 (0–0)	1.94 (0.91–4.01)	0 (0–0)
	Madgwick	3.36 (0.34–7.41)	0.33 (0–0.91)	3.45 (0.25–7.26)	0.27 (0.04–0.76)
	Mahony	1.17 (0.14–3.18)	0.03 (0–0.06)	1.15 (0.48–2.69)	0.29 (0.01–0.71)
iPhone 4	Manufacturer filter	1.40 (0.37–2.20)	1.73 (0.07–2.33)	1.33 (0.32–2.18)	1.78 (0.07–2.45)
	Madgwick	1.97 (1.30–2.32)	0.02 (0–0.03)	1.93 (1.40–2.28)	0.11 (0.06–0.31)
	Mahony	1.12 (0.21–2.07)	0 (0–0.01)	1.05 (0.32–1.90)	0.07 (0–0.21)
Xsens	Manufacturer filter	0.87 (0.30–1.21)	0.02 (0–0.05)	0.94 (0.30–1.23)	0.04 (0–0.14)
	Madgwick	2.55 (1.02–4.10)	0.31 (0–1.20)	2.42 (0.65–4.19)	0.29 (0.01–0.95)
	Mahony	1.24 (0.19–2.94)	0.06 (0–0.11)	1.43 (0.25–3.22)	0.72 (0–2.24)

FIGURE 2.10 – Résultats des mesures en dynamique sur l'angle de tangage (à 20 et 50% de la vitesse maximale)

Mahony. Ceci peut être dû à l'effet de Gimbal Lock qui n'est pas prévu dans ce protocole et qui peut affecter particulièrement l'axe de roulis dans la zone des 90° en fonction de l'angle de lacet qui sera pris. Une solution algorithmique peut être trouvée, mais le plus efficace reste l'utilisation des quaternions (sans passage dans le domaine des angles d'Euler). Il faut noter tout de même que les vitesses testées sont plus importantes que ce qui pourrait être fait par un humain d'une part, mais aussi que les valeurs, dans la plupart des cas, restent relativement bonnes par rapport aux exigences en mesure clinique.
À noter que les résultats présentés pour XSens sont cohérents avec les valeurs annoncées par le constructeur.

2.6.1.1.4.3 Application à la mesure clinique Ces résultats de RMS et de variance doivent être discutés en fonction du contexte applicatif de cette recherche. Dans un contexte clinique, les résultats doivent être en dessous de l'erreur acceptable de $2,7°$ qui est l'erreur pour une mesure d'amplitude de mouvement passive observée récemment pour un goniomètre universel, qui est l'outil de référence (MILANESE et al., 2014). La mesure dite passive d'amplitude de mouvements est une mesure faite sans effort du patient et pour lequel la mobilisation de la partie à mesurer est faite par un praticien. Ceci peut être comparé avec notre protocole statique du fait des conditions dans lesquelles le praticien effectuera la mobilisation du membre. De plus, dans la même étude, une application smartphone qui imite un goniomètre à partir des capteurs a aussi été testé et donnait une erreur de mesure de $1,4°$.
Pour une acceptation et une utilisation de ces objets dans une mesure clinique, l'erreur de mesure entre testeurs devrait être inférieure à 5° (GAJDOSIK et al., 1987) dans la mesure des membres supérieurs et de $5,5°$ (BROSSEAU et al., 1997) dans le cas de mesure active du membre inférieur. La mesure active du membre inférieur est la mesure dans laquelle le patient va activement (sans assistance) mobiliser une articulation à l'aide des muscules adjacents pour atteindre une certaine position. Elle peut se rapprocher dans ce cas de notre protocole dynamique. Dans cette étude, la mesure statique donne des résultats acceptables, quels que soient les filtres dans le cas de la mesure statique. En

mesure dynamique, les filtres donnés par les systèmes d'exploitation sont en dessous des erreurs acceptables alors que les filtres proposés n'ont pas toujours un comportement qui respecte les normes (dépendance des conditions de mesures et de l'axe).
Nous pouvons donc en conclure que le smartphone est un outil suffisant pour de telles mesures en clinique, mais l'effet du filtre en fonction de la vitesse ou selon le type de mesure doit être discuté car va être dépendant de l'application.
Comme nous pouvons le voir dans le tableau, pour le protocole statique, si nous rangeons les filtres par précision croissante, les résultats ne sont pas homogènes en fonction des objets et des axes, contrairement au protocole dynamique. Cependant nous avons également vu que les différences ne sont pas statistiquement significatives. L'effet du filtre dans ce contexte est donc inexistant. L'impact de la vitesse a lui été évalué par un test de rang signé de Wilcoxon. Au contraire de LEBEL et al. (2015), aucune différence significative entre les vitesses n'est trouvée pour un filtre et un objet donné. Cependant, notre protocole n'est pas tout à fait le même. La mesure du RMS est calculée durant une période statique plus longue après un mouvement plus ample que celui de notre protocole statique. Enfin, une comparaison a été effectuée entre le filtre de XSens et les filtres des systèmes d'exploitation des autres objets avec un test de Kruskal-Wallis. XSens est actuellement très utilisé dans un contexte de mesure d'amplitude de mouvement en clinique. Aucune différence significative n'a pu être trouvée lors de ces tests concluant que les résultats de mesure d'amplitude de mouvements par smartphones sont au moins aussi bons que ceux de XSens. Cela confirme les précédentes conclusions que le smartphone peut être utilisé comme un outil clinique pour la mesure de mouvement, sans distinction également de la marque des capteurs utilisés dans lesdits smartphones.
À partir de cette conclusion, nous proposons par la suite deux applications utilisant cette mesure de mouvement par le smartphone.

2.6.2 iBalance

Du fait de l'utilisabilité du smartphone pour mesurer des angles et des mouvements chez l'humain, nous avons testé deux preuves de concepts sur ce sujet. La première est iBalance, qui sera présentée dans cette section et qui travaille sur la capacité des personnes à maintenir une position debout immobile (travail et future rééducation de la proprioception).

2.6.2.1 Architecture du système proposé

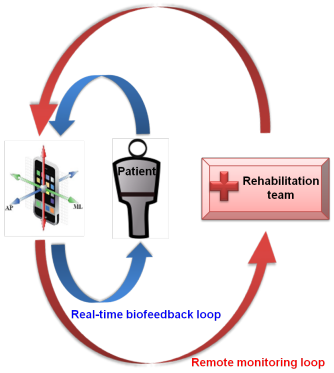


FIGURE 2.11 – Architecture du système iBalance

Le projet que nous avons nommé « iBalance-ABF » est un système, basé sur le smartphone, et qui offre, à l'intérieur de ce système portable, léger et communicant, un moyen d'entraîner et de rééduquer ses capacités de maintien de la posture. Il peut être utilisé par la personne en toute autonomie, chez elle par exemple (2.11, flèche bleue pour le retour sensoriel en temps réel). Le système permet également aux équipes médicales en charge de la rééducation d'évaluer quantitativement et objectivement les capacités de la personne au maintien de sa posture, de créer des programmes d'entraînement au maintien de l'équilibre personnalisé, mais aussi de suivre l'historique des progrès du patient au cours du temps (2.11, flèche rouge pour la boucle de suivi à distance).

Le principe sous-jacent de l'application est de fournir à l'utilisateur des informations supplémentaires à propos de la position de son tronc dans le plan médial-latéral ou bien dans le plan antéro-postérieur (ou les deux). Le type d'information fourni est le fait que la personne soit trop penchée d'un côté ou de l'autre de la position de base et cette information est fournie en fonction d'une zone morte réglable. Le retour sensoriel est fait soit par les écouteurs portés par la personne, soit par le vibreur du téléphone.
Par rapport aux travaux précédents, nous nous retrouvons dans le cas de mesures en quasi-statiques des angles du tronc. Le téléphone est, dans ce cas, attaché à la ceinture de la personne, dans son dos, au niveau du sacrum.

2.6.2.2 Retour sensoriel

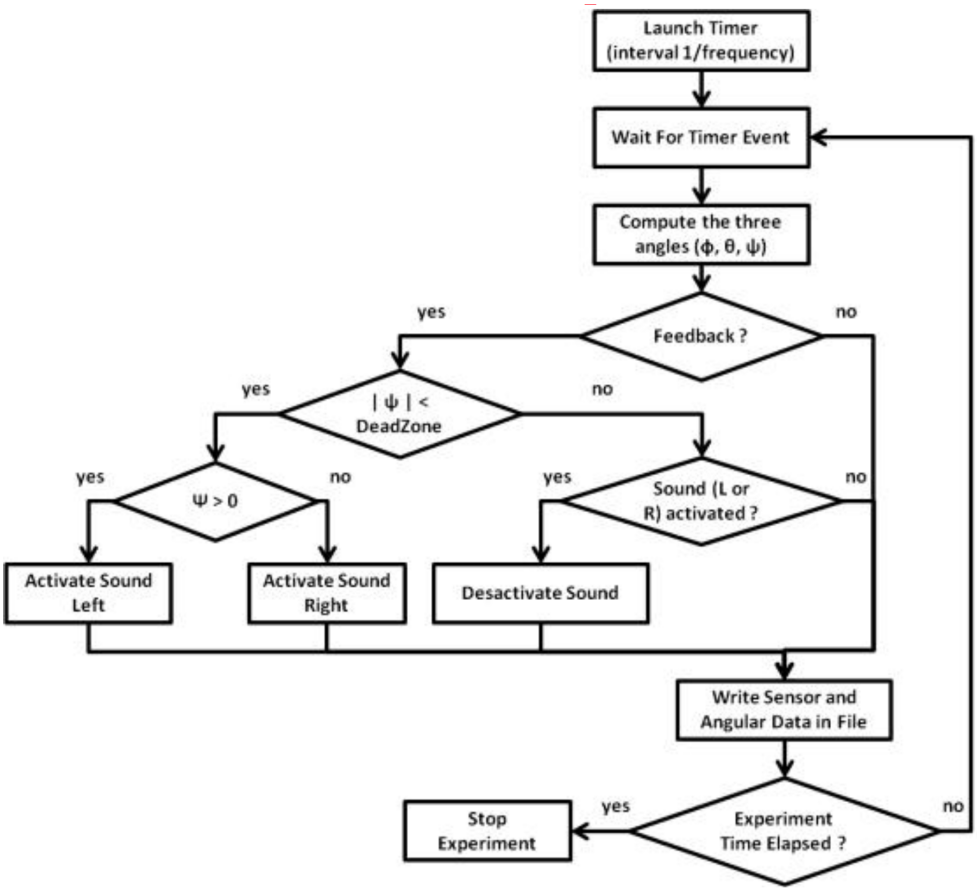


FIGURE 2.12 – Déclenchement du retour sensoriel auditif dans le système iBalance

Le système peut être utilisé avec ou sans retour sensoriel. La présence de ce retour sensoriel permet à la personne de perfectionner ou de réapprendre les sensations proprioceptives, l'hypothèse étant que des entraînements réguliers avec de tels systèmes améliorent durablement les performances même hors utilisation du système.
La figure 2.12 présente le déclenchement de ce système dans le cas d'un retour sensoriel audio et en effectuant des mesures sur le plan Médial-Latéral (ML). Le retour sensoriel dans ce cas est réglé pour donner une indication à la personne dans l'oreille correspondante (ou opposé selon le type de retour) à la dérive qui est mesurée. Lorsque la personne sort de la zone morte, elle est aussitôt prévenue de la direction dans laquelle elle en est sortie par ce retour (DOZZA et al., 2005). Dans cette preuve de concept, nous nous sommes basés sur le plan ML, car il se trouve fortement associé à des risques de chutes chez les personnes âgées (HILLIARD et al., 2008). Nous avons, pour de premiers tests, également réglé la zone morte à 1°. L'intérêt de cette zone morte est de ne pas solliciter la personne pour un écart trop faible avec la normale, sachant que cela entraînerait une plus forte instabilité posturale en obligeant à toujours corriger. Le retour sensoriel utilisé dans ces premiers travaux était de type répulsif. Lorsque l'angle du tronc se trouvait dans la zone morte, aucun signal n'était envoyé. Lorsque par contre l'angle était en dehors de la zone morte à gauche, une sonnerie était envoyée dans l'oreille gauche afin de « repousser » la personne de cette position (au contraire

des retours attractifs pourraient être utilisés, en se disant que l'on va attirer la personne vers la droite en la prévenant du côté droit - ce point sera abordé dans les travaux suivants). Tant que la personne n'a pas suffisamment corrigé sa posture pour retourner dans la zone morte, la sonnerie continue. Il est à noter que le type de retour sensoriel dans l'application que nous proposons peut-être simplement réglé, ainsi que sa zone morte, ses directions, etc.

2.6.2.3 Expérimentations et résultats

2.6.2.3.1 Protocole

Ces mesures effectuées ont pour but d'évaluer :

— L'utilisation du retour sensoriel par la personne, à savoir est-ce que le retour audio va servir pour maintenir une meilleure posture debout.

— La compréhension du système et des informations qu'il apporte par la personne. Compréhension des consignes pour un entraînement seul et un positionnement seul du téléphone.

Pour cela, nous avons fait tester le système par 26 personnes en tout, réparties en deux catégories :

— Vingt personnes (9 hommes et 11 femmes) jeunes (âge $26,5 \pm 3,7$ ans) et en pleine santé. Aucune d'elle n'avait d'historique de problèmes moteurs et/ou sensoriels ou de maladies ou troubles neurologiques.

— Six personnes plus âgées (âge $62,7 \pm 2,7$ ans) en pleine santé également avec les mêmes critères.

FIGURE 2.13 – Illustration du positionnement du smartphone et des deux positions de pieds (à gauche pieds parallèles, à droite pieds tandem)

Le protocole à suivre était simple. Les personnes devaient positionner elles-mêmes le téléphone au niveau de L5 (bas du dos - voir Fig. 2.13 à gauche)). Il était installé dans une ceinture. Elles portaient également des écouteurs attachés à ce téléphone et leur donnant toutes les indications sur le déroulement de l'expérimentation. Ensuite, elle devait maintenir une position debout et le plus stable possible, avec les yeux fermés, les bras le long du corps et ce pour une durée de 30 secondes (auxquelles s'ajoutaient 5 secondes de pré-mesures). Leurs pieds étaient positionnés soit parallèles l'un à l'autre avec un écart d'environ 10 cms, soit en position tandem (voir Fig. 2.13 à droite). Dans les essais, les personnes étaient prévenues qu'elles seraient soit en toute autonomie, soit qu'elles auraient une indication venant du téléphone (le retour audio).
Chacune de ces conditions (position du pied, absence/présence de retour) combinées était répétée 6 fois à la suite. L'ordre de présentation des différentes conditions était aléatoire. L'expérimentation dans son ensemble se déroulait en toute autonomie par la personne, le téléphone mélangeant aléatoirement les conditions et les enchaînant avec des pauses suffisantes.
Pour l'ensemble de ces mesures, le signal était recueilli à une fréquence de $50Hz$ tout au long de la mesure que le retour sensoriel soit actif ou non.
Les variables analysées étaient :

— le RMS de l'angle du tronc en degrés,

— une fréquence caractéristique (Mean Power Frequency dans un cas ou la Spectral Edge Frequency à 95% dans le second cas),

— le pourcentage de temps en dehors de la zone morte.

Les analyses statistiques ont été réalisées avec un t-test jumelé bilatéral. Nous considérons comme significatifs des tests donnant des p-valeurs de 0,05 (*) et 0,01 (**).

2.6.2.3.2 Résultats

Les résultats suivants présentent les différentes variables précédemment citées, dans les deux positions de pieds, et pour les conditions avec retour sensoriel audio (nommé ABF) et sans retour sensoriel audio (No-ABF).

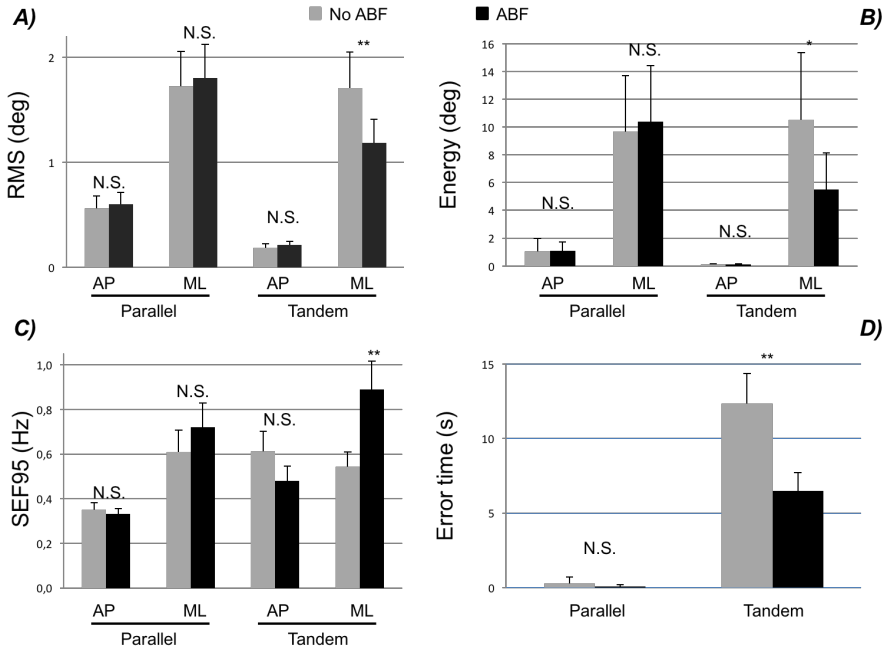


FIGURE 2.14 – Résultats sur les sujets jeunes. Les valeurs grises sont sans retour sensoriel audio (No-ABF) alors que les valeurs noires sont avec ce retour (ABF). La figure (A) présente le RMS en degré (la quantité de mouvement du tronc sur l'axe ML), la figure (B) présente l'énergie, la figure (C) la Spectral Edge Frequency à 95% et la figure (D) le temps en dehors de la zone morte.

Les expérimentations sur les jeunes personnes ayant été faites en premier nous avons plus détaillé les résultats. La figure 2.14 présente ces résultats. Nous y voyons l'effet du retour sensoriel audio qui apparaît clairement à chaque fois sur la position de pieds tandem et sur l'axe Médial-Latéral. Le retour sensoriel se faisant uniquement sur cet axe, nous voyons que les mouvements existants sur l'autre axe ne sont pas corrigés. De plus, la position pieds parallèle étant plus simple à maintenir, le retour est utilisé moins efficacement dans ce cas. L'effet se voit clairement également sur le temps passé hors de la zone morte qui est très bien compensé dans le cas du retour. L'augmentation de la fréquence à 95% se justifie par la réactivité des sujets aux consignes données par le retour sensoriel (condition ABF).
Les résultats pour les sujets plus âgés sont donnés par la figure 2.15. Ils vont exactement dans le même sens que pour les sujets jeunes à savoir qu'en pieds parallèles la différence n'est pas notable et en pieds tandem, les performances sont largement améliorées. Idem la fréquence augmente du fait de la réaction de la personne à la consigne.
Ces résultats démontrent bien une bonne utilisation de l'application et des informations qu'elle apporte par les personnes. C'est cohérent avec ce qui est annoncé par ailleurs dans la littérature, mais avant ces travaux, les résultats étaient obtenus avec du matériel plus conséquent et non utilisable en autonomie au domicile.

2.6.3 iProprio

L'application précédente iBalance visait la rééducation ou l'entraînement au maintien d'une posture debout stable de la personne. Une autre application que nous avons testée consiste en l'évaluation et l'entraînement de la proprioception

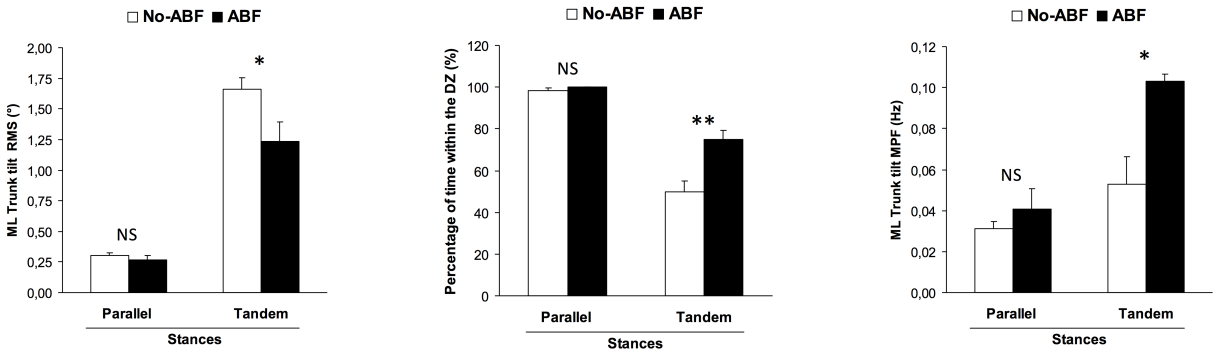


FIGURE 2.15 – Résultats sur les sujets plus âgés. De gauche à droite : Le RMS des mouvements du tronc en degré, le pourcentage de temps en dehors de la zone morte et enfin la fréquence donnant la moitié de la puissance (Mean Power Frequency)

sur un membre précis (donc le ressenti de la position spatiale de ce membre par la personne). Cela a été testé dans le cadre de la thèse de Quentin Mourcou comme application de ses travaux.

2.6.3.1 Présentation de l'application et du protocole de test

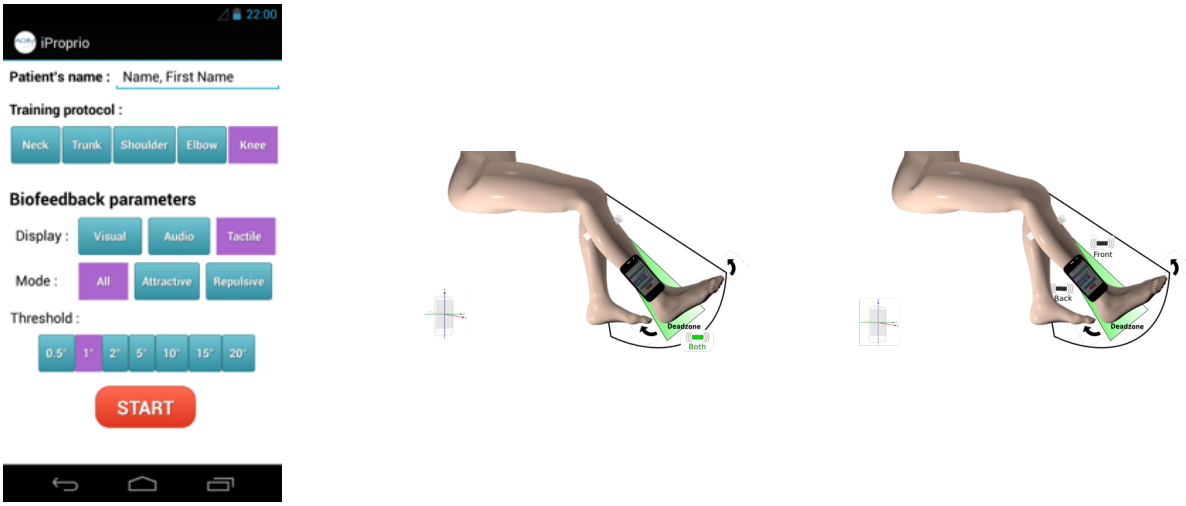


FIGURE 2.16 – Interface graphique et types de retours sensoriels

Le test utilisé pour vérifier la possibilité d'aider les personnes à réaliser des exercices de proprioception sur des membres précis est un test standardisé du genou appelé « active ipsilateral matching » qui est un test couramment utilisé et accepté en médecine (KNOOP et al., 2011). Ce test est l'un des tests embarqué dans l'application que nous proposons et, comme pour iBalance précédemment, un système de Text-To-Speech permet de vocaliser le test et de guider la personne afin qu'elle soit indépendante. Chaque participant est assis sur le coin d'une table, avec le segment tibial dans le vide et les yeux fermés. La position neutre est alors marquée comme la position de repos dans ce cas. Le participant doit alors mobiliser la jambe équipée du smartphone afin d'atteindre une position de 20 degrés par rapport à la position neutre, sans contrainte de temps, et doit marquer cette position. Ensuite, il doit retourner à l'angle précédent.
La tâche est faite dans trois conditions expérimentales : (1) sans retour sensoriel (notée BioFB OFF condition), (2) avec un retour sensoriel vibratoire répulsif (noté BioFB ON / Repulsive condition) et enfin (3) avec un retour sensoriel vibratoire attractif (BioFB ON / Attractive condition). La différence entre attractif et repulsif est montrée sur la figure 2.16, dans un cas le retour est donné dans le sens de l'erreur alors que dans l'autre il est donné dans le sens contraire. Chaque condition consiste en une série de 5 répétitions. L'ordre de présentation des conditions expérimentales est aléatoire et les performances ne sont pas indiquées en cours d'expérimentation. Le smartphone est attaché sur le côté de la fibule, au-dessus de la malléole. Le marquage de la position se fait avec un bouton connecté à part au smartphone.

Deux vibreurs connectés en bluetooth au smartphone sont attachés au niveau du tibia, à l'opposé l'un de l'autre.
Pour chaque essai, le participant démarre donc de sa position de repos, jambe dans le vide. Avant le premier test, cette position est calibrée comme position de base, marquée par un appui sur le bouton. Il reçoit alors comme consigne de se rendre à la position de 20° en étendant la jambe tant qu'il entend encore un bip sur le téléphone. Une fois ce bip entendu, la position est atteinte. Il peut alors retourner à sa position de base. Une fois ceci fait, il doit se repositionner à cet angle sans avoir l'indication sonore. Selon la condition, il n'aura soit rien du tout (1), soit un retour répulsif (2) ou attractif (3). La figure 2.16 montre l'installation du téléphone et de la personne ainsi que les différents retours sensoriels.
Ce protocole a été testé sur deux types de population. La première est un ensemble de seize sujets jeunes et sains (âge : $24,8 \pm 4,9$ ans) ne présentant aucun problème (connu) d'ordre moteur ou mental. La seconde population est de six personnes âgées (âge : $68,8 \pm 4,3$ ans), ne présentant également aucun trouble (connu).
Pour les analyses, deux variables dépendantes sont calculées afin de tester les performances des personnes :

— L'erreur absolue, en degrés, qui est la valeur absolue de la différence entre l'angle cible et la position reproduite. Cela mesure la précision globale du repositionnement du genou.

— L'erreur variable, en degrés, qui est la variance autour de l'erreur moyenne (erreur de repositionnement signée).

Une décroissance dans ces deux mesures indique un meilleur repositionnement du genou et donc une meilleure sensation proprioceptive.

2.6.3.2 Résultats

Pour analyser les résultats, la moyenne sur les cinq essais a été prise dans les trois conditions expérimentales précédemment citées ((1) sans retour sensoriel (notée BioFB OFF condition), (2) avec un retour sensoriel vibratoire répulsif (noté BioFB ON / Repulsive condition) et enfin (3) avec un retour sensoriel vibratoire attractif (BioFB ON / Attractive condition)). Une analyse de variance (ANOVA) unidirectionnelle est réalisée sur les trois conditions séparées. L'analyse post-hoc (Newman–Keuls) a été réalisée lorsque nécessaire et le degré de significativité est fixé à 0,05.

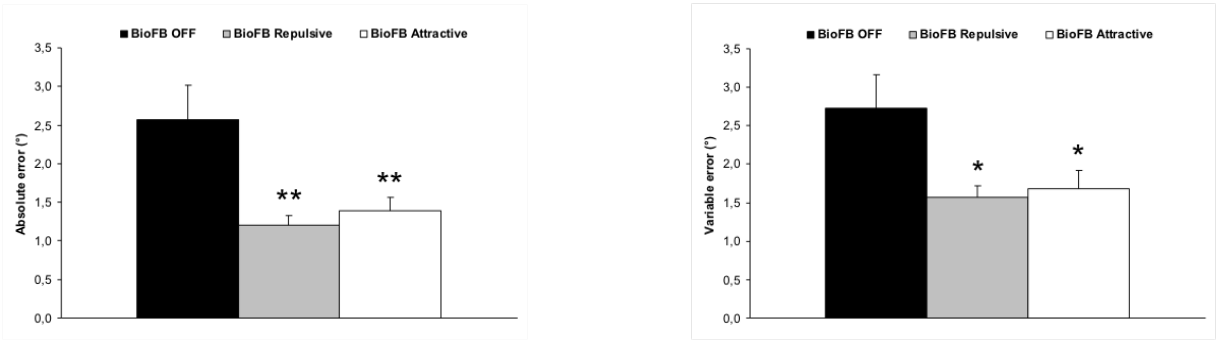


FIGURE 2.17 – Erreurs absolues et variables de repositionnement pour les sujets jeunes, avec ** : $P < 0.01$ et * : $P < 0.05$

Pour les sujets jeunes (Fig. 2.17)ß l'analyse de l'erreur absolue montre un effet principal de la condition ($F(2.30)=7.97$, $p < 0.01$). Idem, l'erreur variable montre un effet de la condition ($F(2.30)=5.15$, $p< 0.05$).
Ces résultats montrent donc une forte amélioration lorsque le retour sensoriel est présent, indiquant encore une fois une bonne utilisation de celui-ci, mais montrent par contre un même comportement entre les deux formes de retours sensoriels, montrant ici que les deux notions sont bien comprises indifféremment par la personne utilisant le système. La faible différence entre attractif et répulsif n'est pas significative. D'un point de vue théorique, et remettant dans le contexte du système de repondération sensoriel, et en se référant au système nerveux central responsable de la capacité à combiner des informations venant de plusieurs modalités sensorielles pour donner un aperçu du positionnement du corps dans l'espace, ces résultats suggèrent que le système nerveux central est capable d'intégrer cette information artificielle à travers la stimulation de la jambe, dans le but d'améliorer les performances proprioceptives de la personne. Ces travaux, au moment de leur publication, étaient originaux, car aucune publication n'avait montré encore l'utilisation de la modalité vibratoire pour améliorer la proprioception du genou.

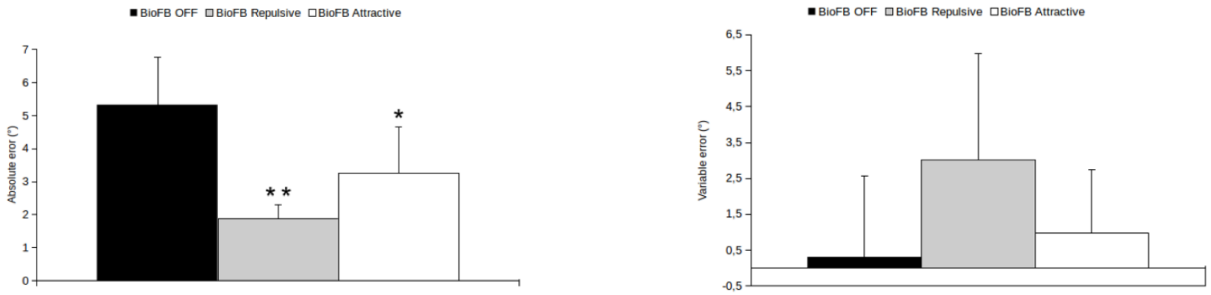


FIGURE 2.18 – Erreurs absolues et variables de repositionnement pour les sujets âgés, avec ** : $P < 0.01$ et * : $P < 0.05$

Les résultats sur les sujets plus âgés (Fig. 2.18) montrent de même que le système est bien utilisé du fait de la réduction significative de l'erreur absolue. L'erreur variable augmentant nous montre par contre une différence de comportement entre les deux à savoir que (comme l'erreur variable prend en compte le signe) les erreurs commises par les personnes plus âgées se concentrent dans l'une des directions lorsque le retour sensoriel est présent. Étant donné le faible nombre de participants, nous ne pouvons pas conclure plus sur ce phénomène observé.
Ces résultats montrent que le système iProprio pourrait avantageusement compléter les exercices réalisés en cabinet de kinésithérapie par exemple, en mesurant et en entraînant la personne à réaliser des mouvements.

2.6.4 Demande attentionnelle

L'un des points importants dans le maintien de la posture ou dans l'exécution d'une tâche de marche, en plus de la performance, est l'automaticité de cette tâche. Essayer d'observer ceci est l'un des défis des travaux précédents. Une seconde manière de faire ceci est d'utiliser le paradigme de la double tâche. Le paradigme de la double tâche requiert que le participant exécute deux tâches simultanées. Exécuter une tâche requiert l'utilisation d'une certaine partie des capacités du système nerveux, le fait d'en ajouter une seconde, avec des capacités limitées forcément, les performances de l'une des deux tâches ou des deux vont s'en trouver dégradées (KAHNEMAN, 1973; WICKENS, 1984). Quantifier la dégradation de la tâche posturale va permettre d'avoir des informations sur son caractère automatique pour la personne. Les personnes à fort risque de chute vont voir ces automatismes diminués et ainsi avoir une plus forte charge pour gérer leur équilibre. Une expérimentation simple a été mise en place pour mesurer la demande attentionnelle de manière automatique et en toute autonomie pour la personne. Cette partie va présenter succinctement les premiers travaux dans ce sens qui ont été effectués.

2.6.4.1 Système de mesure

Comme montré sur la figure 2.19, cette dernière application permet, en se basant sur les travaux précédent et l'architecture iBalance, de mesurer le temps de réaction d'une personne à un stimulus. Cela correspondra à la seconde tâche. La première est le maintien de la posture, tâche que l'on indique comme prioritaire à la personne, et la seconde est de réagir, aussi vite que possible, au stimulus auditif entendu, apparaissant à un moment non déterminable par la personne. Afin de permettre une mesure autonome et automatique, la personne est équipée du smartphone faisant les mesures et d'un casque. Celui-ci va d'une part permettre de donner des indications à la personne sur le protocole à suivre, mais aussi, et dans le cas de la présence d'une seconde tâche, le microphone va mesurer le bruit ambiant et son évolution afin de détecter le moment de réponse de la personne au stimulus auditif.

2.6.4.2 Expérimentation

Les expérimentations ont été effectuées sur une population de huit jeunes personnes en bonne santé (âge moyen 29 ± 3 ans, taille 174 ± 9 cm, poids 66 ± 15 kg, moyenne $\pm$ Écart Type). Les participants portent un smartphone à la ceinture, au niveau de la vertèbre L5, comme dans les précédents travaux iBalance. Le smartphone est équipé d'un casque et d'un microphone.

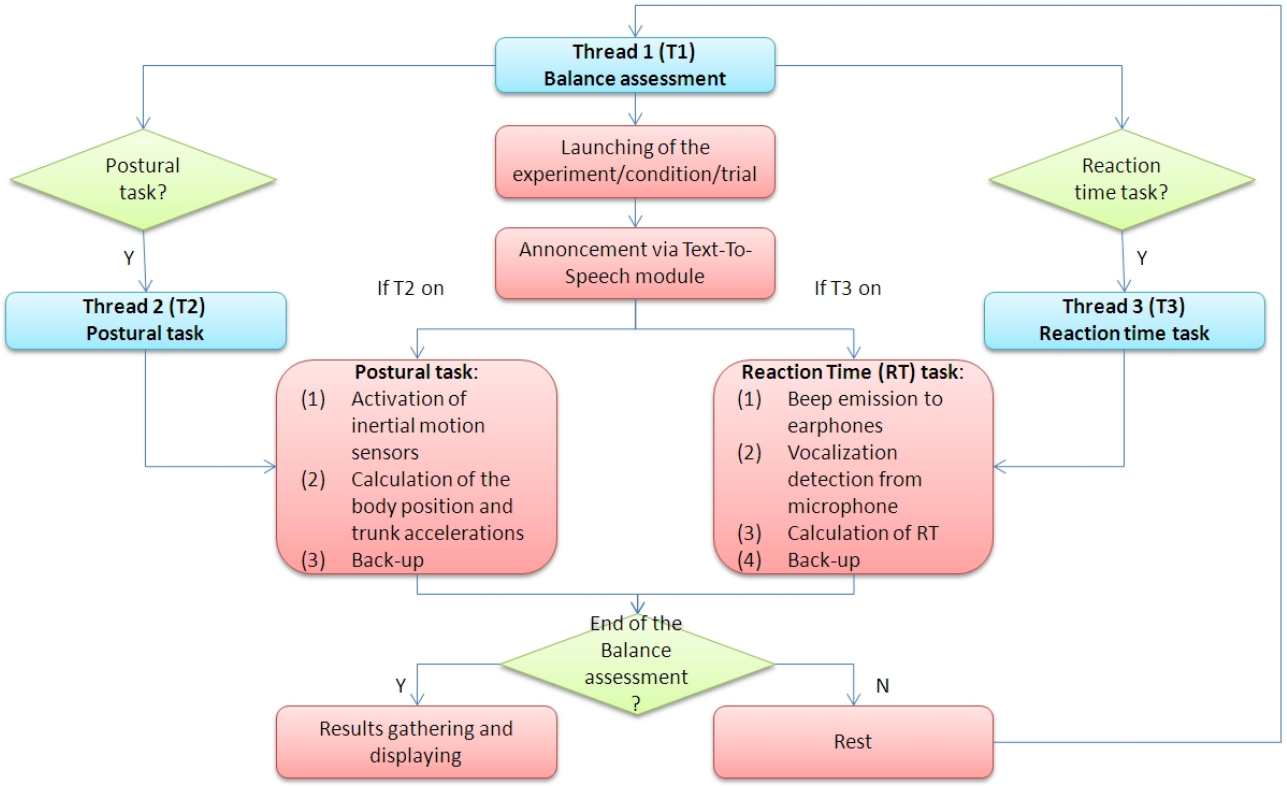


FIGURE 2.19 – Architecture du système de mesure de la demande attentionnelle

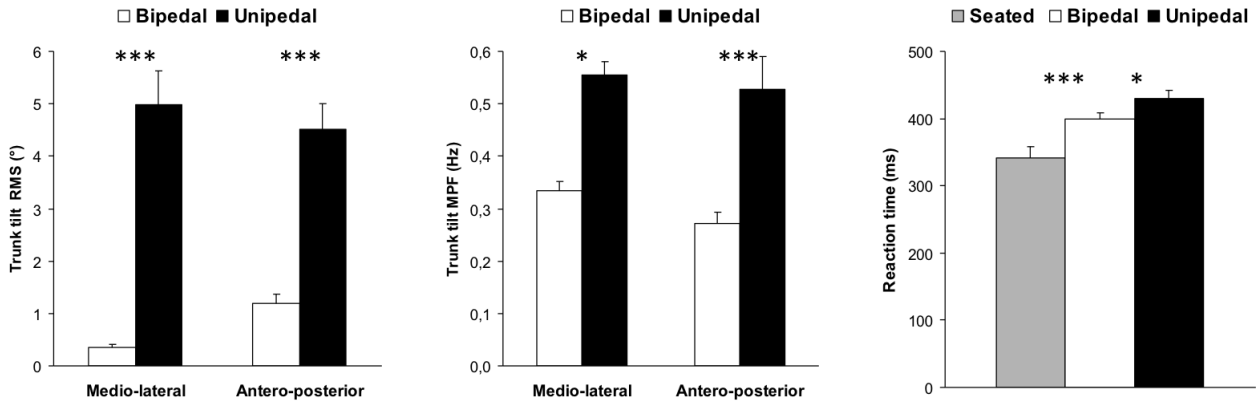


FIGURE 2.20 – Résultats sur les trois variables analysées à savoir le RMS, la Fréquence donnant la moitié de la puissance cumulée et le Temps de réaction, pour différentes position, avec *** : $P < 0.001$ et * : $P < 0.05$

2.6.4.3 Analyse de données et résultats

De la même manière que précédemment, le mouvement du tronc est scruté lors de ces expérimentations, avec notamment le RMS du tronc sur les plans Médial-Latéral et Antero-Postérieurs, la fréquence donnant la moitié de la puissance cumulée, et à tout ceci nous ajoutons le temps de réaction sur la seconde tâche, à savoir le temps mis par la personne en moyenne à répondre au stimulus auditif.

Cinq essais par personne sont effectués dans chacune des conditions expérimentales (avec et sans stimulus auditifs, en position assise, bipédale ou unipédale.

Les premiers résultats de cette expérimentation sont montrés sur la figure 2.20

Des ANOVA avec trois groupes (les trois postures) ont été appliquées sur ces données et ont montré que la difficulté de la tâche engendraient, comme attendu, des temps de réaction plus longs. Nous retrouvons, dans un système totalement intégré et facilement utilisable en autonomie, les résultats de la littérature. Il nous faut continuer ces travaux afin

d'extraire des informations plus intéressantes, notamment sur l'évolution des différences de temps de réaction et l'évolution des valeurs des temps pour une personne donnée, afin de mesurer l'évolution de la personne et de relier ceci à une possible apparition d'une dégradation du bien-être ou de la santé de la personne. Que dire de l'application des notions d'entropie au moment du déclenchement de la seconde tâche ? Cela reste une piste à envisager.

2.7 Conclusion

Cette première partie des travaux, qui se concentre sur des signaux relatifs au fonctionnement « interne » de la personne, a montré que les contributions qui sont présentées ici pour la mesure permettent de faire progresser les connaissances en traitement du signal sur ces domaines et améliorer l'interprétation des signaux. Pour ce qui est de la mesure d'indice à partir de données de plateformes de force, des méthodes non-linéaires basées principalement sur la notion d'entropie et sur son interprétation ont été décrites. Pour les données issues des centrales inertielles des smartphones, ce sont des algorithmes de reconstruction des données qui ont été proposés, avant de se pencher sur les modalités de retours sensoriels et l'extraction de scores à faire afin de pouvoir faire un retour au patient et à l'équipe médicale.
Les contributions principales de ce chapitre sont les suivantes :

— Des contributions méthodologiques en traitement du signal permettant de mieux connaître la personne au travers de mesures plus ou moins simples (notamment basées sur l'entropie), avec des applications futures prévues dans le domaine de la prédiction de risques ou d'effets de maladies chroniques.

— La proposition de systèmes basés sur le smartphone et permettant de mesurer, de manière intégrée, simple, et en toute autonomie pour la personne, diverses variables relatives à la réalisation d'exercices plus ou moins simples liés à la posture ou au mouvements de membres (proprioception). Ces systèmes permettent également la délivrance d'informations à la personne par l'intérimaire de retours sensoriels afin de permettre des entrainements et une amélioration de la condition ou de la rééducation de la personne.

Ces travaux doivent être poursuivis afin de préciser plus finement l'interprétation des variables, de leurs évolutions et leur relation avec un état général de la personne. Une thèse est en cours sur l'acquisition des données smartphone sur un très long terme afin d'en déduire des informations sur le bien-être de la personne. Pour l'instant les résultats se concentrent tout d'abord sur une interprétation du signal en terme de posture et de mouvements quelque soit la position du téléphone.
Il est à noter aussi que les travaux proposés sur smartphone s'inscrivent dans un domaine très porteur actuellement pour l'industrie et les services et pourraient donner lieu à des transferts technologiques notamment par l'intermédiaire de la SATT (suite aux deux brevets déposés). L'amélioration continue d'une preuve de concept est faite pour permettre ce transfert à plus ou moins long terme.
Suite à cette présentation, nous allons maintenant explorer des travaux qui, cette fois, sortent des automatismes de la personne pour aller tout d'abord vers la reconnaissance de la personne et le suivi de ses actions.

CHAPITRE 3

La personne et ses actions : extraction de caractéristiques et reconnaissance de formes

Sommaire

2.1 Introduction . . . 7
2.2 Objectifs de ces travaux et verrous scientifiques à lever . . . 8
2.3 État de l'art sur la mesure de données physiologiques et de données inertielles . . . 8
2.3.1 Les données cardio-respiratoires . . . 8
2.3.2 Les données de posture et de marche . . . 9
2.4 Analyse des signaux posturaux avec l'entropie d'échantillon . . . 12
2.4.1 L'entropie d'échantillon . . . 12
2.4.2 Application au contrôle postural . . . 13
2.5 Analyse non linéaire des signaux par décomposition : modification de l'entropie spectrale pour la construction d'un indice local de complexité . . . 17
2.5.1 L'entropie spectrale (Wavelet Entropy) . . . 17
2.5.2 La transformée par synchrosqueezing . . . 18
2.5.3 Calcul de l'entropie spectrale modifiée WE_{SQT} . . . 18
2.5.4 Tests sur des signaux posturaux et protocole . . . 18
2.5.5 Résultats et discussion . . . 19
2.6 Les signaux du smartphone : vers l'application iProprio . . . 20
2.6.1 Test de la validité de ces mesures . . . 21
2.6.2 iBalance . . . 26
2.6.3 iProprio . . . 29
2.6.4 Demande attentionnelle . . . 32
2.7 Conclusion . . . 34

3.1 Introduction et verrous scientifiques à lever

De nos jours, les villes, les transports en commun, les lieux publics et de nombreux lieux privés sont équipés de systèmes de vidéo-protection (ou vidéo-surveillance). Elle est devenue l'une des solutions les plus considérées (voire la plus) lorsqu'il s'agit de parler de notre sécurité. Cependant, dans un contexte d'une telle utilisation, il est compliqué de scruter les images à tout instant et par conséquent, des thématiques de recherches se sont développées ces dernières années pour tenter d'automatiser ces traitements et de détecter des situations anormales dans ces séquences.
La ré-identification de personnes, de leurs actions et leur suivi dans des réseaux de caméras est alors devenu un fonctionnalité d'intérêt pour de tels systèmes (TRUONG CONG, KHOUDOUR, ACHARD, MEURIE et al., 2010; NILSSON et al., 2016; XU et al., 2016).
Un second domaine de recherche d'intêret lié au traitement vidéos est la reconnaissance de gestes ou d'actions (J. K. AGGARWAL et al., 2014; CHANG et al., 2011; CLARK et al., 2012; XIA, C.-C. CHEN et J. AGGARWAL, 2012). Ces recherches permettent d'avoir diverses applications telles que la sécurité (avec la possible détection de situations

critique sur des flux vidéos), la santé (avec l'analyse de situations de dangers pour une personne surveillée à domicile comme la chute) ou encore l'interaction homme-robot, en analysant le comportement et les gestes de la personne pour entrer en interaction avec elle ou pour proposer des services.
Les travaux décris dans ce chapitre abordent l'extraction d'informations, le traitement du signal et la classification, qu'elle soit en ligne ou hors-ligne. L'objectif de ces travaux est d'améliorer les connaissances en termes de reconnaissances des humains et de leurs actions.
Les verrous scientifiques visés par ces travaux sont les suivants :

1. Comment peut-on utiliser au mieux les données présentes et les attributs extraits pour construire des modèles les plus performants possibles et qui généralisent au mieux la forme ou la modalité que nous recherchons.
2. Comment créer au mieux des algorithmes de classification afin qu'ils soient les plus efficaces possible pour modéliser des distributions non stationnaires (et donc avoir des classifieurs évolutifs).
3. Comment traiter efficacement et en ligne des quantités de données importantes pour obtenir une décision rapide et peu coûteuse en bande passante et en temps de traitement.

Deux domaines de recherche vont être investigués dans ce chapitre, avec une première section décrivant les travaux sur l'identification et la ré-identification de personnes dans des réseaux de caméras (section 3.2) et la reconnaissance de gestes et d'actions avec une caméra RGB-D incluant une analyse en ligne et une reconnaissance précoce (section 3.3). Avant cela, une présentation des objectifs et des verrous scientifiques de ces travaux va être faite.

3.2 Identification/ré-identification de personnes dans des réseaux de caméras : de l'extraction d'information à la classification en ligne

Des chercheurs du département IA avaient par le passé fait des recherches, dans le cadre notamment du projet ANR Canada, sur le suivi d'activité anormales et la reconnaissance de personnes dans des réseaux de caméras de vidéo-surveillance. Étant donné également un long historique de travaux sur les algorithmes de classification incrémentaux, mes recherches ont contribué à davantage explorer ces algorithmes incrémentaux sur des données de type vidéo avec possiblement un grand nombre de classes et peu de variabilité. Dans cette section, les travaux décrits porteront donc sur cette identification/ré-identification de personnes afin de quantifier l'efficacité des algorithmes et les problèmes induits par le type de données sur ceux-ci. Cette description se divisera en une partie extraction d'information, suivie par une partie sélection, puis deux parties apprentissage dynamique, l'une en classification pour la ré-identification et la suivante en clustering pour la l'identification. Dans le premier cas, nous avons un ensemble de personnes connues et nous voulons déterminer, sur une image, laquelle de ces personnes de notre base d'apprentissage est présente. Dans le second cas, nous ne connaissons ni le nombre de personnes ni n'avons d'exemples de l'ensemble de celles-ci et nous voulons, dans chacune des images, tenter de soit retrouver une personne connue des précédentes itérations, soit créer une nouvelle personne dans notre base pour laquelle nous pourrons ensuite retrouver différentes images correspondant à celle-ci.

3.2.1 Acquisition et traitement de données vidéos

Un nombre significatif de recherches a été mené dans le champs de la reconnaissance de personnes, basées sur des caractéristiques biométriques (visage, démarche, iris, etc.) (ROARK et al., 2003 ; KAZISKA et al., 2007 ; ZHOU et al., 2008 ; ZHOU et al., 2007), mais aussi sur l'apparence (MAKRISA et al., 2009 ; YOON et al., 2006 ; HÖRSTER et al., 2007) et spécifiquement sur des applications de détection de personnes à pied dans la rue et dans des réseaux multicaméras (TRUONG CONG, KHOUDOUR, ACHARD et DOUADI, 2010). L'apparence se définit par l'ensemble des parties visibles de la personne (vêtements, parties du corps, etc.) et peut être facilement obtenue par extraction du fond (pour isoler la personne dans l'image) sans requérir une définition d'image trop importante (comme peuvent le faire certaines caractéristiques biométriques). Pour une courte période de temps, l'apparence d'une personne est supposée invariante (en supposant une même orientation vis-à-vis de la caméra, une même condition d'illumination et donc aucune différence dans la perception des couleurs par la caméra, etc.). Lorsque l'on considère des périodes temporelles

de plus grande taille, l'apparence peut varier, surtout dans un réseau de caméras, dans lequel chacune des caméras a son propre environnement lumineux (qui peut de surcroit évoluer dans la journée), mais a aussi ses spécificités techniques (réponse des capteurs CCD). La personne va aussi changer d'orientation, etc. Même une base statique importante d'images ne pourrait pas prévoir toutes les conditions et tous les cas rencontrés. Ainsi, dans nos travaux, nous nous sommes tournés vers des algorithmes d'apprentissage en ligne et adaptatifs afin d'avoir cette capacité de reconnaissance de situations non encore présentées au classifieur.

La reconnaissance d'objets ou de personnes nécessite de travailler sur un ensemble d'attributs caractéristiques de ce qui doit être reconnu. Des attributs adaptés doivent correctement représenter un objet ou une personne tout en différenciant cet objet ou cette personne des autres. Les attributs couramment utilisés pour reconnaitre une personne et la définir sont basés sur le visage (et ses caractéristiques géométriques), la démarche, la forme du corps et l'apparence (YOON et al., 2006; HÖRSTER et al., 2007; ZHOU et al., 2007; TRUONG CONG, KHOUDOUR, ACHARD et DOUADI, 2010). L'apparence d'une personne contient ses vêtements et ses parties visibles du corps. C'est une propriété de la personne. Cependant, comme pour tout application impliquant un capteur, ici la caméra, la perception de cette apparence dans l'image extraite va fortement dépendre des conditions et de l'appareil de mesure. En ce qui concerne les caractéristiques de couleurs, elles sont faciles à obtenir. De plus, les caractéristiques couleurs sont basées sur des caractéristiques de l'image qui sont invariantes en translation, rotation et non sensible à des effets d'échelle si une normalisation correcte est réalisée. Dans les travaux réalisés et présentés dans cette partie, nous nous sommes concentrés d'abord sur la partie extraction d'informations en prenant des attributs basés sur la couleur et sur la texture, pour définir la personne à reconnaître (en considérant un nombre important de personnes à reconnaître).

Les travaux décrits dans cette section sont l'objet d'une partie de la thèse de Yanyun Lu. Ces travaux incluent une partie sur le calcul et la sélection des attributs et une partie sur la classification. La suite de ce court état de l'art concerne les données issues de caméras RGB-D.

3.2.2 But et originalité des travaux

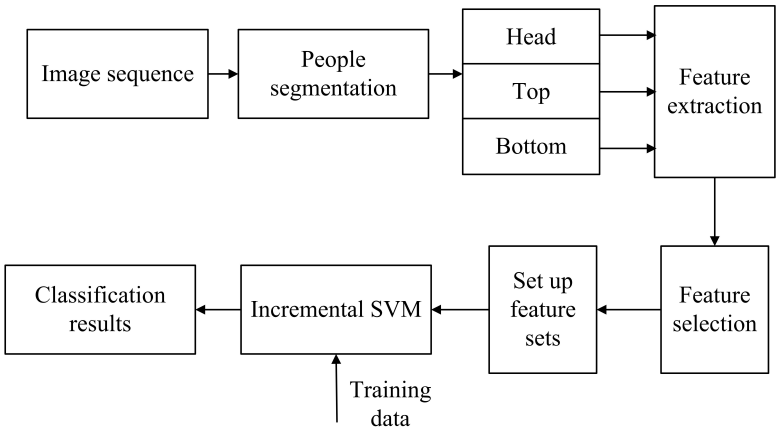


FIGURE 3.1 – Structure du système de reconnaissance de personnes proposé

Comme décrit par la figure 3.1, le but de ces travaux est de proposer une chaîne de traitement complète pour la reconnaissance de personnes dans des réseaux de caméras. L'originalité des travaux réside dans le fait d'appliquer, sur ces données, des algorithmes d'apprentissage dynamique, afin d'être en mesure soit de reconnaître des personnes déjà connues sur leur apparence mais avec des propriétés de captation différentes (caméra, luminosité, angle de vue, etc.), soit « d'identifier » les personnes sans bases de données connues à l'origine (et de créer ainsi un cluster par personne dans l'idéal).

3.2.3 Base de données de test

FIGURE 3.2 – Une personne et les six marches différentes dans la base CASIA

Pour tous les tests de cette section, nous avons utilisé la base de données CASIA [1]. Cette base est à l'origine une base de données de reconnaissance de marche, elle présente vingt personnes différentes, en train de marcher, selon six orientations. Elle nous semblait parfaite pour notre utilisation qui ne consistait qu'à reconnaître des personnes dans différentes situations du fait de ces changements d'orientation. Les six orientations de la base sont montrés sur la figure 3.2.

3.2.4 Extraction des caractéristiques

Comme évoqué dans la section précédente, nous nous sommes basés sur des attributs simples qui sont facilement extraits du signal vidéo. Nous avons donc pris les attributs liés à la couleur et à la texture.

Ces attributs simples ont été pris sur les trois parties du corps à savoir tête, buste et jambes afin de décrire au mieux et au plus simple la personne. Cette séparation en trois parties du corps se base sur la manière dont nous, humains, nous pourrions décrire une personne au plus simplement. Le cas se présente, par exemple, lorsque les autorités lancent une alerte enlèvement et décrivent simplement l'enfant enlevé et ce qu'ils connaissent du suspect. Nous utiliserions alors ces trois parties du corps pour lesquelles nous avons extrait les caractéristiques de couleur et de texture. Les différents attributs que nous avons retenus au départ sont ceux basés sur les valeurs moyennes et les écarts types dans les différentes composantes couleur, ainsi que l'énergie dans des bandes d'histogramme. Les attributs couleur sont extraits sur les trois composantes (Rouge, Vert et Bleu) du signal vidéo. Ces composantes variant en fonction de l'illumination, de la qualité de la caméra, de la réflexion sur les surfaces, etc., une normalisation est nécessaire. Nous nous sommes basés sur la normalisation « Grey-World » qui considère que les changements dans le spectre d'illumination peuvent être modélisés par trois facteurs appliqués aux trois composantes couleurs pour obtenir une invariance et être plus robuste aux variations d'intensité d'illumination (FINLAYSON et al., 1998).

Les attributs de texture sont calculés en prenant en compte la co-occurence spatiale des pixels. Nous nous sommes basés sur les attributs définis par Haralick (HARALICK et al., 1973) qui donne ainsi 13 attributs différents.

Sur ces couleurs et ces textures, nous calculons donc un ensemble de 93 attributs pour chaque image représentant une personne : 54 basés sur la couleur et 39 basés sur la texture. Ces ensembles d'attributs sont très rapides à calculer.

La base de donnée initiale, nommée CASIA-Wholeset, est constituée des 20000 images de cette base (environ 1000 pour chacune des classes) représentées par ces 93 dimensions.

1. Base de données publique « CASIA Gait Database » créée en 2001 et pouvant être récupérée en ligne : http://www.sinobiometrics.com.

Type d'attribut	Description
Couleur	Valeur moyenne pour R', G' et B'
	Écart type pour R', G' et B'
	Histogramme à 4 valeurs pour R', G' et B'
Texture	Énergie
	Correlation
	Inertie
	Entropie
	Inverse du moment de la différence
	Somme des moyenne
	Somme des variances
	Somme de l'entropie
	Différence des moyennes
	Différence des variances
	Différence de l'entropie
	Mesure de l'information de corrélation 1
	Mesure de l'information de corrélation 2

TABLE 3.1 – Ensemble initial d'attributs sur la couleur et la texture

3.2.5 Sélection et résultats

Afin d'appréhender l'importance des attributs, nous avons testé des algorithmes de sélection d'attributs sur ceux-ci.

Pour représenter une personne, nous pourrions utiliser un vecteur de grande dimension (comme ici 93) qui pourrait, nous l'espérons, mener à de bonnes performances de classification. Cependant, plus la dimension du problème augmente, plus la classification est longue (la complexité dépend du nombre d'exemples mais aussi du nombre de dimensions), mais surtout plus celle-ci est rendue difficile. On parle en général de malédiction de la dimension (curse of dimensionality), car dans le cas de grandes dimensions, moins les mesures de distances faites par les classifieurs n'ont de « sens » (compensation entre dimensions, importance de l'une par rapport à l'autre ou encore problème de calcul de normes). Afin d'éviter les dimensions inutiles ou redondantes, et par la même occasion de réduire le temps de calcul, nous utilisons des algorithmes de réduction de la dimension sélection d'attributs.

Les méthodes de sélection d'attributs sont de trois types, selon la manière dont la recherche de l'ensemble optimal est faite : les méthodes dites filtres (boucle ouverte), les méthodes dites Wrapper (boucles fermées) et les méthodes embarquées (boucle fermée aussi, peuvent être vues comme des méthodes Wrapper)(JOHN et al., 1994 ; SAEYS et al., 2007). Les méthodes dites filtres travaillent directement sur les données sans considérer l'algorithme de classification qui sera ensuite utilisé. L'évaluation d'un ensemble d'attributs (avec un critère ou une heuristique), par conséquent, ne dépend que de propriétés inhérentes à l'ensemble des données (distribution des valeurs, corrélation entre attributs, corrélation attribut-classe, etc.). Les méthodes Wrapper utilisent aussi un critère ou une heuristique pour évaluer les différents ensembles, mais, au contraire des méthodes filtres, le critère ou l'heuristique utilisée va dépendre des performances de classification d'un algorithme choisi sur l'ensemble sélectionné. Les méthodes embarquées quant-à-elles vont s'intégrer au classifieur pour faire cette sélection d'attributs.

Un large choix de méthodes de sélection d'attributs existe, dans les travaux de thèse de Yanyun Lu nous avons décidé d'en retenir que trois qui nous paraissaient pertinentes à tester. La première méthode sélectionnée est l'Analyse en Composante Principale (ACP). Cette méthode se base sur les propriétés et la distribution des données pour déterminer un nouvel espace d'attributs (pour lequel chacune des dimensions est une combinaison linéaire des dimensions de départ) dans lequel les données sont efficacement représentées (car dans les directions de plus grande variance). Cette méthode fait partie des méthodes filtres. La seconde, basée sur la corrélation (entre les attributs et avec la classe) utilise une heuristique pour déterminer un meilleur et plus petit ensemble d'attributs. Au contraire de l'ACP, cette seconde sélection va conserver des attributs de base en les retenant et non construire un nouvel espace. Enfin, nous avons également testé une méthode Wrapper. Pour celle-ci, comme énoncé dans le paragraphe précédent, nous cherchons à maximiser les performances de classification. Nous allons donc choisir plus ou moins de dimensions en fonction de ces performances avec un classifieur donné. Le tableau 3.2 résume les résultats obtenus pour cette sélection d'attributs.

3.2.5.1 ACP

La sélection d'attributs basée sur l'ACP vise à réduire le nombre de dimensions en limitant la perte d'information au sens de la variance, en projetant de nouveau les données dans une nouvelle base orthogonale (PEARSON, 1901). Dans les travaux de Yanyun Lu nous avons conservé 95% de la variance globale des données initiales. Nous sélectionnons donc un nombre de dimensions suffisantes pour atteindre ce quota.
Au final nous arrivons à un ensemble de 26 attributs, créés à partir de la distribution des données, et chaque attribut est une combinaison linéaire des 93 attributs initiaux. La projection de nos données initiale dans cette nouvelle base orthogonale constituera l'ensemble « CASIA-PCA ».

3.2.5.2 Sélection basée sur la corrélation

La sélection d'attributs basée sur la corrélation (Correlation-based Feature Selection - CFS) est un filtre simple qui attribue un score à chaque sous-ensemble d'attributs en fonction d'une heuristique appelée le mérite (HALL, 1999). Le mérite a été défini comme :

$$M_s = \frac{k \cdot \overline{r_{cf}}}{\sqrt{k + k \cdot (k-1) \cdot \overline{r_{ff}}}}$$

avec k le nombre d'attributs dans l'ensemble courant, $\overline{r_{cf}}$ la valeur moyenne de la corrélation attribut-classe pour chaque attribut $f \in S$ de notre ensemble courant, et $\overline{r_{ff}}$ la corrélation moyenne entre attributs pris par paires distinctes.
En plus de cette heuristique, il nous faut une méthode de recherche dans les ensembles d'attributs. Nous avons choisi la méthode « Best First » qui est une méthode qui partira d'un ensemble vide d'attributs et va explorer en cherchant toujours à aller vers le meilleur ensemble (meilleure amélioration par rapport à l'actuel). Un critère d'arrêt à définir permet de voir lorsque l'on n'améliore pas pendant un certain nombre d'itérations soit d'aller vers la solution suivante (pas la meilleure, mais la seconde, etc.) soit de s'arrêter si la recherche est terminée. Ce type de recherche permet de trouver des optimaux (pas globaux par contre) et permet d'éviter les petits optimums locaux.
Pour notre ensemble initial, nous arrivons à une solution contenant 40 attributs qui sont trouvés comme les plus représentatifs, avec le moins de redondance possible entre-eux et le plus de corrélation avec la classe à trouver. Ce nouvel ensemble, nommé CASIA-CFS, est décrit dans le tableau 3.2. Nous y voyons que les attributs de texture sont moins présents que les attributs de couleur et que la partie basse du corps est aussi représentée que la partie haute du corps, et ces deux parties beaucoup plus que la tête.

3.2.5.3 Sélection de type Wrapper

La sélection de type Wrapper, initialement décrite par JOHN et al. (1994), va, similairement à la méthode précédente, utiliser un algorithme de recherche afin de parcourir au mieux les différentes combinaisons d'attributs. La différence notable avec la méthode précédente est que l'heuristique permettant de vérifier la performance d'un ensemble sera basée sur les performances de classification en utilisant ledit ensemble (par exemple avec le taux d'erreur global). Par conséquent, le temps d'exécution sera beaucoup plus long (du fait de la nécessité à chaque sous-ensemble donné de construire le modèle et de tester celui-ci). Sur un ensemble de départ à 20000 éléments comme le nôtre, cela peut s'avérer long. L'avantage de cette méthode est que la sélection est meilleure, car elle prend en compte les capacités intrinsèques de l'algorithme d'apprentissage (qui dans notre cas, se basant sur des projections de l'espace de départ, sera peut-être influencé par d'autres paramètres qu'une simple corrélation).
Sur les 93 attributs initiaux, seuls 16 ont été retenus pour créer ce sous-ensemble optimal comme présenté dans le tableau 3.2. Ce nouvel ensemble y est nommé CASIA-Wrapper. Les attributs couleur, surtout par leur moyenne et écart type, y sont bien représentés et pour la texture, l'entropie se retrouve sélectionnée dans toutes les parties du corps.

3.2.6 Discussion sur les attributs extraits

Comme décrit dans les sections précédentes, quatre ensembles de données ont été préparés en vue d'être utilisés dans une future étape d'identification de personnes en utilisant des attributs simples à calculer dans un espace de couleur

Ensemble	Nombre d'attributs	Description
CASIA-Wholeset	93	Attributs de couleurs et de texture initiaux
CASIA-PCA	26	Combinaison linéaire des attributs initiaux
CASIA-CFS	40	5 attributs couleur sur la tête 13 attributs couleur sur le haut du corps 14 attributs couleur sur le bas du corps 1 attribut texture sur la tête 4 attributs couleur sur le haut du corps 3 attributs texture sur le bas du corps
CASIA-Wrapper	16	4 attributs couleur tête 4 attributs couleur haut du corps 4 attributs couleur bas du corps 1 attribut texture tête 2 attributs texture haut du corps 1 attribut texture bas du corps

TABLE 3.2 – Attributs sélectionnés par les différentes méthodes.

simple et compréhensible. Le tableau 3.2 donne un aperçu de la sélection pour chacun de ces ensemble de données. Pour l'analyse en composante principale, nous obtenons 26 attributs sur les 93 initiaux (conservation de 99,6% de la significativité). Cependant, le désavantage de cette méthode est la perte de l'interprétation des attributs (étant maintenant une combinaison linéaire des anciens attributs, et donc plus difficile à exprimer selon un sens physique), et la forte dépendance aux données, un ensemble de données différentes donnera des vecteurs de projection différents alors qu'une sélection d'attributs plus classique pourra donner le même résultat.
Dans CASIA-CFS, la plupart des attributs sélectionnés sont basés sur la couleur. Si l'on compare CASIA-CFS et CASIA-Wrapper, 11 attributs sont en commun dans les deux ensembles (sur 16 pour Wrapper) : h_mean_r, h_std_r, h_mean_b, t_mean_r, t_std_r, t_mean_b, t_std_b, b_mean_r, b_std_r, b_std_g, b_mean_b (h représentant la tête, t le haut et b le bas, mean et std la moyenne et l'écart type, et les r, g et b en troisième partie du nom sont les couleurs). De plus, lorsque nous regardons plus attentivement les valeurs de la matrice de covariance permettant d'effectuer l'ACP, nous pouvons noter que les trois méthodes sont cohérentes étant donné que les plus fortes valeurs se retrouvent sur les attributs sélectionnés dans les deux autres méthodes.
Il était évident que les attributs de couleurs seraient les plus utiles pour la classification de personnes dans notre système (c'est ce qui décrit le mieux la personne lorsque nous la regardons). Nous pouvons aussi remarquer que, pour ce qui est de la texture, les valeurs d'entropie sont sélectionnées à chaque fois comme étant dans les plus informatives. Ces résultats seront utilisés dans la section suivante pour identifier et ré-identifier des personnes à partir d'un réseau de caméras.

3.2.7 Techniques d'apprentissage pour l'identification et la ré-identification de personnes

La reconnaissance de personnes dans le cadre de la vidéosurveillance est un thème très important de nos jours. Nos travaux s'intéressent à la reconnaissance automatique, avec peu de données disponibles lors de la phase initiale d'apprentissage, et donc une reconnaissance qui nécessite des algorithmes de classification en ligne. Sur la base CASIA précédente, contenant 20 personnes présentées avec des orientations différentes et en marchant, dans un premier temps nous donnons, dans cette section, les résultats de la classification de ces personnes avec tout d'abord un algorithme de SVM incrémental, qui a une connaissance de toutes les classes, mais avec un très faible nombre d'éléments pour chaque classe à l'apprentissage puis s'adapte (quelques images servent à initialiser le modèle qui est ensuite renforcé par l'arrivée de nouvelles observations). Dans un second temps, un autre algorithme dénommé SAKM, est exploité afin de permettre l'apprentissage de nouveautés et donc la découverte de nouvelles personnes arrivant.

3.2.7.1 Motivation et origines des travaux

Dans un système permettant de reconnaître des personnes dans un signal vidéo, l'apprentissage de toutes les orientations avec toutes les illuminations ne peut être réalisé. Pour surmonter ceci, il est donc possible d'utiliser des techniques d'apprentissage incrémentales. Cela permet d'apprendre sur une courte séquence et de reconnaître la personne avec des orientations ou des illuminations différentes dans d'autres séquences, en cherchant le modèle le plus proche et en

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	Global
Wholeset	92.16	99.44	100	95.77	99.10	99.37	100	100	100	100	99.78	99.31	98.98	97.34	100	96.87	98.23	91.89	97.42	99.22	98.21
PCA	91.07	99.72	98.74	95.37	99	92.15	93.65	99.76	100	100	99.23	96.11	94	94.42	91.69	97.74	98.11	85.70	85.77	100	95.6
CFS	92.59	99.44	100	97.99	99.10	99.79	100	100	100	100	99.78	98.28	99.39	97.72	100	98.62	98.23	92	96.56	100	98.46
Wrapper	91.83	99.44	99.58	91.15	99.10	92.25	99.08	99.88	100	100	96.92	95.65	86.98	93.03	94.16	95.11	97.99	83.88	93.37	99.67	95.39

TABLE 3.3 – Résultats de classification avec MidSVM, en considérant les ensembles CASIA-Wholeset, CASIA-PCA, CASIA-CFS et CASIA-Wrapper décrits précédemment

mettant à jour avec ces nouvelles données.
Cependant, un autre problème émerge alors, il s'agit de pouvoir, dans des endroits par exemple publics, identifier une personne qui est encore inconnue du système (l'apprentissage de nouveautés). Avec les techniques basées sur les SVM incrémentaux, cela est rendu très compliqué, car il faudrait trouver une condition sur les données qui soit assez spécifique pour permettre de détecter la différence entre une nouvelle classe et une évolution d'une classe connue et ainsi adapter l'ensemble du SVM avec une nouvelle classe à partir de données inconnues. Pour surmonter cet obstacle, nous avons choisi d'appliquer l'algorithme SAKM (Self-Adaptive Kernel Machine) (BOUKHAROUBA et al., 2009) sur les données images et de comparer avec les résultats des exécutions précédentes utilisant les SVM incrémentaux.
Ces travaux font suite, dans l'équipe de recherche, à plusieurs thèses de doctorat travaillant sur la classification dynamique d'un côté et sur la reconnaissance de personnes d'un autre côté et vise à améliorer les connaissances de l'état de l'art et la maîtrise de ces algorithmes de classification ou de clustering dynamiques en l'appliquant sur ces types de données.

3.2.7.2 Apprentissage par SVM Incrémental : MidSVM – ré-identification de personnes

La base de cet algorithme est d'entraîner un SVM avec une portion réduite de l'ensemble de données, de conserver uniquement les vecteurs supports et les vecteurs erreurs à chaque étape du processus d'apprentissage, qui créera un ensemble d'apprentissage pour l'étape suivante.
La classification par SVM est définie par :

$$x_i \in C_k; k = \underset{j=1,\ldots,K}{\arg\max} f_j(x_i)$$

où K est le nombre de classes. Les fonctions de décision sont $f_i(x) = w_i^T \Phi(x) + b_i$ (avec w_i et b_i définissant les coefficients de la séparatrice et Φ est la fonction noyau). Afin d'obtenir la meilleure marge entre les classes, la méthode des multiplieurs de Lagrange est appliquée et les fonctions de décisions sont réécrites avec les coefficients de Lagrange α_i. Ensuite, en se basant sur les conditions dites KKT en un point x_m, l'ensemble des données d'apprentissage que nous nommerons T est divisé en trois catégories en fonction des valeurs de g_m :

$$g_m = \begin{cases} > 0; & if \quad \alpha_m = 0; \quad \text{Ensemble des vecteurs de données} \quad DV \\ = 0; & if \quad 0 < \alpha_m < C; \quad \text{Ensemble des vecteurs supports} \quad SV \\ < 0; & if \quad \alpha_m = C; \quad \text{Ensemble des vecteurs erreurs} \quad EV \end{cases}$$

où $C \geq 0$ permet de paramétrer le nombre d'erreurs admissibles.
L'apprentissage incrémental par SVM repose sur le fait de préserver les conditions KKT sur tous les vecteurs de la base d'entraînement lorsqu'un nouveau vecteur est ajouté de manière adiabatique. La procédure incrémentale est décrite intégralement par (BOUKHAROUBA et al., 2009). Succinctement, elle consiste à déterminer à quel sous-ensemble la nouvelle donnée sera ajoutée, puis à mettre à jour la fonction de décision entre les classes si besoin.
Pour tester la méthode basée sur les MidSVM, 10% des images de la base de données sont réservés à l'apprentissage initial, et le reste est utilisé pour mettre à jour la base de connaissances. En effet, nouvel échantillon, le model actuel sert à déterminer sa classe puis il sera ajouté avec cette classe pour créer le modèle mis à jour (en considérant cette décision comme juste). Les résultats sont montrés dans le tableau 3.3. Les résultats sont très bons et très proches des 99.5% obtenus par les SVM avec un apprentissage plus classique (60% de la base de départ) et sans incrémentation.

Cependant, cette méthode, très performante pour l'adaptation des modèles de classe, ne permet pas de découvrir de nouvelles classes.

3.2.7.3 Self-Adaptive Kernel Machine (SAKM) – Identification et ré-identification de personnes

SAKM (pour Self-Adaptive Kernel Machine) (BOUBACAR et al., 2008) est un algorithme de clustering en ligne en mesure de recevoir des données non-stationnaires dans un contexte multiclasse. L'idée principale est de calculer le degré de proximité d'un nouvel échantillon avec les plus proches vecteurs supports de chacun des clusters dans l'espace de données, ceci avec une métrique induite par le noyau.
Lorsqu'une nouvelle donnée est présentée, la procédure à suivre (création, adaptation, fusion) est sélectionnée suivant un critère permettant de choisir l'action à réaliser. Ceci va prendre en compte d'une part des informations sur les clusters (leur taille, etc.) mais aussi les proximités et les évolutions des données présentées.
Le tableau 3.4 présente les résultats avec l'algorithme SAKM. Pour ces résultats, trois classes sont au départ considérées comme connues. Pour ces trois classes, 10% des données sont prises (en début de la première séquence) pour former une base d'apprentissage et apprendre trois clusters. Une fois cet apprentissage réalisé, le reste des éléments est utilisé pour l'apprentissage en ligne. Afin de tester la découverte de nouveautés, nous avons ainsi ajouté à cet ensemble de test quatre autres classes qui n'ont jamais été apprises. Le but étant que l'algorithme puisse détecter ces nouveautés afin de créer automatiquement une nouvelle classe. La dernière colonne du tableau nous indique le nombre total de clusters trouvés dans les données. Une étape de fusion est ensuite nécessaire pour rassembler les clusters appartenant à une même personne. À noter aussi que pour cette expérimentation nous avons choisi aléatoirement 7 individus dans l'ensemble des données de départ. Ainsi, P1 ne correspond pas forcément à C1 du premier tableau.

	P1	P2	P3	P4	P5	P6	P7	Nb Cluster
Wholeset	100	100	97.48	97.53	95.34	90	100	23
PCA	100	100	95.27	98.02	88.34	84.33	100	16
CFS	100	100	99.26	100	96.64	90.83	100	16
Wrapper	100	100	97.69	100	95.96	89.17	100	9

TABLE 3.4 – Résultats de classification avec SAKM et la découverte de nouveautés (3 classes apprises et 7 testées).

Les deux résultats sont très bons et se rapprochent de ceux de la classification non incrémentale bien qu'un nombre plus important de données d'apprentissage ait été utilisé. Il est à noter cependant que plus de clusters ont été créés qu'il n'y a de classes à identifier. L'algorithme sépare certaines personnes en plusieurs clusters être en mesure de rassembler ces données.
D'autres tests ont été lancés pour vérifier les performances de SAKM, en se concentrant cette fois-ci uniquement sur l'ensemble CFS. Ainsi, le tableau 3.5 montre les résultats de l'utilisation des classes $P1$, $P2$ et $P3$ pour l'apprentissage et le test avec un nombre différent de classes et à chaque fois un nombre croissant de nouvelles classes. La dernière ligne du tableau donne à chaque fois le taux de bonne classification globale. Pour ces nouveaux tests cela va de 98.10% à 78.84% au fur et à mesure que le nombre de classes nouvellement créées augmente.

CASIA-CFS	4 nouvelles classes	5 nouvelles classes	6 nouvelles classes	7 nouvelles classes	8 nouvelles classes
P1	100	100	99.80	99.80	99.71
P2	100	100	100	100	100
P3	99.26	93.59	59.41	56.89	57.31
P4	100	97.92	89.82	97.83	97.73
P5	96.64	88.34	88.34	88.34	88.34
P6	90.83	84.64	82.26	86.07	89.88
P7	100	96.83	96.83	96.83	96.83
P8	-	36.61	42.26	42.26	42.03
P9	-	-	84.16	78.47	77.33
P10	-	-	-	69.22	72.08
P11	-	-	-	-	46.02
Global	98.10	87.24	82.54	81.57	78.84

TABLE 3.5 – Classification en ligne avec SAKM. 5 Expérimentations sur l'ensemble CFS, avec trois classes en apprentissage et un nombre croissant de classes ajoutées en test.

Les tests précédents ont montré la capacité de l'algorithme à détecter des nouveautés. Le test suivant va éprouver les performances de l'algorithme SAKM dans la classification en ligne et incrémentale de vingt personnes avec un nombre de classes dans l'ensemble d'apprentissage croissant. Les résultats pour chacune des personnes de la base ($P1, \cdots, P20$) avec les attributs issus de CFS, sont donnés dans le tableau 3.6.
Dans chacun des tests, nous avons un certain nombre de classes dans la base d'apprentissage (un certain nombre de personnes), avec à chaque fois environ 10% (100 images) utilisés pour l'apprentissage et le reste pour le test. Dans tous les cas pour le test nous avons les 20 classes qui sont présentées. Pour le premier cas avec trois classes d'apprentissage seulement, le taux de bonne reconnaissance est de 61.15%. De plus, nous pouvons noter, toujours pour ce cas, que $P3$ est reconnue de manière non satisfaisante en considérant qu'elle était dans l'ensemble d'apprentissage. Cela s'explique par le fait que ces données ont été classées par l'algorithme d'apprentissage avec des données de la classe $P9$.
Pour les expérimentations suivantes, avec 5, 8 ou 10 classes d'apprentissage, les résultats augmentent avec le nombre de classes, comme nous pouvons logiquement nous y attendre. Il est logique aussi de noter qu'en général les classes connues dans la base d'apprentissage sont les mieux reconnues. Nous pouvons aussi noter que certaines classes représentant la nouveauté sont tout le temps mal reconnues ($P14$, $P15$ et $P19$). Dans chacun des tests, $P19$ est toujours confondue avec $P3$, $P9$ ou $P12$. Cela est logique, car les vêtements de ces personnes ont des similarités flagrantes en termes de couleur et/ou de texture. De la même manière, $P15$ est fréquemment confondue avec $P5$ et $P11$. Des exemples d'image de ces personnes sont montrés à la figure 3.3. Certaines personnes de la base sont mieux reconnues, car sont vraiment différentes en termes d'attributs extraits. Des exemples sont donnés à la figure 3.4. Enfin, le dernier test qui est présenté (dernière colonne) contient uniquement 18 classes de test (sans $P14$ et $P15$). Avec ce dernier changement, les résultats passent de 67.9% à 74.08%.

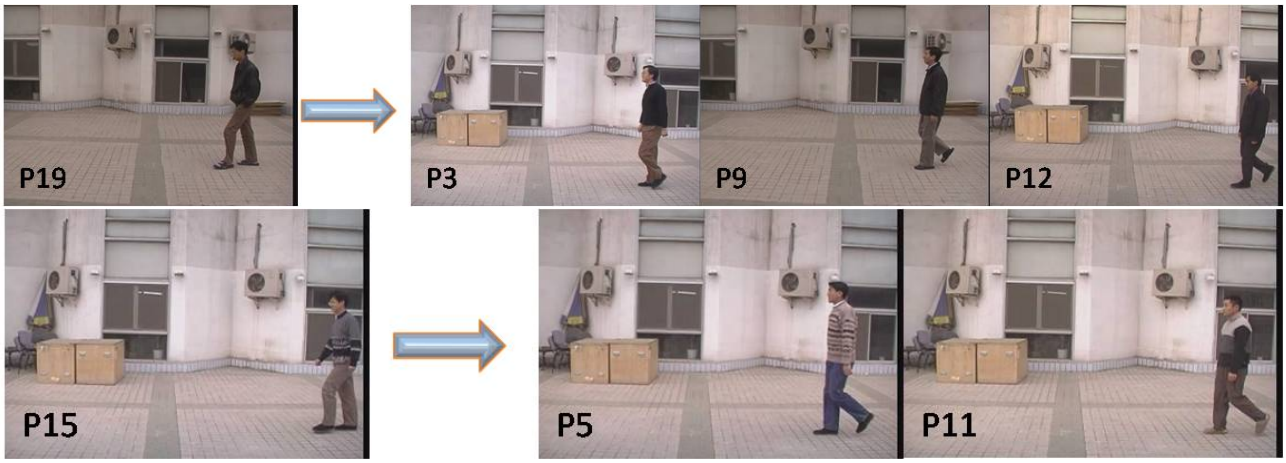


FIGURE 3.3 – Deux exemples ($P19$ et $P15$) de classes de tests qui sont difficilement reconnues.

FIGURE 3.4 – Exemples de classes avec un fort taux de reconnaissance.

3.2.8 Conclusions des travaux

La question qui se posait était de savoir si les algorithmes de classification en ligne seraient capables de généraliser des classes apprises avec des données différentes mais correspondant à la même personne, tout simplement en adaptant la forme de la classe. Les premiers résultats testés ont montré que oui, même si l'algorithme SAKM permettant ensuite

CASIA-CFS	3 classes d'entraînement	5 classes d'entraînement	8 classes d'entraînement	10 classes d'entraînement	10 classes d'entraînement
P1	99.71	99.80	99.80	99.8	99.80
P2	96.45	97.78	98.89	98.89	98.89
P3	62.57	98.42	96.53	89.27	90.85
P4	97.53	98.03	97.59	88.6	88.6
P5	87.78	91.04	98.23	98.23	98.23
P6	85	87.14	86.35	86.35	86.35
P7	64.38	66.38	82.15	81.09	83.22
P8	42.26	22.6	58.6	58.6	58.6
P9	86.96	77.23	37.27	44.57	45.2
P10	71.22	40.73	76	75.37	75.05
P11	44.53	33.9	64.91	75.94	74.65
P12	55.01	83.06	74.78	89.29	90.17
P13	98.78	98.53	100	97.8	96.21
P14	0	33.19	11.59	0	-
P15	42.11	25.03	12.19	0	-
P16	82.50	94.29	81.72	94.2	82.5
P17	0	46.57	43.44	52.59	58
P18	66.81	68.41	12.81	62.54	27.96
P19	0	21.60	44.22	22.41	10.34
P20	39.39	47.9	59.44	42.54	68.88
Global	61.15	66.58	66.83	67.9	74.08

TABLE 3.6 – Classification pour les 20 personnes avec l'ensemble CFS. L'apprentissage s'est fait avec 3, 5, 8 ou 10 personnes selon les cas.

d'apprendre des nouveautés donne des résultats qu'il faut réinterpréter car il créé plus de clusters que de classes.
Ces résultats ont montré l'intérêt des algorithmes incrémentaux pour d'une part l'adaptation à des données non stationnaires et/ou incomplètes (lorsque l'ensemble d'apprentissage ne représente pas assez bien l'ensemble des possibilités), mais aussi pour la découverte de nouveautés. Des résultats encourageants ont été trouvés sur des bancs de tests connus avec des attributs simples à extraire et compréhensibles. La découverte de nouveauté, elle, doit par contre encore être améliorée afin de donner des résultats plus fiables lorsque le nombre d'individus augmente.
La partie suivante va s'intéresser à la reconnaissance d'actions avec une volonté d'application dans le domaine de la robotique et notamment dans un but prochain d'assistance à la personne.

3.3 Reconnaissance d'actions

Dans les différents travaux, nous considérons qu'une action est définie comme un ensemble simple de gestes. Dans ces travaux, nous avons donc travaillé sur différentes actions telles que taper dans les mains, sauter, lever les bras, etc. Mes travaux de recherche portent beaucoup sur l'assistance à la personne. Un moyen de fournir cette assistance qui est exploré par d'autres personnes de l'équipe est de proposer un robot qui permettrait d'aider la personne à son domicile par exemple. Le but de ces travaux était de permettre, par exemple à un tel robot, de reconnaître certains gestes ou certaines attitudes d'une personne afin de pouvoir détecter une intentionnalité de rentrer en interaction avec celui-ci. Pour faire ceci, il faut lever deux problèmes. Le premier est que les actions sont bien reconnues (comme précédemment pour les personnes) lorsque l'orientation de la personne est proche de celle de départ. Or, pour un robot d'assistance qui peut être n'importe où, nous n'avons aucune idée de l'orientation qui sera faite. Nous cherchons donc également un algorithme qui sera indépendant de l'orientation. Le second est que la décision doit être prise le plus vite possible. En effet, le robot se devra d'être réactif et donc pour ceci il faut que le système le soit et que la prise de décision se fasse au plus vite, y compris peut être avant la fin de l'action. Après un court état de l'art sur les caméras RGB-D et leurs données, ces deux problématiques seront abordés par l'intermédiaire de deux travaux qui ont été réalisés consécutivement sur le sujet.

3.3.1 Les cameras RGB-D

En plus de capteurs 2D (caméras RGB classiques qui donnent une image en deux dimensions), nous nous sommes tournés également vers les capteurs 3D, RGB-D (nous donnant en plus une notion de profondeur – Depth) afin de permettre des analyses plus complètes et de diriger les travaux vers l'analyse de comportement et l'interaction avec une personne.

Les avancées dans le domaine de la vision par ordinateur ont été un atour majeur dans le développement des interactions homme-machine. En effet, de nombreux algorithmes permettant d'analyser, de reconnaître ou d'interpréter des actions humaines ont été proposés ces dernières années. La démocratisation des capteurs 3D a également aidé à ce développement. En effet, l'arrivée de la Kinect (au départ créée uniquement pour les jeux) et l'extension de son utilisation grâce à la disponibilité d'un SDK et à son prix raisonnable ont permis aux chercheurs d'élargir l'utilisation de cet outil (l'arrêt de commercialisation de la Kinect n'a fait que reporter l'utilisation sur les nombreux concurrents apparus depuis sa sortie). Les capteurs 3D sont maintenant en grand nombre, avec des prix et des « services » offerts différents (couleur vs noir et blanc, extraction du squelette en temps réel, reconnaissance faciale, détection des pulsations cardiaques, finesse de la détection ou suivi multi-personnes etc.).

L'état de l'art montre une utilisation croissante de ce type de capteurs (HAN et al., 2013) dans des domaines diversifiés tels que l'assistance ambiante et notamment destinée aux personnes âgées (SEVRIN, 2016), la reconnaissance d'actions (XIA, C.-C. CHEN et J. AGGARWAL, 2012), la reconnaissance d'activités (J. K. AGGARWAL et al., 2014), la réhabilitation (CHANG et al., 2011 ; CLARK et al., 2012), la détection de chute (STONE et al., 2015), etc.

En ce qui concerne la reconnaissance d'actions, celle-ci peut se faire avec deux signaux différents (et de volumétrie différente) à savoir les signaux de profondeur (signal RGB-D) ou l'extraction dans celui-ci du squelette de la personne (fait en temps réel par certain modèles).

Les premières approches se basent sur le signal complet RGB-D. Dans leurs travaux, XIA, C.-C. CHEN et J. K. AGGARWAL (2011) ont créé un modèle de graphes d'actions pour décrire les postures saillantes. OREIFEJ et al. (2013) ont proposé HON4D, un histogramme en quatre dimensions sur le temps, l'espace et la profondeur utilisés pour détecter les changements dans les surfaces normales à l'orientation. C. CHEN et al. (2016) ont extrait des attributs en utilisant ce qu'ils ont appelé les « Depth Motion Maps » (DMM). Pour chaque image en profondeur reçue, elle est projetée sur trois repères cartésiens orthogonaux. Les différences (absolues) entre deux images consécutives projetées sont cumulées ensuite sur la séquence complète.

Les approches suivantes tentent de se baser sur des données moins complètes, mais aussi moins lourdes à traiter. En effet, traiter une image en trois dimensions à des fréquences suffisantes est un processus très couteux en temps de calcul, mais aussi et surtout en bande passante. Ainsi, à partir des images de profondeur, il est possible d'identifier des corps humains et de reconstruire le squelette de celui-ci. C'est le processus qui est embarqué dans la Kinect et qui se base sur les travaux de SHOTTON et al. (2011). Le squelette est vu comme un ensemble de composantes rigides reliant les articulations du corps humain. Le nombre de segments (et donc d'articulations) dépend du capteur utilisé et donc de l'algorithme qui est intégré. Certains travaux ont ainsi montré l'utilisation de ces données pour la reconnaissance d'actions ou d'activités. CIPPITELLI et al. (2016) ont extrait des attributs sur la position du squelette en le normalisant, par distance Euclidienne, à partir de la distance entre le torse et le cou. Ensuite, un algorithme de type k-means groupe les différentes positions du squelette dans des clusters et un vecteur contenant les centroides de ces clusters est créé. Un SVM multiclasses va, à partir de ces données de position, classifier des actions humaines. DEVANNE et al. (2015) se basent eux sur des similarités dans la forme des trajectoires des articulations dans un espace Riemannien. La classification est faite avec les K plus proches voisins. MIRANDA et al. (2014) présentent une méthode de reconnaissance de gestes en temps réel. Les coordonnées sphériques des articulations sont utilisées pour représenter la posture de la personne. Un SVM multiclasses avec un noyau adapté permet d'identifier les positions les plus importantes et une forêt aléatoire reconnaît ensuite les gestes. Dans sa thèse effectuée en partie dans l'équipe, GHORBEL (2017) a également travaillé sur la reconnaissance d'actions de manière rapide avec des caméras RGB-D. Deux descripteurs sont alors proposés (l'un pour travailler hors-ligne et l'autre en-ligne), avec des méthodes de calculs optimales afin de réduire le temps d'exécution.

Les mémoires de Master de Mounir Hammouche et de Sid Ahmed Walid Talha ont porté sur la reconnaissance de gestes avec les données issues du squelette en créant deux nouveaux types d'attributs qui seront présentés dans la suite de cette section. Les résultats de classification sur une base acquise dans le département et sur des bases connues suivront et concluront ce chapitre.

3.3.2 Extraction d'informations dans les signaux vidéos RGB-D

3.3.2.1 Travail sur l'indépendance à l'orientation

Pour ce travail, aucune nouvelle méthode d'extraction de caractéristiques n'a été proposée. Différents attributs disponibles dans la littérature ont été réimplémentés et ont été testés. Il en ressort l'utilisation des deux types d'attributs qui semblaient les plus adaptés et les plus performants selon le type de signal considéré :

— Pour les données de profondeur/couleur, les descripteurs HOG^2 (OHN-BAR et al., 2013) ont été retenus. HOG (pour Histogram of Oriented Gradients) calcule un ensemble d'histogrammes sur une grille dense sur l'image (dans notre cas comme dans d'autres applications en se restreignant à la partie contenant la personne faisait l'action), en considérant le sens (l'orientation) de la variation (du gradient) afin d'analyser la direction des contours. L'idée de ces descripteurs est de donner de l'importance à cette direction. Dans HOG^2, les auteurs ont modifié la grille afin qu'elle se positionne autour des articulations du squelette, calculent les descripteurs HOG sur cette image, concatènent les résultats sur une séquence et recalculent les descripteurs HOG sur cette séquence temporelle.

— Pour les données du squelette, le descripteur RJP (pour Pairwise relative positions of the joints) (VEMULAPALLI et al., 2014). Comme son nom l'indique, dans cette représentation, la position de chacune des articulations va être soustraite à la position de toutes les autres articulations afin d'avoir la position relative entre chaque paire d'articulations différentes (par articulation nous entendons point sur le squelette).

Ces deux ensembles d'attributs ont été retenus pour les résultats qu'ils ont donné sur une première tâche de classification testée lors des premiers travaux présentés précédemment. Leur efficacité de calcul a également été prise en compte.
Sur ce premier travail, une plus longue description de la tâche de classification sera faite par la suite. La section suivante décrit l'introduction d'un nouveau descripteur sur les données du squelette.

3.3.2.2 Le descripteur Body-part Directional Velocity (BDV) : direction et vitesse

Pour ces seconds travaux, notre but était toujours le même, s'inscrire dans une logique de détection d'actions par un robot afin de pouvoir gérer les interactions avec celui-ci. Pour ce travail, nous nous sommes demandé s'il était possible de détecter assez rapidement une action, y compris avant sa fin.
Pour ceci, nous avons introduit un nouveau descripteur, nommé BDV pour Body-part Directional Velocity, qui va prendre en compte la direction et la vitesse de chacune des articulations. L'avantage de ce descripteur, contrairement à celui utilisé précédemment, est le fait de pouvoir se calculer en ligne et d'être causal (ie. ne nécessitant pas de connaissances du futur).
Une séquence de squelettes $\mathbf{p}$ représente une série de N poses ordonnées au cours du temps $\mathbf{p} = [\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_t, ..., \mathbf{p}_N]$. À chaque instant, le squelette $\mathbf{p}_t$ est composé de n positions d'articulations $\mathbf{p}_t = [\mathbf{p}_t^1, \mathbf{p}_t^2, ..., \mathbf{p}_t^i, ..., \mathbf{p}_t^n]$. Du fait des occlusions ou du bruit d'extraction du squelette, celui-ci n'est pas toujours parfait. Il est tout d'abord lissé par un filtre de Savitzky-Golay, appliqué sur chacun des points du squelette :

$$\mathbf{P}_t^i = \frac{1}{35}(-3\mathbf{p}_{t-2}^i + 12\mathbf{p}_{t-1}^i + 17\mathbf{p}_t^i + 12\mathbf{p}_{t+1}^i - 3\mathbf{p}_{t+2}^i) \tag{3.1}$$

Avec $\mathbf{P}_t^i$ la position du point i à l'instant t après filtrage. La vitesse du point i à un instant t est ensuite calculée comme $\mathbf{V}_t^i = \mathbf{P}_{t+1}^i - \mathbf{P}_{t-1}^i$.
Ensuite, nous proposons de diviser le squelette en cinq segments qui sont présentés sur la figure 3.5. Nous y trouvons le bras gauche (B_1), le bras droit (B_2), la jambe gauche (B_3), la jambe droite (B_4) et le tronc (B_5). L'ensemble de ces parties du corps est noté $B = \{B_1, B_2, B_3, B_4, B_5\}$
Pour chacune de ces parties du corps, nous prenons la somme des vitesses de chacun des points d'un segment considéré en prenant ceux ayant une somme négative et ceux ayant une somme positive séparément (notés $D_{B_l}^+$ et $D_{B_l}^-$). La séparation entre les sommes positives et les sommes négatives est faite pour éviter les problèmes de compensations de mouvements et pour indiquer la direction principale du mouvement. Ces deux sommes sont définies comme :

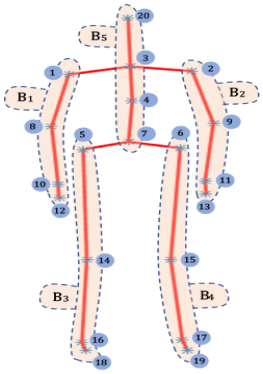


FIGURE 3.5 – Subdivision du corps humain en cinq parties (B_1, B_2, B_3, B_4 et B_5).

$$\mathbf{D}^{+}_{B_l}(t) = \sum_{i \in B_l} (\mathbf{V}^i_t \geq 0) \tag{3.2}$$

$$\mathbf{D}^{-}_{B_l}(t) = \sum_{i \in B_l} (\mathbf{V}^i_t \leq 0) \tag{3.3}$$

Le descripteur final est alors construit comme :

$$\mathbf{D}(t) = \bigcup_{l=1}^{5} [\mathbf{D}^{+}_{B_l}(t), \mathbf{D}^{-}_{B_l}(t)] \tag{3.4}$$

Ce descripteur proposé a montré une reconnaissance à la fois bonne et rapide sur deux ensembles de données habituellement utilisés en reconnaissance d'actions, comme cela sera décrit dans les sections suivantes avec les résultats de classification.

3.3.3 Classification indépendante de l'orientation

Comme énoncé précédemment, nous voulons une classification d'activités qui fonctionne sur un robot mobile, donc pour laquelle l'orientation n'est pas fixe au départ, et qui soit en plus rapide. Cette première partie va présenter des travaux permettant d'avoir une classification indépendante de l'orientation de la personne.

3.3.3.1 Approche proposée

Comme expliqué précédemment, le problème des approches actuelles en reconnaissance d'actions est une trop forte dépendance à l'orientation de la caméra par rapport à la position de la personne. Les approches utilisant les données de profondeur sont plus précises que les approches basées sur le squelette, mais les approches sur le squelette sont aussi plus indépendantes de cette position (tant que l'algorithme d'extraction du squelette est toujours valide). Afin de garder une bonne précision et d'avoir une invariance à l'orientation, les travaux de Mounir Hammouche ont porté sur une nouvelle méthode pour fusionner ces deux types d'approches de manière plus efficace. Pour tester ceci, un nouvel ensemble de données a également été créé, nommé Multiview3D [2] dans lequel des actions sont faites avec trois orientations différentes (−30, 0 et +30 degrés).

Dans ces travaux, nous avons proposé une méthode qui se base sur quatre classifieurs de base. Trois d'entre eux ont comme données d'entrée les données de profondeur (chacun avec l'une des orientations) et le dernier les données de squelette. Pour les données de profondeur, après une rapide comparaison des performances en termes de précision et de rapidité, nous avons finalement choisi HOG^2 (OHN-BAR et al., 2013) comme descripteur (voir la section 3.3.2.1 qui présentait ces sélections). Pour les données de squelette, comme expliqué dans le chapitre précédent, nous retenons RJP (VEMULAPALLI et al., 2014) (Relative Joint Position). Les sorties des différents classifieurs de base sont ensuite combinées pour améliorer la précision de la reconnaissance. Pour la fusion des informations en sortie des classifieurs,

2. Ensemble de données acquis et rendu publique à l'adresse http://ia.ur.mines-douai.fr/en/datasets/.

deux algorithmes sont proposés afin de réduire l'effet néfaste de la mauvaise orientation de la personne et de prendre en compte correctement les informations sur la précision de l'information du squelette. Ces deux algorithmes sont présentés dans la section suivante.

3.3.3.2 Fusion des informations

3.3.3.2.1 Depth-skeleton Fuzzy Switcher Algorithm (DSFSA)

Afin de créer un algorithme de reconnaissance d'actions précis même en présence de mauvaise reconnaissance du squelette ou de variations de l'orientation du corps humain, la fusion des informations du squelette et des informations de profondeur est faite avec un « fuzzy switcher » comme décrit par la figure 3.6. Dans notre algorithme, nommé « Depth-skeleton Fuzzy switcher Algorithm » (DSFSA), un ensemble de règles expertes choisissent automatiquement le meilleur classifieur de base. DFSDA utilise les informations de squelette quand tous les segments (ou du moins une majeure partie d'entre eux) sont relativement bien reconnus. Dans le cas contraire, le système se base sur les informations de profondeur. Comme montré dans la Figure 3.7, à partir des deux informations de profondeur et de squelette, nous avons une sortie finale pour la classification. Cette architecture va être détailler dans les paragraphes suivants.

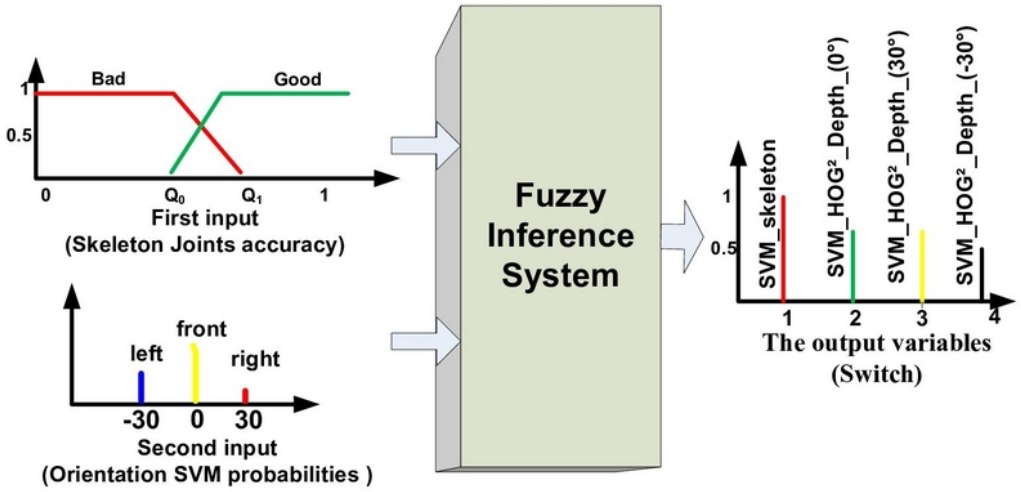


FIGURE 3.6 – Architecture du système logique flou. Les seuils Q_0 et Q_1 sont deux seuils flous déterminés sur les bases expérimentales et qui vont déterminer les fonctions floues utilisées

La relation entre les entrées et la sortie peut être décrite par les règles suivantes :

$$\begin{cases} \text{Skeleton_Accuracy} = \text{Good} \implies \text{SW} = \text{SVM}_1 \\ \text{Skeleton_Accuracy} = \text{Bad} \wedge \text{Orientation} = \text{Front} \implies \text{SW} = \text{SVM}_2 \\ \text{Skeleton_Accuracy} = \text{Bad} \wedge \text{Orientation} = \text{Right} \implies \text{SW} = \text{SVM}_3 \\ \text{Skeleton_Accuracy} = \text{Bad} \wedge \text{Orientation} = \text{Left} \implies \text{SW} = \text{SVM}_4 \end{cases}$$

Les règles floues sont basées sur la précision du squelette (une donnée de l'algorithme connue et interne de la Kinect) et sur l'orientation de la personne, donnée inconnue. La sortie du système d'inférence flou présenté est un vecteur de poids qui détermine lequel des quatre classifieurs semble le plus adéquat pour reconnaître au mieux l'action en cours. La précision de détection des segments du squelette est une donnée à deux dimensions, comme montré par la figure 3.6 et dont les paramètres (Q_0, Q_1) dont définis par un algorithme de Fuzzy C-Mean (FCM). FCM est une technique de clustering qui donne les centres des clusters et le degré d'appartenance de chaque point (DUNN, 1973; BEZDEK, 1981). Cette technique peut être utilisée pour construire des systèmes d'inférence et plus précisément un partitionnement flou en utilisant une fonction objectif d'optimisation itérative comme décrit par l'équation 3.5 :

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^m ||x_i - c_j||^2, \quad 1 \leq m < \infty \tag{3.5}$$

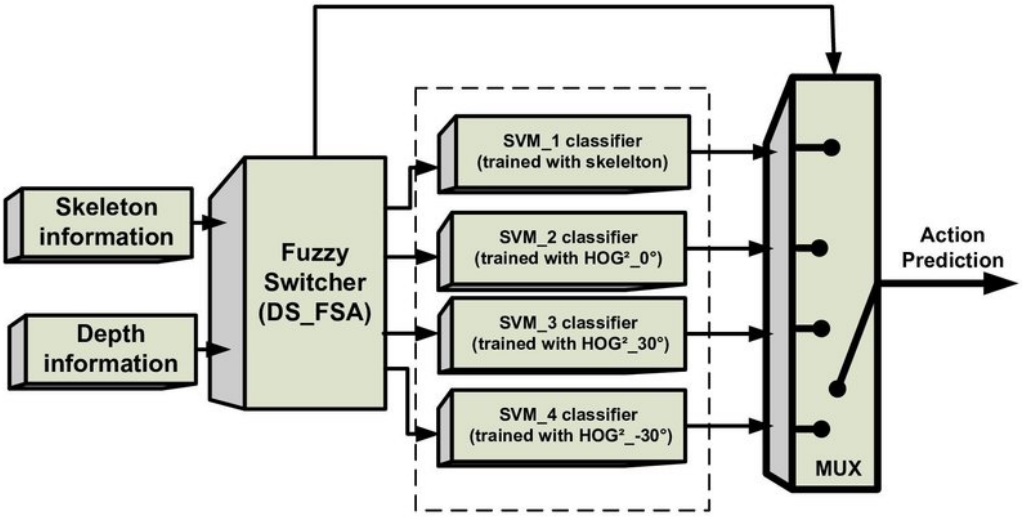


FIGURE 3.7 – Algorithme DFSDA

m est un nombre réel supérieur à 1, x_i est la i^{eme} mesure, c_j le centre du cluster et u_{ij} le degré d'appartenance de x_i au cluster j. Les centres des clusters c_j et le degré d'appartenance u_{ij} sont calculés par l'équation (3.6).

$$u_{ij}^m = \frac{1}{\sum_{k=1}^{C} (\frac{||x_i - c_j||}{||x_i - c_k||})^{\frac{2}{m-1}}}, \qquad c_j = \frac{\sum_{i=1}^{N} u_{ij}^m \cdot x_i}{\sum_{i=1}^{N} u_{ij}^m} \tag{3.6}$$

L'algorithme s'arrête lorsque la condition $max_{ij}|u_{ij}^{k+1} - u_{ij}^k| < \Im$ est satisfaite, avec k le nombre d'itérations et $\Im$ un seuil de convergence fixé empiriquement. Ce processus itératif converge vers un minimum local de J_m. Dans notre expérimentation, nous avons pris les valeurs suivantes : $\Im = 1e-5$, et $k = 100$. Ainsi, nous obtenons un couple (Q_0, Q_1)=(0.356, 0.716).

Plusieurs approches pour « rendre non floues » sont disponibles dans la littérature telles que la méthode des centroïdes, la méthode des moyennes pondérées, le principe du maximum d'appartenance, etc. (SINGHALA et al., 2014). Dans nos travaux, nous avons utilisé la méthode des moyennes pondérées, résumée par l'équation (3.7). Nous avons utilisé cette méthode du fait de la sortie de notre système d'inférence flou qui est un vecteur de pondération indiquant l'utilisation des différents classifieurs.

$$Defuz = \frac{\sum_{i=1}^{4} P[i] \cdot W[i]}{\sum_{i=1}^{4} W[i]} \tag{3.7}$$

Defuz est la sortie du système d'inférence flou, *P[i]* la valeur extrême de la i^{eme} sortie (fonction d'appartenance) et $W[i]$ le poids de la i^{eme} règle.

3.3.3.2.1.1 Estimation de l'orientation du corps Pour prédire l'action faite par une personne, utiliser le classifieur le plus adapté, car appris avec la même orientation (ou une orientation proche) de celle utilisée pour faire le mouvement actuellement donnera les meilleurs résultats. Ainsi, une bonne estimation de l'orientation du corps est nécessaire. L'estimation de l'orientation du corps humain est un défi du fait de la variabilité de la localisation par rapport à la caméra, mais aussi des variabilités anthropométriques et de la différence dans la vitesse d'exécution. Ces facteurs détériorent la précision de l'estimation. De plus, les approches récentes sont souvent basées sur des informations liées à la marche, en prenant en compte que l'orientation est en rapport avec la direction du mouvement (SHINMURA et al., 2015; W. LIU et al., 2013). Cependant, dans le contexte d'interaction avec un robot, seuls les membres supérieurs vont avoir une utilité pour effectuer l'action. Cela rend l'estimation de l'orientation non triviale.

Ici, nous utilisons les attributs « Shape Context » et « Scale-Invariant Feature Transform (SIFT) ». Nous utilisons les informations de profondeur en projetant ces données sur la séquence dans un repère cartésien orthogonal (vue de haut) (X. YANG et al., 2012 ; C. CHEN et al., 2016). La carte de profondeur des données de projection est calculée pour estimer l'orientation du corps (Figure 3.8). La carte de profondeur est calculée en sommant les différences en valeur absolue entre deux images de profondeur consécutives dans la séquence comme suit (Equation (3.8)) :

$$DM_{top} = \sum_{i=a}^{b} |map^{i}_{top} - map^{i-1}_{top}| \tag{3.8}$$

Avec i le numéro de l'image, map^{i}_{top} la projection vue de haut de la i^{eme} image et $(a,b) \in [2,n]^2$ le premier et le dernier indice de l'image dans la séquence.

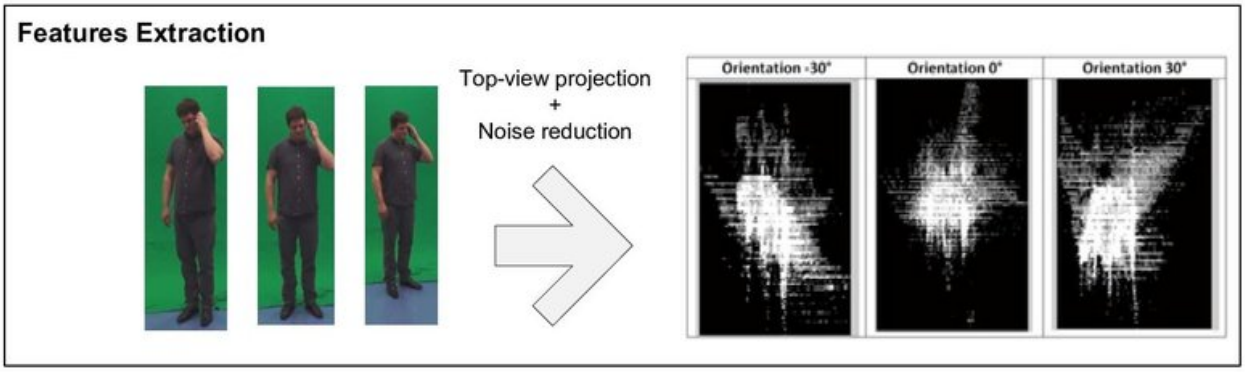


FIGURE 3.8 – Trois exemples de cartes de profondeur obtenues pour différents points de vue dans l'ensemble Multiview3D.

Un classifieur SVM est utilisé pour estimer l'orientation du corps. Cependant, pour l'entraîner et le tester, les attributs extraits des cartes de profondeur sont normalisés à partir de la connaissance de l'ensemble des exemples. Une ACP a été réalisée pour minimiser l'ensemble des attributs et ainsi le temps d'exécution de la tâche de classification. Cette approche a été testée en utilisant la moitié de l'ensemble des données pour l'entraînement et l'autre moitié pour le test et a permis un taux de reconnaissance de l'orientation de 86.1%.

3.3.3.2.2 Agrégation des résultats des différents classifieurs

Une dernière étape permet d'encore améliorer notre algorithme, se basant sur la sortie des différents classifieurs et l'agrégeant, comme le montre la figure 3.9. Nous conservons cependant la même stratégie que précédemment qui est de considérer les données issues du squelette ou les données issues de la profondeur en fonction de la précision donnée par l'algorithme de reconstruction du squelette. Un vote majoritaire (KIM et al., 2002) et un algorithme basé sur la notion d'investissement (Y. LI et al., 2015 ; PASTERNACK et al., 2010) sont utilisés pour fusionner les données issues des trois classifieurs basés sur la profondeur.
Ces deux méthodes ont été adaptées et testées dans les travaux de master de Mounir Hammouche.

3.3.4 Expérimentations

Afin de tester l'algorithme de décision flou et les deux méthodes précédemment proposées pour l'agrégation des décisions en termes d'une part de précision, mais aussi de temps d'exécution, des tests ont été conduits sur l'ensemble de données Multiview3D que nous avons créé. Cet ensemble de données contient douze actions différentes (remuer une main, boxer à deux mains, s'asseoir sur une chaise, remuer les deux mains, se ternir la tête, répondre au téléphone, ramasser un objet, frapper avec le pied, se ternir le dos, regarder sa montre, sauter et enfin lancer un objet par dessus sa tête), répétées deux fois par huit personnes différentes, avec trois orientations différentes par rapport à la Kinect (0°, 30°,

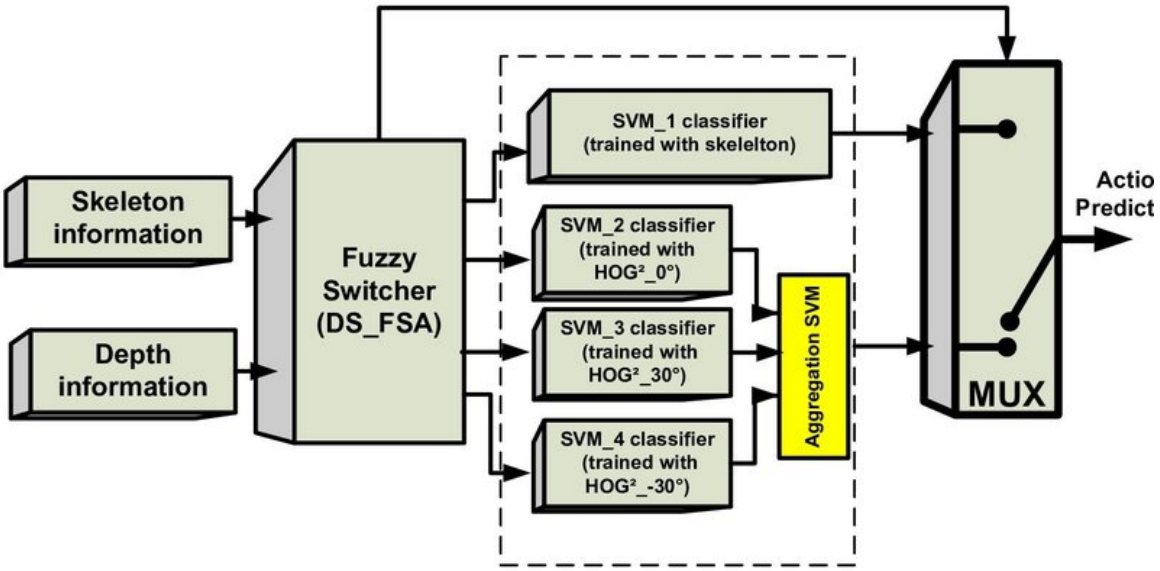


FIGURE 3.9 – Architecture de l'algorithme DSFDA auquel nous avons ajouté l'algorithme d'agrégation.

-30°). Cet ensemble de données a été créé pour ce défi de reconnaître correctement les actions avec des orientations aussi différentes les unes des autres. Dans nos expérimentations, chacun des SVMs avait un noyau linéaire.

En ce qui concerne la base d'entraînement et celle d'évaluation, quatre sujets étaient pris au hasard pour composer la base d'apprentissage, et les quatre autres étaient dans la base de test. Le tableau 3.7 résume les résultats obtenus par les différentes approches proposées en termes de précision dans la reconnaissance.

Algorithme	Résultats de classification
DMM (C. CHEN et al., 2016)	69.25
HON4D (OREIFEJ et al., 2013)	82.92
HOG2 (OHN-BAR et al., 2013)	81
RJP (VEMULAPALLI et al., 2014)	95.20
SVM_1_Skelton_RJT (1/2 skeletons noisy 5dB)	**79.17**
SVM_2_HOG^2_0°	**89.58**
SVM_3_HOG^2_30°	**83.63**
SVM_4_HOG^2_-30°	**79.66**
Fuzzy + **viewpoint estimation** (Depth+Skelton)	**95.83**
Fuzzy + **Aggregation (Investment)**	**97.91**

TABLE 3.7 – Comparaison entre les différentes approches

Les résultats obtenus montrent que la combinaison des informations de profondeur et de squelette améliorent considérablement les performances en termes de précision. De plus, nous remarquons clairement que l'agrégation basée sur l'investissement fournit la meilleur précision moyenne avec 97.91%, en comparaison avec l'état de l'art et les autres méthodes proposées. En ce qui concerne le temps d'exécution, nous constatons que l'ensemble du processus, de l'extraction de caractéristiques jusqu'à la classification, atteint un maximum de 3,5 secondes par séquence pour une moyenne de 35 images. Ces résultats montrent que l'approche proposée est applicable en temps réel. De plus, nous pouvons conclure que la combinaison DSFSA et agrégation par investissement dépasse les autres approches en termes de performances tant en précision qu'en temps.

Ces résultats encourageants montrent la faisabilité de classification efficace des actions même lorsque la position de la personne par rapport à la caméra peut changer. Comme expliqué en introduction, par rapport à l'application spécifique que nous visons (robotique mobile d'assistance à la personne), le système doit pouvoir reconnaître l'action dans toutes les conditions (donc avec possiblement une orientation inconnue) mais aussi être capable de prendre une décision rapidement afin d'être le plus réactif possible. Ceci sera la base des seconds travaux décrits dans la section suivante et qui s'intéressent à une classification anticipée de l'action en cours.

Algorithme	Temps d'exécution (sec.)
Classifieur basé sur le squelette (RJP)	1,357 (par séquence)
Classifieur basé sur la profondeur (HOG^2)	3,083 (par séquence)
Estimation de la précision du squelette	0,006
Estimation de l'orientation du squelette (10 images)	0,290
Système flou	0,0181
Agrégation (investissement)	0,1312
Agrégation + vote	0,1067

TABLE 3.8 – Performances temporelles

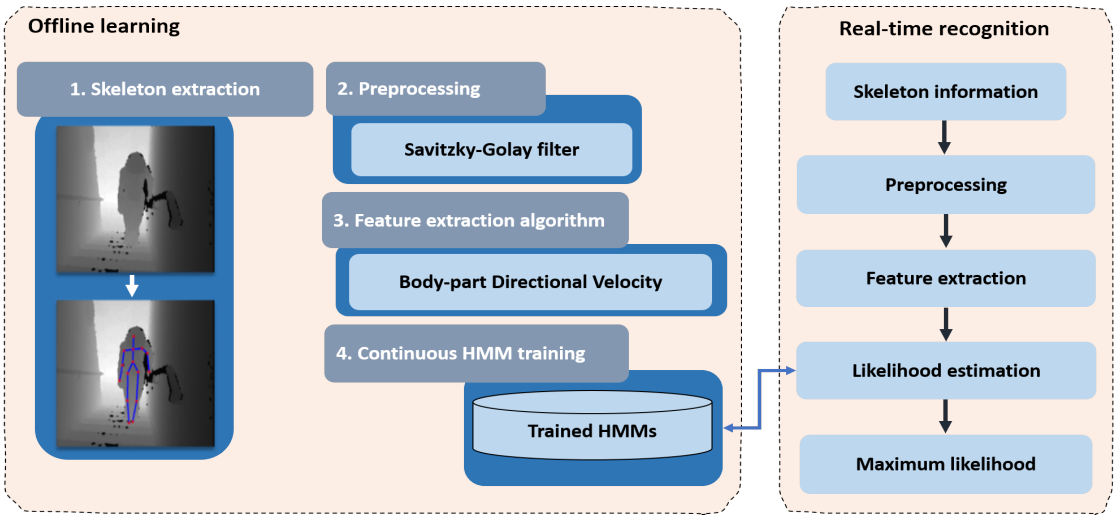


FIGURE 3.10 – Vision générale de l'architecture de reconnaissance proposée

3.3.5 Classification au fil de l'eau et anticipée

Les résultats présentés précédemment ainsi que les différents travaux de l'état de l'art ont montré que les capteurs RGB-D (pour les trois couleurs et la profondeur) pouvaient être efficaces pour reconnaître des actions. Nous avons également montré qu'il était possible de classifier efficacement des actions même lorsque la personne n'était pas précisément en face de la caméra.
Dans des applications qui se veulent de fonctionner en temps réel, il est important de détecter au plus vite et au mieux l'action en cours. Ainsi, il peut être pénalisant d'analyser des séquences complètes avant de pouvoir classifier (comme le requièrent les travaux précédents). Dans les travaux faits avec Sid Ahmed Walid Talha en Master, nous avons étudié cette piste afin de permettre de classifier plus vite et au fil de l'eau les différentes actions.
Pour cela, il était important de travailler sur un descripteur original, calculable sur chacune des images au fur et à mesure de leur arrivée. Ce descripteur a été présenté précédemment, dans la section 3.3.2.2.

3.3.5.1 Système proposé : classification avec des HMMs

La figure 3.10 présente l'architecture générale du système proposé pour reconnaître les actions. Cette architecture générale comprend deux blocs. Le premier, apprentissage hors-ligne, permet d'extraire les informations du squelette, de faire un pré-traitement pour diminuer le bruit présent dans ces données, d'extraire le descripteur Body-part Directional Velocity puis d'entraîner des classifieurs basés sur des HMMs pour chacune des actions présente dans l'ensemble de données. Une fois ces modèles construits, la partie reconnaissance en temps-réel va se charger d'analyser le flux des données squelette arrivant de la Kinect.
Comme expliqué dans la partie concernant l'extraction d'information de ce chapitre, nous calculons un descripteur que nous avons appelé « Body-part Directional Velocity » et qui est adapté à une classification en ligne, car ne nécessitant que la connaissance de l'image en cours et de l'image et de la position précédente pour fonctionner. Par la suite, des Modèles de Markov Cachés (HMMs) avec des distributions des états de sortie modélisés par des Mixtures de Gaussiennes (GMMs) sont utilisés pour la classification (illustré par la figure 3.11).

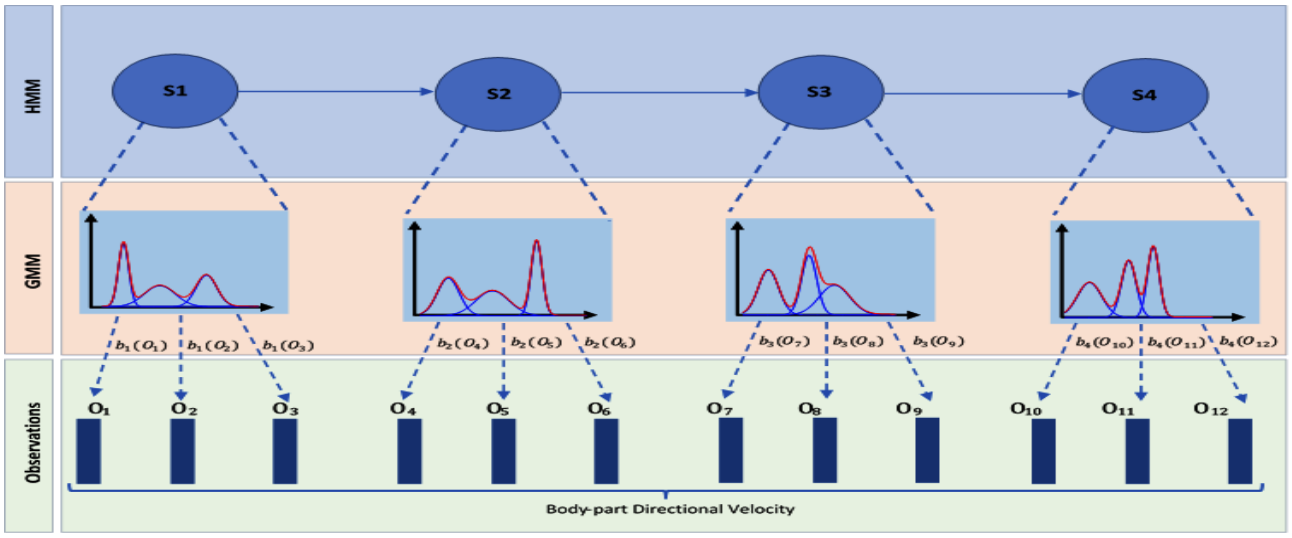


FIGURE 3.11 – Illustration du processus de classification mis en place

Pour rappel, un modèle de Markov Caché (RABINER et JUANG, 1986) est un modèle statistique utilisé pour décrire l'évolution de variables observables. Ces modèles sont principalement mis en œuvre pour travailler sur des séquences de données temporelles telles que la parole, le geste ou la reconnaissance d'activité. Les HMM sont basés sur deux processus stochastiques. Le premier est le processus des données observables, qui représente les mesures acquises. Le second est un processus modélisant ce qui n'est pas observable (caché) et qui peut être indirectement déduit en analysant la suite des données observées. Dans ce travail, nous avons choisi d'apprendre un HMM pour chaque action a à reconnaître.
La décision sera prise en utilisant le maximum de vraisemblance, celle-ci étant calculée grâce à l'algorithme Forward.

3.3.5.2 Résultats expérimentaux

Les résultats de ces travaux vont être présentés en deux parties. Après avoir rapidement présenté les deux ensembles de données utilisés (des bancs de tests de références), les premiers résultats vont comparer la méthode de classification (et notamment le nouveau descripteur) avec l'état de l'art, avant de présenter des résultats sur la « vitesse » (au sens précocité) de classification de l'algorithme.

3.3.5.2.1 Bancs de tests utilisés

Le premier ensemble de données utilisé, MSRAction3D, est le plus utilisé en reconnaissance d'actions. Il contient les données de profondeur et les données de squelettes et a été collecté par Microsoft Research. Il contient 20 actions différentes (dont le nom est laissé en anglais par clarté et cohérence avec la littérature) : *high arm wave, horizontal arm wave, hammer, hand catch, forward punch, high throw, draw x, draw tick, draw circle, hand clap, two hand wave, side-boxing, bend, forward kick, side kick, jogging, tennis swing, tennis serve, golf swing, pick up and throw*. Chacune de ces actions est faite par 10 personnes et répétée deux ou trois fois. Au total, l'ensemble contient 567 séquences. Comme discuté par W. LI et al. (2010) et afin de réalisé une comparaison avec l'état de l'art qui ait du sens, la subdivision en trois sous ensembles de données a été conservée : AS1, AS2, AS3 comme montré par le tableau 3.9. L'apprentissage et le test sur les différents sous-ensembles sont faits séparément, les résultats sont ensuite additionnés pour avoir des données sur la base complète. Comme fait par W. LI et al. (2010), la moitié des sujets sont retenus dans une partie apprentissage et l'autre moitié dans la partie test, faisant ainsi une validation croisée relative au sujet.
Le second ensemble de données, Florence3D, a été, comme son nom l'indique, créé par l'université de Florence. Il contient également les séquences avec les données de profondeur ainsi que les données de squelettes. 9 actions différentes sont proposées : *arm wave, drink from a bottle, answer phone, clap, tight lace, sit down, stand up, read watch, bow*. Chaque action est réalisée 2 ou 3 fois par 10 sujets pour un total de 215 séquences. Nous avons, dans ce cas, suivi le protocole expérimental de SEIDENARI et al. (2013) avec une validation croisée par sujets retenant à chaque tour un

AS1	AS2	AS3
[a02] Horizontal arm wave	[a01] High arm wave	[a06] High throw
[a03] Hammer	[a04] Hand Catch	[a14] Forward kick
[a05] Forward punch	[a07] Draw X	[a15] Side kick
[a06] High throw	[a08] Draw tick	[a16] Jogging
[a10] Hand clap	[a09] Draw circle	[a17] Tennis swing
[a13] Bend	[a11] Two-hand wave	[a18] Tennis serve
[a18] Tennis serve	[a12] Side boxing	[a19] Golf swing
[a20] Pick up and throw	[a14] Forward kick	[a20] Pick up and throw

TABLE 3.9 – Subdivision de MSRAction3D en trois sous ensembles AS1, AS2 et AS3

Algorithme	AS1 (%)	AS2 (%)	AS3 (%)	Global (%)
W. LI et al. (2010)	72,90	71,90	79,20	74,70
VENKATARAMAN et al. (2013)	77,50	63,10	87,00	75,90
C. CHEN et al. (2016)	96,20	83,20	92,00	90,50
MIRANDA et al. (2014)	96,00	57,10	97,30	83,50
CHAARAOUI et al. (2014)	91,59	90,83	97,28	93,23
VEMULAPALLI et al. (2014)	95,29	83,87	98,22	92,46
DU et al. (2015)	93,33	94,64	95,50	94,49
CIPPITELLI et al. (2016)	79,50	71,90	92,30	81,50
Z. LIU et al. (2016)	86,79	76,11	89,29	84,07
Notre algorithme	91,40	91,07	96,23	**92,90**

TABLE 3.10 – Résultats (précision) de notre algorithme sur MSRAction3D en comparaison à différentes méthodes de la littérature

sujet pour le test et le reste pour l'entraînement.

3.3.5.2.2 Reconnaissance d'actions humaines avec les HMMs et le descripteur BDV

Nous avons testé notre algorithme suivant le protocole décrit ci-dessus sur les deux ensembles de données. Les résultats sur le premier ensemble, MSRAction3D, sont donnés dans le tableau 3.10. Dans ce tableau, la précision de reconnaissance est comparée entre notre algorithme et différentes méthodes de la littérature, en utilisant bien sûr le même protocole de test. Les résultats montrent que notre méthode permet d'atteindre 92,9% de bonne classification, allant au-delà de quasiment tous les résultats de la littérature sur le sujet. La figure 3.12 présente en plus la matrice de confusion dans le cas des trois sous-ensembles de données. Nous y voyons que la plupart des actions sont bien reconnues. Des confusions « plus importantes » (mais restant raisonnables) se retrouvent sur des actions très similaires. Par exemple, dans le sous-ensemble AS2, « Hand Catch » n'est bien reconnue que dans la moitié des cas. On peut noter pour cette action des confusions, dans une séquence sur quatre, avec les actions « draw X » et « High arm wave ». Pour expliquer ceci, nous pouvons remarquer que l'action « Hand Catch » est caractérisée principalement lors de la partie « catching ». Comme dans ce cas nous ne travaillons qu'avec les données de squelette, ce moment est mal caractérisé, car le squelette n'a qu'un segment pour représenter chacune des mains. Ainsi, cette information manquante donne des résultats moins bons sur cette action.

Pour continuer cette validation, nous l'avons également lancée sur un second ensemble de données, Florence3D, sur lequel nous avons également comparé nos résultats avec l'état de l'art. Le tableau 3.11 résume cette analyse.

Algorithme	Précision (%)
SEIDENARI et al. (2013)	82,00
ANIRUDH et al. (2015)	89,67
DEVANNE et al. (2015)	87,04
CIPPITELLI et al. (2016)	76,10
VEMULAPALLI et al. (2014)	90,88
Notre algorithme	**90,32**

TABLE 3.11 – Résultats de différents algorithmes sur la base Florence3D

	a02	a03	a05	a06	a10	a13	a18	a20
a02	.92	0	0	0	0	0	0	.08
a03	0	1	0	0	0	0	0	0
a05	0	.18	.73	0	0	0	.09	0
a06	0	.18	0	.82	0	0	0	0
a10	0	0	0	0	1	0	0	0
a13	0	0	0	0	0	1	0	0
a18	0	0	0	0	0	0	1	0
a20	0	0	0	0	0	.22	0	.78

	a01	a04	a07	a08	a09	a11	a12	a14
a01	.76	0	0	0	.08	.08	.08	0
a04	.25	.5	.25	0	0	0	0	0
a07	0	0	1	0	0	0	0	0
a08	0	0	0	.93	0	0	.07	0
a09	0	0	0	0	1	0	0	0
a11	0	0	0	0	0	1	0	0
a12	0	0	0	0	0	0	1	0
a14	0	0	0	0	0	0	0	1

	a06	a14	a15	a16	a17	a18	a19	a20
a06	.73	0	0	0	0	.27	0	0
a14	0	1	0	0	0	0	0	0
a15	0	.09	.91	0	0	0	0	0
a16	0	0	0	1	0	0	0	0
a17	0	0	0	0	1	0	0	0
a18	0	0	0	0	0	1	0	0
a19	0	0	0	0	0	0	1	0
a20	0	0	0	0	0	0	0	1

FIGURE 3.12 – Matrices de confusion sur la base MSRAction3D

	Arm wave	Drink	Phone call	Clap	Tight lace	Sit down	Stand up	Read watch	Bow
Arm wave	.92	.04	0	0	0	0	0	.04	0
Drink	.04	.77	.15	0	0	0	0	.04	0
Phone call	0	.18	.73	0	0	0	0	.09	0
Clap	0	0	0	.94	0	0	0	.06	0
Tight lace	0	0	0	0	1	0	0	0	0
Sit down	0	0	0	0	0	1	0	0	0
Stand up	0	0	0	0	0	0	1	0	0
Read watch	.09	.04	.04	.13	0	0	0	.70	0
Bow	0	0	0	0	0	0	0	0	1

FIGURE 3.13 – Matrice de confusion de notre algorithme sur la base Florence3D

Comme précédemment, nous notons que notre algorithme donne quasiment les meilleurs résultats ((VEMULAPALLI et al., 2014) reste premier avec une légère avance sur notre algorithme). La matrice de confusion est donnée par la figure 3.13. Nous y voyons que les classes (5, 6, 7, 9) sont parfaitement reconnues alors que les classes 2 et 3 (*Drink* contre *Phone Call*), incluant toutes deux l'utilisation d'un objet, sont confondues plus fréquemment.

3.3.5.2.3 Vers une reconnaissance anticipée d'actions

Les résultats précédents ont montré l'efficacité du nouvel attribut proposé, le descripteur BDV. Afin que nous puissions reconnaître les activités au fil de l'eau et le plus tôt possible, le descripteur est généré image par image en calculant les positions et angles de chaque segments du squelette et leurs évolutions. La classification se fait ensuite dans la foulée. Pour valider cette approche d'un point de vue reconnaissance précoce d'actions, nous proposons donc un nouveau protocole expérimental sur les données que nous avions précédemment.

La figure 3.14 montre la valeur de la vraisemblance pour chaque HMM au fur et à mesure de l'exécution de l'action dans le cas de l'action « Side kick ». Dans ce cas, nous remarquons que, à partir de 20% de la séquence totale, la vraisemblance de la bonne action est supérieure à celle des autres HMMs et que ceci ne change pas d'ici à la fin de l'action. Dans cet exemple, nous sommes donc capables de reconnaître très rapidement l'action en cours. Nous pouvons également

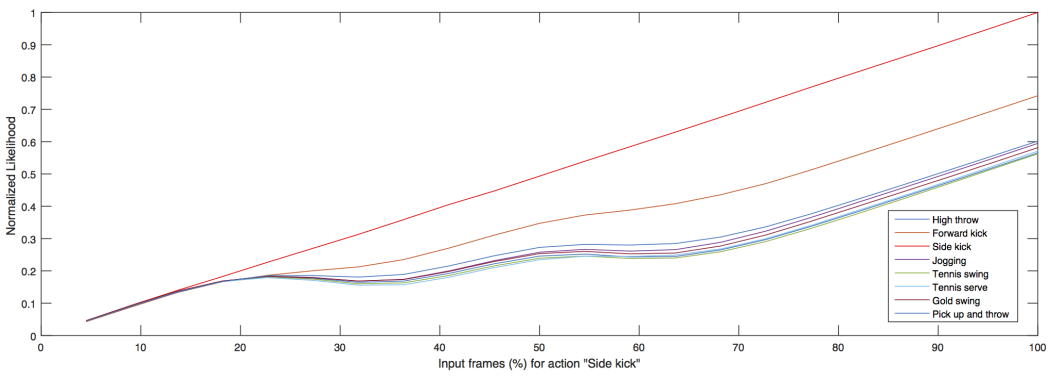


FIGURE 3.14 – Évolution de la vraisemblance dans le cas d'une action « Side kick »

noter que dans ce cas, la seconde vraisemblance en termes de valeur est celle de « Forward kick », ce qui s'explique simplement par le fait que les deux actions font entrer en jeu un mouvement important de la jambe.
Afin d'avoir ces résultats de manière globale, le même protocole a été suivi pour tout l'ensemble de données MSRAction3D, à savoir noter la dernière image pour laquelle la décision change. Ainsi, nous avons le pourcentage de temps nécessaire pour reconnaître chacune des activités. Un ensemble de boîtes à moustaches représente, sur la figure 3.15, ce pourcentage pour chacun des ensembles de données AS1, AS2 et AS3. Une première remarque générale sur ces trois ensembles est que les valeurs sont différentes en moyenne en fonction de l'ensemble de données (et donc logiquement des actions qui les composent). Pour AS1, les valeurs moyennes pour chaque action sont $Moy \in [4\%, 52\%]$, pour AS2 $Moy \in [15\%, 42\%]$, et enfin pour AS3 $Moy \in [4\%, 41\%]$. En regardant les médianes pour chacun de ces ensembles, nous voyons que la plus forte médiane pour AS1 est 52%, alors qu'elle est à 42% pour AS2 et 41% pour AS3. Ces valeurs maximales correspondent aux actions « Hammer », « Draw circle » et « Tennis swing ».

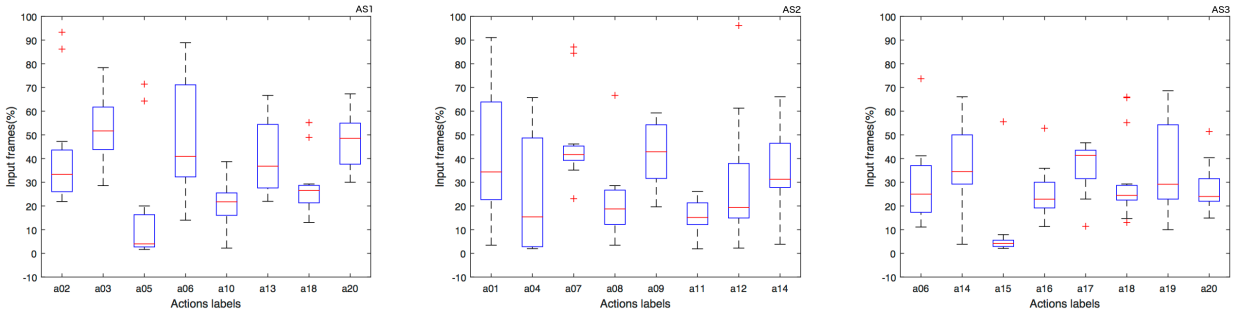


FIGURE 3.15 – Boîtes à moustache indiquant la vitesse de convergence de la décision (pour obtenir une classification stable) sur les trois sous-ensembles de MSRAction3D

Nous voyons donc avec ces résultats que nous sommes capables de converger assez rapidement vers une solution en grande majorité et que, en prenant une petite sécurité sur la plus grande moyenne, nous pouvons avoir une reconnaissance rapide de l'action pour améliorer l'expérience utilisateur dans nos applications.

3.3.6 Conclusion de ces travaux

Les travaux des deux étudiants présentés ci-dessus et publiés dans des conférences internationales et dans un journal ont permis de montrer comment efficacement reconnaître des actions courantes en relâchant quelques contraintes autour de la position, de la prise de vue et de la durée de la séquence. Les travaux ont porté sur une reconnaissance invariante à l'orientation de l'action faite et sur une reconnaissance précoce. Les attributs extraits (dont l'un a été créé pour cette application) et les classifieurs utilisés (que ce soient des SVM ou des HMM) ont montré des performances très bonnes (de l'ordre de plus de 90% de bonne classification) et surpassant la littérature actuelle sur la reconnaissance d'activités sur des bases connues et publiques. Il a été aussi montré, sur les mêmes bases publiques, que seulement les 40% de début de l'action sont nécessaires à une reconnaissance, ceci permettant ainsi de réagir au plus vite dans le cas d'un système en interaction avec des humains.
Avec tout ceci, nous avons posé les bases nécessaires pour la partie interaction que nous voulons mettre en place par la suite, en levant les deux problèmes posés à savoir reconnaître rapidement et aussi bien que lorsque l'orientation ne change pas des actions sur une caméra RGB-D, et la reconnaître au plus vite afin d'avoir une utilisation fluide du système embarquant cette reconnaissance.

3.4 Conclusions

Ce chapitre s'est intéressé à la reconnaissance de la personne et de ses actions. Pour atteindre cet objectif, les travaux ont porté sur d'une part l'extraction d'information (trouver les attributs les plus pertinents pour les deux problèmes posés) mais aussi sur les résultats de classification (avec des algorithmes séquentiels ou avec des algorithmes incrémentaux selon les besoins).

Chaque travail a montré de bons résultats de classification dans l'application concernée. Cependant ces travaux doivent tous être prolongés pour pouvoir améliorer leur utilisabilité dans des conditions réalistes. Pour les travaux concernant la reconnaissance de personnes, nos modèles à base de classification incrémentale ont montré de bonnes performances qui se rapprochent de celles des algorithmes statiques pour lesquels la répartition de données de départ est plus étendue. L'utilisation ensuite de SAKM a montré également que nous étions capables d'apprendre l'apparition de nouvelles personnes non connues au départ et garder de bons résultats. Cependant, cela se fait au détriment de la simplicité de lecture des résultats car une même personne va être divisée souvent en plusieurs clusters. Une thématique qui reste par contre inexplorée est de savoir quel serait le comportement « sur des grands nombres » de ces algorithmes, et notamment avec un nombre très important de personnes. Cela est une piste à explorer par la suite.
En ce qui concerne les travaux sur la reconnaissance d'actions, comme montré dans les sections précédentes, nous avons réussi à extraire des caractéristiques intéressantes de notre signal RGB-D pour obtenir une classification d'actions efficace et rapide. De plus, une décision anticipée est même possible, améliorant la réactivité des systèmes sur lesquels ceci pourrait être présenté. Un travail pour le futur pour cette thématique est de combiner les deux modèles réalisés, à savoir un multi-classifieur pour l'indépendance à la position de la personne et une reconnaissance en ligne, afin de proposer un système réel pouvant reconnaître au mieux comment interagir avec la personne et ensuite de tester tout ceci en condition réelle sur un robot d'assistance. Cela fera également une collaboration supplémentaire dans l'équipe de recherche, en y associant les chercheurs en génie logiciel qui travaillent sur les architectures modulaires pour la robotique mobile.
Dans le chapitre suivant, nous allons nous intéresser au problème de la reconnaissance d'activités et la création de profils utilisateurs dans les habitats intelligents. Nous distinguerons les actions des activités et nous décrirons la manière dont sont extraites les données et dont est effectuée la classification.

CHAPITRE 4

La personne et ses activités : analyse/reconnaissance dans les habitats intelligents et création de profils

Sommaire

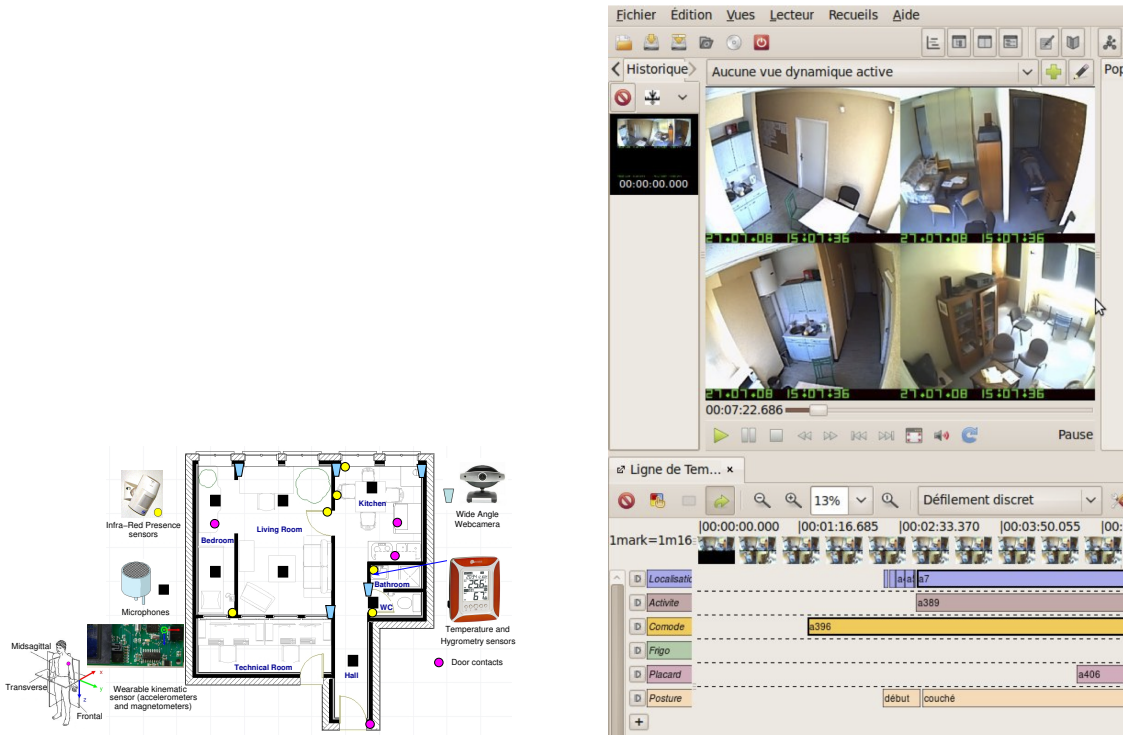
3.1 Introduction et verrous scientifiques à lever . . . 35
3.2 Identification/ré-identification de personnes dans des réseaux de caméras : de l'extraction d'information à la classification en ligne . . . 36
3.2.1 Acquisition et traitement de données vidéos . . . 36
3.2.2 But et originalité des travaux . . . 37
3.2.3 Base de données de test . . . 38
3.2.4 Extraction des caractéristiques . . . 38
3.2.5 Sélection et résultats . . . 39
3.2.6 Discussion sur les attributs extraits . . . 40
3.2.7 Techniques d'apprentissage pour l'identification et la ré-identification de personnes . . . 41
3.2.8 Conclusions des travaux . . . 44
3.3 Reconnaissance d'actions . . . 45
3.3.1 Les cameras RGB-D . . . 45
3.3.2 Extraction d'informations dans les signaux vidéos RGB-D . . . 47
3.3.3 Classification indépendante de l'orientation . . . 48
3.3.4 Expérimentations . . . 51
3.3.5 Classification au fil de l'eau et anticipée . . . 53
3.3.6 Conclusion de ces travaux . . . 57
3.4 Conclusions . . . 57

4.1 Introduction et verrous scientifiques à lever

Nous avons introduit, dans le chapitre précédent, la reconnaissance d'actions à partir de données issues de caméras dites 4D (pour laquelle nous avons le plan, la profondeur et le temps). Dans le chapitre précédent, comme dans l'ensemble des travaux que j'ai effectué, ce que nous définissons comme action est un ensemble de gestes incluant plusieurs membres de la personne et pouvant être plus ou moins complexe. Ces actions peuvent être reconnues comme précédemment montré par des caméras (2D + t ou 3D + t) ou par des capteurs inertiels.
Dans ce dernier chapitre, nous nous intéressons à des installations de capteurs plus complexes présents dans les habitats intelligents (Fig. 4.1). Dans un habitat dit intelligent, un ensemble de capteurs et d'actionneurs a été intégré dès sa construction (ou en rénovation) afin de :

— Mesurer un ensemble de paramètres liés aux conditions environnementales dans différentes pièces (température, hygrométrie, intensité lumineuse).

— D'évaluer l'utilisation de certaines fonctionnalités du logement (capteur de débit d'eau, compteur électrique équipé donnant la consommation instantanée, indicateurs donnant l'état branché ou débranché d'une prise ou donnant sa consommation en temps réel, indicateurs d'ouverture/fermeture de fenêtres ou volets, etc.).



(a) (b)

FIGURE 4.1 – Exemple d'un habitat intelligent (Laboratoire TIMC-IMAG, Grenoble) avec la vue des caméras.

— D'agir automatiquement sur des éléments de l'habitat (ouverture ou fermeture des volets, allumage ou extinction des lumières, contrôles d'éléments importants comme des plaques gaz, contrôle des différentes prises de la maison, etc.).

À partir de ces différentes informations, il nous faut extraire des vecteurs de caractéristiques (rappelons que ces données ne sont pas homogènes et n'ont bien sûr pas la même fréquence d'échantillonnage (lorsqu'ils en ont une, souvent ils se déclenchent sur évènement). Avec ces vecteurs de caractéristique, nous voulons ici reconnaître des activités. Une activité est définie comme une suite complexe d'actions menant à la réalisation d'une tâche précise de la vie quotidienne (faire un repas, prendre un repas, se laver, s'habiller, etc.). L'intérêt de cette reconnaissance d'activités est de pouvoir caractériser la journée d'une personne, et surtout de caractériser l'évolution de ces journées au cours du temps. Ces deux caractérisations seront utiles ensuite pour tenter d'évaluer le bien-être ou l'autonomie de la personne.

La reconnaissance d'activités basée sur les données capteurs d'un environnement intelligent est donc une capacité clé des applications liées à la santé, au handicap et au vieillissement. De nombreux projets autour du monde se sont attaqués à cette tâche qui consiste donc à identifier l'activité en cours à partir des données diverses récupérées (D. COOK et al., 2009 ; INTILLE et al., 2006). Ce fut aussi une partie importante de mes travaux de thèse (FLEURY, VACHER et NOURY, 2010).

Indépendamment des capteurs utilisés, pour extraire les informations sur ces données, les algorithmes de reconnaissance d'activités discrétisent les données segmentées en fenêtres temporelles, que ces fenêtres soient de taille fixe ou variable (temporellement parlant). Chacune de ces fenêtres est alors étiquetée avec une activité, celle en cours (ou celle majoritaire dans la fenêtre ou de classe inconnue ou de rejet si c'est une transition ou une superposition d'activités, etc.). Lorsque les activités sont bien réalisées séquentiellement, ce fenêtrage ne pose que relativement peu de problème. Cependant, lorsque les activités sont entrelacées, cela est plus problématique, car une fenêtre pourra contenir plusieurs activités et il faudra alors prendre une décision concernant la manière de traiter ces cas. Dans les travaux qui seront présentés, nous avons choisi deux approches différentes : (1) créer une classe de « rejet » qui se charge de traiter ces fenêtres ou ces activités de transitions ou inconnues (FLEURY, VACHER et NOURY, 2010) et (2) travailler sur des algorithmes traitant ces données comme des flux et avec des algorithmes incrémentaux pour l'apprentissage (KRISHNAN et al., 2014). De plus, un problème de déséquilibre entre les classes se pose, sur lequel

nous reviendrons dans ce chapitre.
La partie segmentation se charge de diviser les données afin de les envoyer (ou leur version attribut du moins) au classifieur pour obtenir une réponse sur celle-ci. Or, il n'existe pas de manière « simple » de connaître le début et la fin d'une activité. Sans cela, nous devons donc faire un choix sur la manière de découper les données pour savoir à quel moment la décision sera prise. Pour cela plusieurs méthodes existent et vont être présentées succinctement.
Pour reprendre de manière concise et rassemblée, les verrous scientifiques à lever dans les travaux présentés dans ce chapitre sont les suivants :

— Comment acquérir, mettre en forme et traiter efficacement des données issues d'un habitat intelligent, données qui sont souvent hétérogènes, pour la plupart à valeur donnée sur événement (et non échantillonnées périodiquement) et bruitées ? L'acquisition/la mise en forme de ces données va inclure le fait de construire des fenêtres temporelles ou bien de traiter de ces données comme un flux. L'extraction de caractéristiques informatives fera partie également des traitement à appliquer à ces données.

— À partir de ces données, sommes-nous en mesure de déterminer des informations intéressantes sur les activités de la personne puis sur ses habitudes de vie et les possibles dérives par rapport à ces habitudes ?

— Peut-on utiliser ces dernières informations (sur les activités et les habitudes) pour permettre à la personne de mieux vivre au quotidien chez elle, en adaptant le comportement de l'habitat intelligent aux besoins et aux souhaits de l'occupant ?

Pour répondre à ces questionnements, dans ce chapitre, une première partie état de l'art va revenir sur les différentes étapes du traitement des données dans les habitats, puis seront abordés successivement des résultats en termes d'extraction et de sélection d'information puis de classification. Enfin, une section abordera la construction et la mise à jour de profils utilisateurs dans ces habitats afin d'être en mesure d'offrir un meilleur service individualisé dans les automatisations.

4.2 État de l'art sur l'extraction d'informations dans les habitats intelligents

4.2.1 Segmentation des données

La figure 4.2 (Figure 1 de la publication (YALA et al., 2017)) illustre les trois méthodes de segmentation possibles de ces données.

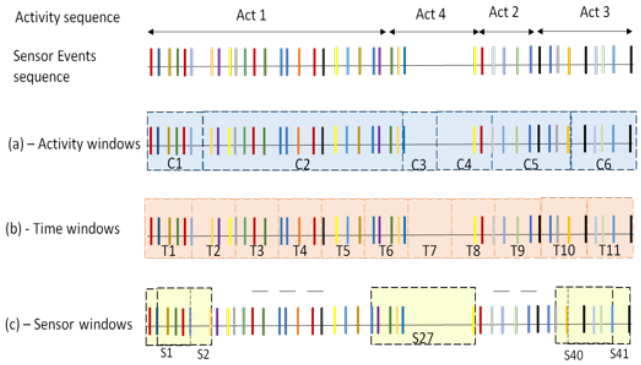


FIGURE 4.2 – Présentation succincte des différentes approches de segmentation. Le haut présente la séquence de déclenchement/mesure des capteurs et les activités en cours de réalisation au moment de ces déclenchements (de 1 à 4). (a) : une fenêtre pour une activité, (b) : fenêtre fixée par le temps et (c) : fenêtre avec un nombre fixe de déclenchements/mesures de capteurs.

4.2.1.1 Segmentation basée sur l'activité

Comme énoncé précédemment, il n'y a pas de moyen « simple » de faire cette segmentation, c'est à dire de détecter précisément le début et la fin de chacune des activités. Cependant, certains travaux nécessitent l'emploi d'une telle

segmentation pour effectuer la reconnaissance. Il faut donc réussir à identifier le point de changement dans l'activité de la personne (BAO et al., 2004). À partir de cela, chaque fenêtre correspondra à une activité, mais plusieurs inconvénients se posent. Le premier est que les activités ne sont pas toujours distinctes (une personne pourra commencer une activité en finissant la précédente), à partir de ceci, il est compliqué de trouver le point de séparation le plus pertinent pendant la phase d'apprentissage. Le second désavantage est que cette technique n'est pas « en ligne ». Dans les algorithmes que nous tentons de proposer, nous voulons toujours nous baser sur des méthodes qui ne considèrent que les données précédentes, et qui n'ont pas besoin à un moment donné des évènements futurs pour travailler. Or ce n'est pas le cas de cette méthode qui va, pour prendre sa décision d'appartenance pour un point donné, exploiter la connaissance du ou des points suivants.

4.2.1.2 Segmentation basée sur le temps

Cette méthode est la plus simple et la plus utilisée. Pour faire la séparation des données, l'algorithme va travailler sur des fenêtres temporelles de taille fixe. Cela revient à faire un traitement par lot (après avoir reçu un certain nombre d'échantillons ou après avoir attendu un certain temps). Deux raisons ressortent à la prédominance de cette méthode. La première est sa simplicité d'implémentation (qui ne nécessite qu'un horodatage des données). La seconde raison est sa « cohérence ». En effet, dans ce cas, tous les attributs traités sont basés sur le même temps de mesure par exemple, ce qui peut être important notamment dans le cas de données continues (BAO et al., 2004 ; TAPIA et al., 2004 ; WANG et al., 2012)(FLEURY, VACHER et NOURY, 2010).
L'un des points importants de cette méthode reste la sélection du temps de fenêtrage (GU et al., 2009). Des résultats l'horizon de fenêtrage et sur le recouvrement des fenêtres temporelles ont été publiés dans (CHAHUARA, FLEURY, VACHER et al., 2012) et seront présentés à la fin de la section concernant l'extraction et la sélection d'attributs (page 69).

4.2.1.3 Segmentation basée sur les événements (déclenchement de capteurs)

Pour cette méthode, les données sont segmentées en fonction du nombre d'événements mesurés (KRISHNAN et al., 2014). Un événement est un déclenchement d'un capteur, donc l'obtention dans la base de données d'une nouvelle valeur pour l'un d'entre eux. En choisissant un nombre N d'événements, l'algorithme retient toujours ce même nombre, quel que soit le temps qui s'écoule entre les événements. Il est donc évident que le temps de chacune des fenêtres va être différent. Ainsi, pendant la réalisation d'une activité qui demande de nombreux mouvements et des interactions avec l'environnement, plusieurs événements vont se déclencher. Au contraire, pendant des périodes plus « silencieuses », ce nombre d'événements va diminuer (lorsque par exemple la personne se repose). La prise en compte de certaines données est remise en cause selon les cas et il peut être utile de pondérer ces données en fonction notamment de leur écart temporel avec la dernière donnée.
Les travaux de YALA et al. (2017) se sont basés sur cette méthode et notamment sur les travaux de KRISHNAN et al. (2014) tout en proposant des modifications sur l'extraction des données. Par choix, ces travaux ne seront pas présentés dans ce manuscrit. Ces travaux donnent lieu à une extraction puis une classification avec des algorithmes incrémentaux. Cependant ces travaux sont en cours de finalisation et ne peuvent ainsi pas être présentés de manière complète et satisfaisante. La partie précédente sur l'extraction d'information concernant ces travaux a donc été omise également de ce chapitre.

4.2.2 Extraction des caractéristiques par fenêtrage

4.2.2.1 Fenêtrage basé sur le temps

Dans le cas d'un fenêtrage temporel, nous choisissons donc une durée de la fenêtre pour laquelle il nous faut ensuite extraire des caractéristiques pertinentes des signaux ainsi délimités.
Certains auteurs se sont basés sur des règles de logique simples pour créer des descripteurs (e.g. (TAPIA et al., 2004)). Dans FLEURY, VACHER et NOURY (2010), issus de mes travaux de thèse, l'extraction a porté sur des caractéristiques statistiques dépendant des types de capteurs. Pour les capteurs donnant des valeurs réelles, la moyenne et l'écart type

des valeurs recueillies dans la fenêtre temporelle étaient extraits. En ce qui concerne les capteurs ayant des valeurs binaires (ouverture de porte, présence, etc.), le pourcentage de temps passé, dans la fenêtre, dans l'une des deux positions était utilisé. Enfin, pour les capteurs liés entre eux ou ayant un nombre de valeurs finies, le pourcentage de temps dans chaque état (pour la présence) ou le nombre d'évènements pour chaque type (pour le son par exemple) étaient pris en compte.

4.2.2.2 Fenêtrage basé sur les évènements

Soit un ensemble d'évènements mesurés (de valeur de capteurs reçues) $[E_1, E_2, \ldots, E_N]$ dans un habitat équipé. Nous avons, en général, pour chacun des items, des informations sur le capteur qui l'a généré, sur l'heure et sur la valeur. En considérant N comme le nombre d'évènements, nous avons donc le vecteur représentant les valeurs de ce capteur à un moment donné pour chaque fenêtre. Un évènement E_i est par ailleurs caractérisé par une séquence (des valeurs la précédant) $[E_{i-N}, E_{i-N+1}, \ldots, E_i]$. N est choisi empiriquement, mais influencé par le nombre d'évènements moyens dans les activités recherchées.
À partir de ceci, une méthode proposée par (KRISHNAN et al., 2014) pour extraire des caractéristiques à partir de ces mesures est simple. Nous prenons en compte la taille de la fenêtre (en prenant l'heure des évènements extrêmes nous donnant la durée de la fenêtre également) et nous comptons ensuite le nombre d'évènements qui sont apparus pour chacun des capteurs de la maison. Ainsi dans un habitat à, par exemple, 20 capteurs, nous aurons $20+3$ attributs (20 attributs pour le comptage de chacun des capteurs, deux attributs pour les heures de début et de fin et un attribut pour la durée). La décision à prendre pour la base d'apprentissage est choisie en fonction de l'activité lors de l'évènement E_i. Cette méthode ne prenant pas en compte les écarts temporels entre les évènements (par exemple nous pouvons avoir quatre heures, voire bien plus, entre deux évènements lors de la nuit) une seconde méthode est proposée en pondérant les évènements. Cependant, lorsque des fenêtres représentant le passage d'une activité à une autre apparaissent (transitions), les évènements, même peu distants, peuvent ne pas avoir de liens entre eux. Pour corriger ceci, une seconde méthode est proposée dans la littérature pour pondérer en fonction d'informations mutuelles apportées par une suite de deux évènements. L'information mutuelle sur deux variables aléatoires mesure à quel point l'une des valeurs nous donne des informations sur l'autre. Dans ce contexte d'habitat intelligent, chaque capteur est considéré comme une variable aléatoire avec ses réalisations. En considérant E_i et E_j deux évènements, alors l'information mutuelle $MI(i,j)$, est définie par :

$$MI(i,j) = \frac{1}{N}\sum_{l=1}^{N-1} \delta(E_l, E_i) \cdot \delta(E_{l+1}, E_j) \tag{4.1}$$

$$s.t.\ \delta(E_l, E_i) = \begin{cases} 0 & if \quad E_l \neq E_j \\ 1 & else \end{cases} \tag{4.2}$$

Ce terme prend la valeur de 1 lorsqu'un évènement apparu sur le capteur i est suivi par un évènement sur le capteur j. La valeur de cette information mutuelle est liée à la proximité des deux déclenchements. Elle nous indique le fait que deux capteurs sont concernés par la même activité. La matrice d'information mutuelle est calculée hors ligne sur les données d'apprentissage. Elle est ensuite utilisée pour ajouter un poids donnant l'influence d'un évènement dans une fenêtre lorsque l'on construit le vecteur d'attributs. À la place de compter le nombre d'évènements, sont sommées les contributions de chaque capteur basées sur l'information mutuelle.
Les travaux de YALA et al. (2017) ont proposé d'améliorer ces vecteurs mais ne seront pas présentés dans ce manuscrit. Des travaux sont en cours et suivent ceux-ci.

4.3 Ensemble de données utilisés dans ce chapitre

Dans cette première section, nous allons d'abord décrire les deux ensembles de données utilisés dans la section suivante de ce chapitre. Ceux-ci ont été acquis à l'université de Grenoble et sont devenus depuis des bases de données publiques qui ont été mises à disposition de la communauté scientifique.

Suite à mes travaux de thèse, nous avons continué à travailler avec Michel Vacher, François Portet et leur doctorant Pedro Chahuara [1] sur des descripteurs basés sur un découpage temporel du signal issu des habitats intelligents.
Les données sur lesquelles se sont concentrés ces travaux sont issus des deux plateformes présentées sur la figure 4.3. La plateforme HIS du TIMC-IMAG a été utilisée lors de mes travaux de thèse et l'ensemble de données acquises dans ce cadre a été rendu public (FLEURY, VACHER, PORTET et al., 2013) [2]. Le second, celui du laboratoire LIG, a été utilisé dans le cadre du projet SWEET-HOME (projet ANR) et des données d'activités y ont de même été acquises (VACHER, LECOUTEUX et al., 2014).

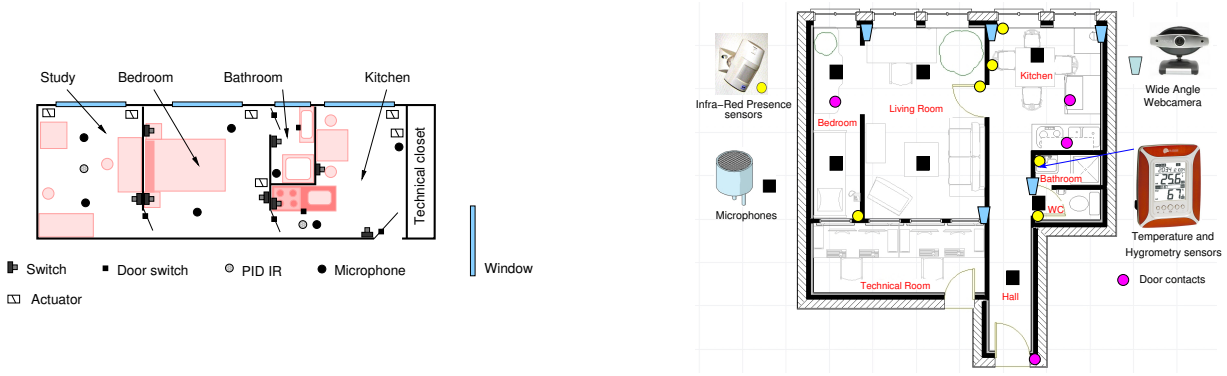


FIGURE 4.3 – Les deux plateformes d'expérimentation. À gauche la plateforme DOMUS du LIG, à droite la plateforme HIS du laboratoire TIMC-IMAG

Ces deux plateformes vont être présentées plus précisément dans les deux sous-sections suivantes. À ces données vont s'ajouter un travail, dans la dernière section (section 4.6, page 78) avec la plateforme Eco-Confort de l'IMT Lille Douai. Cependant, dans ce cas, nous n'utilisons pas un ensemble de données publiques comme les deux décrits ici, cette plateforme sera donc présentée à part dans la dernière partie de ce chapitre.

4.3.1 HIS

Le corpus HIS a été acquis dans le cadre des travaux de thèse que j'ai effectués sur la surveillance d'une personne sur son lieu de vie, dans le but d'aider les gériatres à évaluer la dépendance d'une personne (FLEURY, VACHER et NOURY, 2010). Le corpus a été acquis à la faculté de médecine de Grenoble, dans l'appartement spécialement aménagé de 47m^2 contenant l'ensemble des commodités communes (une chambre, un salon, une cuisine (avec placards de rangements, réfrigérateur, etc.), une salle de bain avec une douche et des toilettes. Nous y retrouvons un certain nombre de capteurs : des capteurs infrarouges pour la présence, placés dans chacune des pièces de l'appartement ; des contacteurs de portes sur les portes de l'appartement ainsi que pour la détection de l'utilisation de certaines commodités (réfrigérateur, placards de la cuisine, commode dans la chambre ; des microphones qui permettent d'enregistrer les sons de la vie courante ainsi que les paroles ; un capteur de température et d'humidité et enfin des caméras grand-angle installées dans quasiment toutes les pièces et permettant d'indexer les signaux uniquement. La figure 4.3 (droite) présente l'organisation de l'appartement ainsi que la répartition des capteurs dans celui-ci.
Sept activités avaient été sélectionnées dans les grilles couramment utilisées par les gériatres : préparation et prise d'un repas, activités d'hygiène, aller aux toilettes, s'habiller/se déshabiller, se reposer ou faire une sieste, se détendre et enfin communiquer avec un téléphone. Chaque personne ayant réalisé une mesure dans cet habitat devait réaliser ces différentes activités, sans contrainte d'ordre ou de temps, de nombre de réalisations, ou de manière de la réaliser. Au total quinze personnes saines et en pleine forme (six femmes et neuf hommes, jeunes), ont réalisé ce scénario qui était fait pour se rapprocher au mieux de la réalité.
Au total treize heures de données ont été acquises. Les personnes étaient âgées de 31,9 ± 9 ans (minimum 24, maximum 57), d'une taille moyenne de 1,74 ± 0,11 m (minimum 1,62, maximum 1,92), avec un poids moyen de 68,5 ± 9,11

1. Doctorant pour lequel je n'ai pas participé à l'encadrement de la thèse mais avec lequel j'ai eu une longue collaboration scientifique (lui et ses deux encadrants), menant à la publication de plusieurs articles
2. Corpus Grenoble Health Smart Home Dataset accessible publiquement en ligne suite à signature d'un accord à l'adresse https://sites.google.com/site/thonyfleury/health-smart-home-his-datasets.

kg (minimum 50, maximum 81). En plus des capteurs de l'habitat, les personnes portaient une centrale inertielle enregistrant leurs mouvements et en charge de déterminer leur posture (assis, debout, couché, en train de marcher) à chaque instant.
Le temps d'exécution moyen par participant pour l'expérimentation a été de 51 min 40 s (minimum 23 min 11 s, maximum 1 h 35 min 44 s). Les données ont été annotées en utilisant les sept activités retenues à l'aide du logiciel Advene [3] permettant de rassembler l'ensemble des vidéos en une et d'annoter sous forme d'un fichier XML les différents capteurs ou les différentes activités. L'ensemble de données a été divisé en deux, une partie pour le paramétrage (configuration pour la reconnaissance de son et la localisation) et une partie pour l'apprentissage-test fait par la suite. Dans cette partie des travaux, nous nous concentrons sur la partie paramétrage du corpus.

Classe	Paramétrage	Apprentissage/test	Global
Habillage/déshabillage	3.2%	2.1%	2.4%
Repas	19.4%	15.7%	16.6%
Aller aux toilettes	4.5%	4.6%	4.6%
Hygiène	3.5%	5.2%	4.8%
Téléphone	7.1%	6.1%	6.2%
Lecture/ordinateur/radio (détente)	22.1%	26.5%	25.5%
Repos	22.9%	21.2%	21.6%
Classe « *Autre* »	17.3%	18.6%	18.3%
Nombre de participants	4	11	15
Nombre de fenêtres	376	1221	1597
Durée	3h08m00s	10h10m30s	13h18m30s

TABLE 4.1 – Distribution des fenêtres temporelles dans le corpus HIS ($T = 60s$).

Le tableau 4.1 présente la répartition dans ce corpus.
Les activités considérées dans ces travaux sont inspirées des Activités de la Vie Quotidienne de la grille introduite par Katz (KATZ, 1983) qui est souvent utilisée dans le milieu des gériatres pour permettre de connaître les capacités d'autonomie d'une personne. Les activités effectués dans les expérimentations ont été choisies pour donner un contexte de vie dans la maison, mais aussi pour permettre d'aller plus loin sur la compréhension des habitudes de vie de la personne. Une autre classe peut être considérée dans l'étude, la classe que nous allons appeler *« Autre »* (présente dans les différents ensembles), qui représente aussi l'hypothèse NULL. Elle représente une période de temps dans laquelle nous n'avons actuellement aucune information sur l'activité en cours. Comme reporté dans certaines études (exemple (FLEURY, VACHER, PORTET et al., 2013)(KRISHNAN et al., 2014)) une grande partie des activités enregistrées dans un habitat sont soit des activités de transition, pas identifiable par un classifieur aisément (mouvement entre pièces, etc.) ou des activités qui ne nous intéressent pas. Dans notre approche, cette classe est considérée comme un label. Construire un unique modèle pour cette classe permet de se rapprocher d'une situation réelle.
Dans la suite, lorsque des résultats sont considérés sans cette classe *« Autre »*, toute fenêtre qui est étiquetée avec cette classe sera exclue tant de l'apprentissage que du test.

4.3.2 Domus

Le corpus multimodal Sweet-home (VACHER, LECOUTEUX et al., 2014) (noté SH) a été créé dans l'appartement Domus, créé par le Laboratoire d'Informatique de Grenoble (équipe Multicom) (GALLISSOT et al., 2013). Cet appartement fait 30m^2 et comprend une salle de bain, une cuisine, une chambre et un salon. Toutes ces pièces sont équipées avec un ensemble de capteurs et d'actionneurs telles que des capteurs de présence infrarouges, des contacteurs de

3. liris.cnrs.fr/advene

Classe	Paramétrage	Apprentissage/test	Global
Nettoyage	19.4%	20.6%	20.2%
Habillage/déshabillage	2%	2.6%	2.4%
Repas	31.2%	27.2%	28.3%
Hygiène	6.5%	6.8%	6.7%
Téléphone	2.3%	2.9%	2.7%
Lecture/ordinateur/radio	23.5%	22.9%	23.1%
Repos	10.8%	12.8%	12.3%
Classe « *Autre* »	4.3%	4.2%	4.2%
Nombre de participants	7	14	21
Nombre de fenêtres	603	1526	2129
Durée	5h01m30s	12h43m00s	17h44m30s

TABLE 4.2 – Distribution des fenêtres temporelles pour le corpus Sweet-Home ($T = 60s$).

porte, des caméras (utilisées également uniquement pour l'annotation), etc. En plus de ceci, sept microphones sont installés dans le plafond. L'appartement est complet et permet d'accueillir une personne pour y vivre plusieurs jours.

Contrairement au corpus HIS principalement basé sur des capteurs sur bus CAN et sur l'USB, Domus, de conception plus récente, est lui basé sur une installation KNX. En plus de KNX, nous y trouvons de l'UPnP pour l'audio et la vidéo ainsi que X2D pour la détection des ouvertures et fermetures de portes, fenêtres et placards.

Au total plus de 150 capteurs, actionneurs et fournisseurs d'information sont présents dans l'appartement et coexistent. Le son est traité et enregistré à part grâce à des cartes d'acquisition National Instrument permettant une acquisition temps réel et un traitement multi-canaux efficace (VACHER, FLEURY et al., 2010).

Les deux appartements, Domus comme le HIS, ont été créés pour être aussi proches que possible d'un appartement standard et ainsi permettre à la personne d'agir et de faire des activités de la vie quotidienne comme elle le ferait chez elle dans son propre appartement.

Pour cet ensemble de données, 21 personnes (dont 7 femmes) ont réalisé un ensemble d'activité, de même dans des conditions les plus proches possibles du réel. Ces personnes étaient âgées de 38,5 ± 13 ans (minimum 22, maximum 63), une taille de 1,72 ± 0,13 m (minimum 1,43, maximum 2,10), et un poids de 70 ± 14 kg (minimum 50, maximum 105). Les participants devaient rentrer dans l'appartement et agir comme ils le feraient chez eux. Précédemment à cette entrée et comme dans le cas de l'autre ensemble de données, un tour de l'appartement était fait afin de le présenter à la personne, afin qu'ils ne cherchent rien des outils de la vie quotidienne disponibles. Ensuite, les personnes devaient réaliser au moins une fois chacune des activités définies. La liste des activités était la suivante : (1) fermer la porte avec la gâche électrique, (2) s'habiller et/ou se déshabiller, (3) se laver les mains, (4) préparer un repas et le manger, (5) laver la vaisselle, (6) se laver les dents, (7) faire une sieste, (8) refaire le lit, (9) faire le ménage, (10) lire un livre ou écouter de la musique, (11) passer un appel téléphonique, (12) sortir de l'appartement – comme pour faire des courses – puis revenir, (13) utiliser le PC et appeler un proche, (14) prendre une douche et aller dormir.

Au total, 26 heures d'enregistrement ont été acquises dans cet ensemble de données avec une durée moyenne de scénario d'une heure. Environ 18h ont pu être conservées pour de la reconnaissance d'activités alors que le reste a été spécifiquement conservé pour de l'analyse audio. De la même manière que précédemment, les 7 activités précédentes ont été annotées dans Advene. Le corpus a de même été divisé en deux, une partie pour le réglage des paramètres (configuration pour la reconnaissance de son et la localisation) et une partie pour l'apprentissage et le test de la reconnaissance d'activités.

Le tableau 4.2 présente la répartition dans ce corpus.

Suite à la présentation de ces différents ensembles de données qui vont être utilisés par la suite, la section suivante va s'intéresser à l'extraction puis à la sélection des attributs dans les habitats intelligents, en vue d'une reconnaissance

d'activités efficace.

4.4 Sélection d'attributs dans les signaux issus des environnements intelligents

Nous avons effectué deux tests différents. Le premier est de savoir quels paramètres sont intéressants pour caractériser des activités de la vie quotidienne, en effectuant une sélection d'attributs sur les deux ensembles de données. Le second test portera sur les différentes durées de fenêtres temporelles. Ces deux tests seront décrits dans les deux sections suivantes.

4.4.1 Descripteurs

À partir des données extraites des lectures des différents capteurs (pour la plupart évènementiels mais avec certains capteurs à lecture périodique comme les capteurs d'ambiance ou d'activité électrique), un vecteur V d'attributs est créé dans le cas des deux corpus de données. La durée T de la fenêtre W est fixée à 60s. Les attributs sont calculés à partir des données des différents signaux.
Ainsi, pour les données Sweet-Home, nous retrouvons :

— Des données concernant la localisation avec entre autres le pourcentage de temps dans chaque pièce, la pièce prédominante, la pièce précédente, le temps depuis lequel la personne est dans la pièce, le nombre de détections de chaque pièce, etc.

— Les données concernant les contacteurs, avec la position de ceux-ci dans la fenêtre, etc.

— Les données concernant la lumière (ou les volets, les portes, les fenêtres), avec par exemple le nombre d'allumages/extinctions (ou de fermetures/ouvertures), le temps passé allumé (ouvert), etc.

— Les données concernant l'utilisation en courant dans l'appartement, avec par exemple les dernières mesures.

— Les données sonores avec par exemple le nombre de sons dans chacune des pièces, la répartition son/parole, etc.

— Des données de capteurs d'ambiance avec par exemple l'utilisation en eau chaude et froide, le capteur de CO_2, les capteurs de température, etc.

Au total, dans le corpus sweet-home, 94 éléments sont extraits.
En ce qui concerne le second ensemble (HIS) les attributs extraits des signaux sont les mêmes, en se restreignent aux capteurs disponibles (capteurs de porte, d'utilisation de certaines commodités, présence, son, etc.). Cela résulte ainsi à un ensemble de données de 26 éléments.
À partir de ces descripteurs, nous allons en sélectionner les plus importants dans les deux cas de ces deux corpus.

4.4.2 Sélection des attributs

Dans les deux ensembles de données précédents, toute fenêtre qui n'est pas identifiée comme appartenant complètement à l'une des sept activités sélectionnées a été classée dans la classe de rejet (la classe « *Autre* »). Les activités ont été indexées à la main à partir des vidéos. Toute fenêtre temporelle hors des activités indexées appartiendra à cette classe. Dans la suite nous allons décrire la sélection d'attributs faite sur ces deux ensembles de données.

4.4.2.1 Méthode de sélection

Comme expliqué précédemment, la sélection d'attributs est une étape importante afin de réduire la dimension du problème et ainsi d'en réduire d'une part le temps de résolution, mais aussi et surtout de diminuer les risques liés à l'interprétation des distances dans les espaces à grande dimension.
La méthode de sélection d'attrributs utilisée se base sur le Ratio du Gain d'Information (Information Gain Ratio – IGR). Cette méthode en pratique donne en général de bons résultats (MOLINA et al., 2002) sur le type de données que nous avons et, en tant que méthode filtre (comme vu dans le chapitre précédent), est indépendante du modèle de classification utilisé. Le gain d'information est la base de certains algorithmes d'arbre par exemple C4.5 (QUINLAN, 1996) (pour le choix des paramètres à utiliser pour créer un nouveau nœud en prenant l'attribut qui maximise ce ratio.

Le gain d'information se base sur la notion d'entropie. L'entropie d'une variable V prenant un ensemble de valeurs différentes v, notée $H(V)$ est définie par :

$$H(V) = -\sum_{v} p(v) \cdot log_2(p(v))$$

où $p(v)$ est ici la probabilité d'apparition du symbole v.
L'entropie de cette variable V, sachant la variable X (et ses différentes valeurs x) est définie par :

$$H(V|X) = -\sum_{x} p(x) \cdot \sum_{v} p(v|x) \cdot log_2(p(v|x))$$

Le ratio du gain d'information d'un attribut $a \in A$, en considérant la classe $c \in C$, est obtenu par :

$$IGR(a,c) = \frac{H(c) - H(c|a)}{H(a)} \tag{4.3}$$

L'équation 4.3 est appliquée à chaque attribut et un seuil est fixé pour retenir un certain nombre d'attributs (soit un seuil sur la valeur de l'IGR soit un seuil sur le nombre d'attributs voulu).
Le calcul de l'IGR est fait sur chacun des attributs et sur l'ensemble des données considéré pour la paramétrisation. Au final, nous avons choisi d'être tout de même relativement souples sur la sélection et donc tout attribut qui a un gain d'information non nul (donc qui apporte réellement une information) sera retenu.

4.4.2.2 L'ensemble de données Sweet-Home

Type d'attribut	Nom des attributs retenus (ayant un IGR strictement positif)
Location	**PercentageLocationRoom1**, **PercentageLocationRoom2**, **PercentageLocationRoom3**, **PercentageLocationRoom4**, **PredominantRoom**, **LastRoomBeforeWindow**, NumberOfDetectionPIROffice, NumberOfDetectionPIRKitchen, **TimeSinceInThisRoom**, **PercentageAgitationRooms**
Switches	SwitchBathroomUse, SwitchBedroomBed, SwitchBedroom, SwitchOffice, SwitchSinkKitchen
Lights	**PercentageTimeLightBathroomOn**, ActivationDeactivationLightBathroomSink, ActivationDeactivationLightBedLeft, ActivationDeactivationLightBedRight, ActivationDeactivationLightKitchenSink, PercentageLightOfficeOn, PercentageLightKitchenSinkOn, PercentageLightKitchenTableOn, PercentageLightBedLeftOn, PercentageLightBedRightOn
Shutter	**PercentageShutterBedroom**, **PercentageShutterBedroom2**, PercentageShutterDesk2, PercentageShutterKitchen, ActivationDeactivationShutterBedroom, ActivationDeactivationShutterDesk, **PercentageCurtain**, PercentageShutterOffice
Power	PowerLastUse, PowerLastLastUse, PowerLastLastLastUse
Doors and Windows	ActivationDeactivationNumberOfDoorBedroom, ActivationDeactivationNumberOfDoorBathroom, ActivationDeactivationDoorCupboardKitchen, ActivationDeactivationDoorFridge, ActivationDeactivationNumberOfWindowBedroomBathroom, PercentageAgitationDoors
Sounds	**SoundsKitchen**, **SoundsDinningRoom**, SoundsBathroom, SoundsOfficeDoor, SoundsBedroomWindow, SoundsOfficeWindow, SoundsBedroomDoor, SpeechBedroomDoor, SpeechBedroomWindow, SpeechBathroom, **SpeechKitchen**, SpeechOfficeDoor, SpeechOfficeWindow, SpeechDinningRoom, **PercentageAgitationSounds**, **PercentageAgitationSpeech**, **PercentageTimeSound**, **PercentageTimeSpeech**
Divers	ColdWaterTotal, HotWaterTotal, **TotalAgitation**, AmbientSensorCO2Bedroom, AmbientSensorTemperatureOffice, AmbientSensorTemperatureBedroom
Class	One of : cleaning, dressing up, eating, hygiene, phone, sleeping, reading/computer/radio, unknown activity/transition

TABLE 4.3 – Attributs sélectionnés dans le corpus Sweet-Home. Les vingt meilleurs (avec comme critère la valeur de leur IGR) sont mis en gras

La sélection d'attributs sur l'ensemble Sweet-Home a réduit l'ensemble des données de 94 à 65 attributs.
Le tableau 4.3 montre les attributs conservés par cette sélection. Dans ce tableau, les vingt attributs ayant le plus important gain d'information (donc sélectionnés en premier par un algorithme d'arbre et également étant les plus informatifs sur la classe d'un échantillon).
Ces résultats montrent surtout l'importance cruciale de deux types de données que sont les données de localisation et les données de son. Ce qui représente la base des travaux avec le HIS. Les capteurs ajoutés dans DOMUS apportent une information supplémentaire mais moins primordiale que les informations apportées par les capteurs présent dans le HIS.

4.4.2.3 L'ensemble de données HIS

Type	Nom des attributs retenus (ayant un IGR strictement positif)
Location	PredominantRoom, LastRoomBeforeWindow, PercentageLocation (in every room), TimeSinceInThisRoom
Doors	ActivationDeactivationCupboardDoor, ActivationDeactivationDressingDoor
Sounds	Sound on all the microphones
Speech	Speech in Entrance, Hall, Shower, WC, Kitchen
Class	One of : dressing up, eating, elimination, hygiene, phone, sleeping, reading/computer/radio, unknown activity/transition

TABLE 4.4 – Attributs sélectionnés dans le corpus HIS (ici nous ne précisons pas les 20 meilleurs car les écarts sont plus faibles et le nombre de dimensions éliminées moins important)

En suivant la même méthode sur ce second corpus, nous retrouvons, à partir des 26 dimensions d'origine, une sélection avec 23 dimensions par IGR. Ce corpus avait un nombre d'attributs très restreint, et beaucoup étaient portés sur la localisation ou le son. Il est donc logique de retrouver un plus grand nombre des attributs d'origine dans l'ensemble de données final, comme montré par le tableau 4.4.

4.4.3 Temps du fenêtrage et son influence

Dans (CHAHUARA, FLEURY, VACHER et al., 2012), nous nous sommes intéressés, à partir de ces données présentées précédemment, à l'influence du temps de fenêtrage sur les performances de classification. Le temps a varié entre 1 minute, 2 minutes, 3 minutes et 5 minutes. Le taux de recouvrement entre chaque fenêtre prenait trois valeurs différentes : 0%, 25% ou 50%.

Ainsi, en testant des versions optimisées de nos classifieurs (ici MLN et SVM qui seront présentés dans la section suivante), nous avons remarqué ce dont nous étions persuadés :

— Lorsque nous diminuons le temps de fenêtrage, les résultats sont meilleurs. Cela s'explique simplement par deux facteurs. Le premier est l'augmentation du nombre d'exemples disponibles dans la base d'apprentissage et donc la meilleure création de modèles, mais surtout la plus grande cohérence de chacune des fenêtres. En effet, lorsque nous diminuons le temps de fenêtrage, dans des activités complexes qui impliquent plusieurs phases, chacune des phases va ainsi pouvoir être mieux représentée. Par conséquent, si une autre réalisation de la même activité se fait avec un rythme différent, mais avec des étapes similaires, la reconnaissance sera alors meilleure.

— Le second point, sur le taux de recouvrement, donne les conclusions similaires. Plus le recouvrement est grand, meilleurs sont les résultats pour des raisons similaires. En augmentant le recouvrement d'une part nous augmentons artificiellement notre nombre de données, mais également nous avons plus de chance de mieux représenter les différentes étapes d'une activité.

4.4.4 Conclusions

Cette partie a présenté l'extraction d'attributs puis leur sélection dans les habitats intelligents. Concernant les deux ensemble de données que nous utilisons, HIS et Domus, la réduction effective du nombre de dimensions par le gain d'information a été donné et nous permet d'obtenir deux ensembles de données que nous pourrons utiliser dans les sections suivantes pour une tâche de classification. Le fenêtrage temporel a également été analysé, montrant qu'une fenêtre de petite taille (1 minute) avec un recouvrement important (50%) était une combinaison efficace. Ainsi, ces paramètres de fenêtrage ainsi que les attributs sélectionnés (23 dans le cas de HIS et 64 dans le cas de Domus) seront la base des travaux suivants.

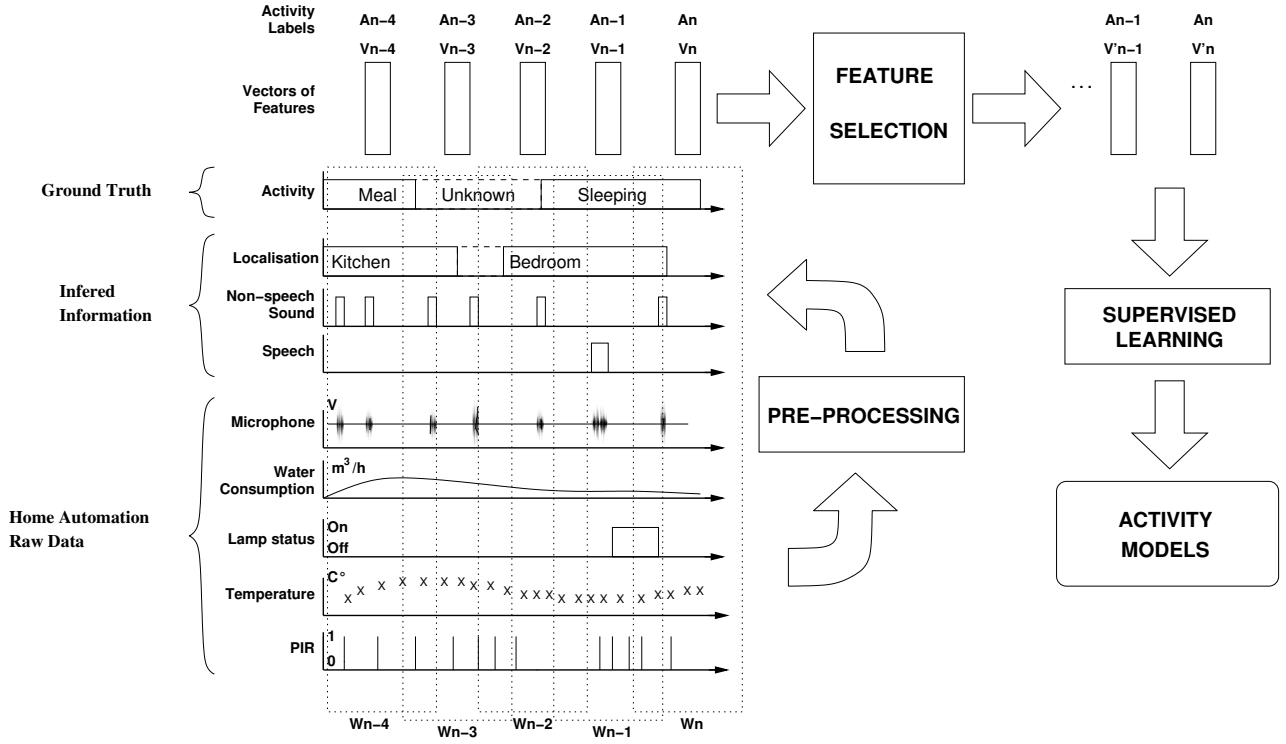


FIGURE 4.4 – Méthodologie générale pour le système de reconnaissance d'activités humaines.

4.5 Un aperçu de la classification d'activités dans les habitats intelligents avec des méthodes en ligne

Les données des habitats intelligents nous indiquent ce qu'il se passe à l'intérieur à un instant t. Ceci est vu au travers des attributs issus de la sélection de paramètres sur des données multimodales puis ensuite de la classification d'activités sur ces données. Cette section reprend des travaux qui ont été effectués ces dernières années, à savoir une classification en ligne de ces activités (causale – c'est à dire ne prenant en compte que des données passées dans le cas de flux/séries temporelles). Nous nous sommes basés sur des méthodes de classification nous permettant de prendre les décisions au fur et à mesure afin de pouvoir mettre ce genre de système dans un habitat réel et l'utiliser au quotidien. Le but final de cette utilisation est de cerner les habitudes de vie d'une personne afin d'en déduire toute évolution dans ces habitudes et toute dégradation de la réalisation des activités.

4.5.1 Introduction

Dans cette section, nous avons travaillé sur la reconnaissance en ligne d'activités en comparant les résultats d'un certain nombre de classifieurs. Notre but ici était de mieux comprendre l'impact de la méthode de classification, du déséquilibre des classes, de la présentation des données et de la considération ou non de certaines classes dans les résultats de reconnaissance d'activités et dans leur généralisation.

La méthode pour reconnaître des activités à partir de données brutes de capteurs qui arrivent sous forme d'un flot de données va faire voyager les données à travers différents niveaux d'abstraction, comme présentés par la figure 4.4. Les données brutes sont horodatées et peuvent se présenter sous la forme de valeurs symboliques, de valeurs d'état, de signaux régulièrement échantillonnés ou de signaux irrégulièrement échantillonnés (à chaque changement par exemple). Certaines de ces données sont prétraitées afin d'en extraire des informations de plus haut niveau, comme présenté dans le chapitre précédent avec les extractions d'attributs. Toutes les informations sont ensuite mises bout à bout afin de créer un unique vecteur de caractéristiques V_n (en rappel du début, ce vecteur serait un ensemble de statistiques sur les valeurs des capteurs E_i recueillies dans l'habitat), qui va correspondre à une fenêtre temporelle W_n de durée T. Le vecteur d'attribut peut être associé au label A_n qui peut être généré à partir de la vérité terrain quand

celle-ci est disponible (indispensable dans le cas d'apprentissage supervisé pour une partie des données). Dans ce cas, est prise l'activité qui a la plus longue durée dans W_n comme label. Comme présenté précédemment, si le nombre d'attributs est large, une sélection des plus pertinents peut être faite.
La reconnaissance d'activités dans les habitats intelligents peut être effectuée à l'aide d'un ensemble hétérogène de données brutes de capteurs, tels que des capteurs binaires donnant la présence (PIR), des microphones, des capteurs de températures, de la vidéo, des capteurs portés, etc. Pour supporter cette hétérogénéité, la stratégie adoptée est d'extraire un ensemble d'attributs à partir de chacun de ces capteurs, sur des fenêtres temporelles, afin de « résumer » les informations disponibles à un moment donné. Cette approche repose sur l'hypothèse que chacun instance d'une activité est composée d'un ensemble d'événements dont l'observation est réalisée par un ensemble de capteurs. Ces observations seront la signature de l'activité et peuvent être décrites par des statistiques sur les variables mesurées.
Les sections suivantes vont présenter ces résultats en termes de classification d'activités. Nous commencerons par une présentation des méthodes utilisées pour classifier les données reçues, puis le conditions expérimentales seront décrites avant de discuter sur les résultats de classification obtenus.

4.5.2 Différentes méthodes de modélisation d'activités

Nous allons considérer cinq méthodes dans cette partie, trois d'entre-elles, présentées en premières (HMM, MLN et CRF) sont séquentielles, les deux autres (SVM et Random Forest) ne le sont pas. Ces cinq méthodes ont été choisies pour les raisons suivantes : les HMM et CRF sont parmi les méthodes les plus utilisés en reconnaissance d'activités depuis longtemps, en faisant des méthodes de comparaison intéressantes. Random Forest quant-à-elle était une méthode qui commençait à être utilisée au moment de l'écriture de ces résultats, même si la tendance n'a pas continué. Enfin, SVM et MLN sont les méthodes que les laboratoires impliqués dans les travaux de recherche présentés tentent d'utiliser et veulent comparer avec la littérature. Une brève description des trois méthodes séquentielles (et surtout de leur utilisation dans un cadre de reconnaissance d'activités) est présentée ici. La partie état de l'art de (CHAHUARA, FLEURY, PORTET et al., 2016) donne un aperçu des travaux antérieurs utilisant ces méthodes. Une description plus poussée des différentes méthodes et de leur utilisation en reconnaissance d'activités est également disponible dans cette même référence.

4.5.2.1 Modèles de Markov Cachés (HMM)

Les modèles de Markov Cachés (HMM) (RABINER, 1989) sont beaucoup utilisés dans le domaine de la reconnaissance d'activités et en sont même devenus une approche relativement standard (DUONG et al., 2009 ; T. L. M. v. KASTEREN et al., 2011). L'une des utilisations la plus répandue consiste à calculer la séquence la plus probable d'états cachés $Y = \{y_1, y_2, \ldots, y_n\}$ en considérant une séquence d'observation $X = \{x_1, x_2, \ldots, x_n\}$ (ici dans notre cas les entrées sont les valeurs des attributs pour la fenêtre considérée et les sorties les probabilités qu'une activité soit en cours dans la fenêtre). Dans la suite, nous avons considéré des HMM ergodiques avec des états complètement connectés entre eux. Les HMMs ergodiques se basent sur deux hypothèses. La première, qui est vraie pour tout HMM, est l'indépendance conditionnelle des observations. Une observation x_t émise à un instant t ne dépend d'aucune autre observation, sachant l'état qui l'a généré. Dans de nombreux cas, cette hypothèse est fausse, mais en pratique et en première approximation, la classification fonctionne bien tout de même. Par exemple, si l'une des variables est la localisation de la personne, elle n'est pas indépendante entre des états consécutifs. En fonction de l'organisation de l'appartement, la localisation au temps $t+1$ ne peut être, par rapport au temps précédent, que dans la même pièce ou dans les pièces adjacentes. La seconde hypothèse, importante dans les HMM d'ordre 1 (tels que ceux considérés ici) suit le principe de Markov énonçant que les probabilités du HMM en un instant t ne dépendent que de l'état courant et de celui au temps $t-1$. L'activité réalisée dans une fenêtre en cours ne dépend que de la fenêtre précédente. En reconnaissance d'activité c'est une hypothèse crédible et acceptable.
Afin de modéliser les activités, un modèle séparé pour chacune des activités est créé. Chaque état caché est modélisé par un mélange de Gaussiennes. Le processus d'apprentissage a été réalisé pour chacune des activités et a consisté en l'estimation des probabilités initiales, des paramètres des mélanges de Gaussiennes (en utilisant l'algorithme EM) et des probabilités des observations pour chaque état et de la matrice de transition entre états. La convergence

vers les paramètres finaux est obtenue par l'algorithme Baum-Welch (algorithme itératif, qui permet d'estimer les paramètres du modèle qui maximisent la probabilité d'une séquence d'observables). Finalement, nous avons choisi empiriquement des modèles avec 2 états et 3 GMM par état. Pour choisir l'activité en cours, le HMM répondant avec la vraisemblance la plus élevée est choisi comme activité pour la fenêtre.

4.5.2.2 Les Conditional Random Fields (CRFs)

Les CRFs sont des modèles graphiques qui permettent de l'inférence probabiliste discriminante sur un ensemble de variables pour effectuer une tâche de classification (LAFFERTY et al., 2001).
Comme pour les HMMs, les CRFs peuvent classifier une séquence de variables $Y = \{y_1, \ldots, y_n\}$ à partir d'une séquence d'observations données $X = \{x_1, \ldots, x_n\}$. Cependant, les CRFs ne sont pas des modèles génératifs et donc n'ont pas pour vocation de modéliser les distributions conjointes $p(X, Y)$. À la place, les CRFs sont des modèles discriminants qui modélisent les distributions conditionnelles, $p(Y|X)$ mais sans avoir besoin de modéliser la distribution de la variable X. Graphiquement, un CRF se présente comme un graphe unidirectionnel. Dans le cas de la reconnaissance d'activités, nous considérons X comme étant un vecteur de caractéristiques décrivant une fenêtre temporelle et la variable cachée Y est inférée et correspond à l'activité la plus probable qui a permis de générer les observations.

4.5.2.3 Modélisation par des réseaux logiques de Markov (Markov Logic Network – MLN)

Un réseau logique de Markov (MLN) (RICHARDSON et al., 2006) est un modèle relationnel statistique qui combine une logique de premier ordre (First Order Logic – FOL) et les réseaux de Markov. Un modèle MLN est assez expressif pour inclure, par le biais du FOL, les principales relations qui existent parmi les éléments d'un environnement intelligent. Chaque relation logique se voit de plus attribuer un poids indiquant le degré de confiance en celle-ci. Cette représentation logique associée à son ensemble de poids peut être considérée comme un méta modèle qui, lors du processus d'inférence, permet la construction d'un réseau de Markov, un modèle probabiliste pur qui peut en même temps s'accommoder de variables incertaines. Un MLN est composé d'un ensemble de formules de logiques de premier ordre, chacune associée à un poids qui exprime le degré de vérité. Cette approche atténue le fait qu'une équation logique est vraie ou fausse. Nous allons considérer des MLNs Naïfs, en référence au modèles Bayesiens Naïfs, en considérant que la probabilité qu'une activité soit effectuée dans une fenêtre temporelle est indépendante des autres fenêtres temporelles. La structure sera ainsi plus rapide à apprendre. En complément, un modèle dynamique qui considère le problème de reconnaissance d'activités comme un processus stochastique dans le temps a été proposé dans le cadre de ces travaux. Ce modèle est appelé MLN dynamique et noté MLN dans la suite.

4.5.2.4 Modélisation d'activités par des méthodes non séquentielles : SVMs and Random Forests

Les deux dernières méthodes considérées sont les SVM et les forêts aléatoires (Random Forest – RF). Ces algorithmes, non séquentiels, sont exécutés sur les fenêtres temporelles indépendamment les unes des autres. Ces deux algorithmes ont déjà par le passé été utilisés dans le domaine de la reconnaissance d'activités en montrant de bonnes performances Random Forest et SVM étant des méthodes beaucoup plus courantes et facilement applicables à tout problème de classification, elles ne seront pas détaillées ici.
Une validation croisée à 10 partitions est faite pour optimiser les paramètres du classifieur en utilisant un ensemble d'évaluation. Pour ce qui est du réglage des paramètres de chacun de ces modèles, dans le cadre de SVMs, étant donné que nous utilisons un noyau gaussien, ce sont les paramètres C (relaxation) et σ (définition de la gaussienne) qui sont optimisés à l'aide d'une recherche sur un pavage de l'espace à 2 dimensions. Pour RF, le nombre d'arbres est testé afin de voir l'évolution de la performance en fonction de celui-ci

4.5.3 Expérimentations

Des expérimentations ont été effectués sur les deux ensemble de données présentés précédemment et leurs résultats sont présentés par la suite.

4.5.3.1 Évaluation des performances

La méthode utilisée pour évaluer les performances se base sur la validation croisée, mais en étant plus spécifique, car cette validation croisée est faite par sujets (Leave-One-Subject-Out-Cross-Validation – LOSOCV). Nous considérons l'ensemble des données (apprentissage + test) comme un ensemble d'enregistrements de différentes personnes (N participants). Pour chaque tour de validation croisée, 1 sujet est utilisé pour le test et les $N-1$ autres pour l'apprentissage. Cela nous permet aussi de tester la capacité de généralisation de notre algorithme en fonction des personnes, donc de voir que l'on peut reconnaître les activités d'une personne sans avoir perçu comment elle réalisait ses activités par une partie de l'ensemble d'apprentissage.
Pour évaluer la performance, nous calculons les performances globales en termes de précision de classification :

$$Acc_{Global} = \frac{\sum_i V_i}{\sum_i N_i}$$

avec V_i le nombre de fenêtres de la classe i correctement reconnues et N_i la cardinalité (le nombre total de fenêtres) de la classe i. La précision moyenne par classe est également calculée pour voir la capacité des méthodes d'apprentissage à reconnaître les classes indépendamment, sans biais de déséquilibre entre les classes. Elle est définie par

$$Acc_{Class} = \frac{\sum_i Acc_i}{N_c}$$

avec N_c le nombre de classes et $Acc_i = \frac{V_i}{S_i}$ la précision de la classe i.
Dans tous les tableaux de résultats présentés par la suite, la précision globale est donnée ainsi que la moyenne et l'écart type de cette précision calculée par participants.

4.5.3.2 Résultats globaux

Le tableau 4.5 résume synthétiquement les résultats de performance des différents algorithmes testés sur les deux ensembles de données.

	Sweet-Home			HIS		
Modèle	Exp 1 Sans «*Autre*»	Exp 2 Avec «*Autre*»	Différence	Exp 3 Sans «*Autre*»	Exp 4 Avec «*Autre*»	Différence
SVM	75,00	71,90	3,10	74,86	64,90	9,96
RF	82,96	80,14	2,82	70,72	62,32	8,40
MLN Naïf	79,20	76,73	2,47	75,45	66,81	8,64
HMM	74,76	72,45	2,31	77,26	67,11	10,15
CRF	85,43	83,57	1,86	75,85	69,29	6,56
MLN	82,22	78,11	4,11	75,95	65,82	10,13

TABLE 4.5 – Précision globale (%) sur les deux ensembles de données, avec et sans la classe «*Autre*» (la classe «*Autre*» comprenant, pour rappel, toute fenêtre qui n'est pas étiquetée dans la base de données)

Nous pouvons voir que l'approche basée sur CRF donne les meilleurs résultats dans 3 conditions sur 4 (Exp 1, 2 et 4) et que les HMMs donnent les meilleurs résultats dans le dernier cas (Exp 3). Les MLNs sont toujours en seconde ou troisième position avec de bons résultats. Le pire classifieur dans ce cas est SVM, mais HMM et RF arrivent à de moins bons résultats dans le cas du HIS en incluant la classe «*Autre*» (Exp 4). Pour l'ensemble sweet-home sans classe «*Autre*», un test de Kruskal-Wallis montre un effet significatif de la méthode sur le résultat ($\chi^2 = 16.22$, $p = 0.006$). Un test post-hoc avec un test de rang de Wilcoxon montre que cette différence significative est surtout apportée par CRF et HMM ($p = 0.032$). Quand la classe «*Autre*» est incluse, la significativité décroît ($\chi^2 = 17.78$, $p = 0.003$), mais HMM et CRF sont toujours moteurs dans cette différence ($p = 0.028$). Aucun des résultats sur le HIS ne montre une différence significative entre les algorithmes. CRF montre la plus faible différence avec ou sans l'utilisation de la classe «*Autre*».
De manière générale, CRF apparaît comme la méthode la plus performante (d'où son utilisation croissante ces dernières années dans le domaine). Par la suite, nous nous concentrons sur CRF et sur les autres modèles dynamiques (HMM, MLN et MLN Naïf) pour discuter des résultats plus en détail dans chacun des cas. Cela a pour but de montrer la manière dont les classifieurs modélisent bien ou non les classes d'activités et les problèmes posés par les contraintes de chacune

des activités ou des classes sur la reconnaissance globale.

4.5.3.3 Résultats sur les données Sweet-Home

Des résultats détaillés par classe avec ou sans la classe *« Autre »* sont présentés respectivement dans les tableaux 4.6 et 4.7. Sans la classe *« Autre »*, CRF montre les meilleurs résultats (85,43%) et la meilleure moyenne par classe (76,26%), suivi de près par MLN (82,22% globalement et 75,41% par classe). Les deux sont très largement au-dessus des HMM (74,8% globalement et 63,25% par classe). Dans le détail, CRF obtient les meilleurs résultats pour quasiment toutes les classes (*Nettoyage*, *Habillage*, *Repas*, *Hygiène*, *Repos*). MLN a de meilleurs résultats pour deux classes, à savoir *Téléphone* et *Lecture*. Une large supériorité de CRF est visible sur les classes *Habillage* (56,11±31,88%) et *Nettoyage* (84,25±12,93%) et de MLN sur la classe *Téléphone* (79,76±23,6%). Les HMMs montrent des performances correctes sur les classes *Hygiène* et *Lecture*, mais de très mauvaises performances dans le cas de *Habillage*.

Classe	HMM	CRF	MLN	MLN Naïf
Nettoyage	64,8% (66,1 ±18,9%)	**82,80% (84,25±12,93%)**	75,16%(76,71±12,39%)	74,2% (75,83±11 %)
Habillage/déshabillage	2,7% (7,1 ±26,7%)	**53,84% (56,11±31,88%)**	30,77%(28,33±32,4%)	25,6%(23,33 ±30%)
Repas	76,9% (76,1 ±28,5%)	**85,43% (87,76±16,49%)**	83,37%(82,67±18,52%)	75,42% (75,33±19%)
Hygiène	79,1% (77,6 ±25,7%)	79,80% (79,78±23,84%)	78,85%(76,44±26,4%)	**81,73% (80,49±21%)**
Téléphone	54,8% (55,5 ±33,8%)	50% (51,97±37,72%)	**81,82%(79,76±23,6%)**	65,9% (69,25±30%)
Lecture/ordinateur/radio	**92,1% (90,2 ±13,1%)**	91,14 % (91,08±8,67%)	91,71%(92,76±7,15%)	91,1% (91,99 ±8%)
Repos	72,4% (72,1 ±25,8%)	**90,81% (88,81±13,29%)**	86,22%(87,29±12,62%)	86,2% (87,14±11%)
Global	74,8%	**85,43**	82,22 %	79,2%
Classe	63,25%	**76,26%**	75,41%	71,45%

TABLE 4.6 – Résultats de classification pour l'ensemble des données Sweet-Home sans classe *« Autre »* (Exp 1) : forme résultats (moyenne par participant $\pm EcType$)

Lorsque la classe *« Autre »* est mise dans les ensembles d'apprentissage et de test, les résultats restent similaires. Toutes les performances diminuent sauf les HMM et MLN dans le cas de la classe nettoyage où les résultats sont quelque peu améliorés. Dans ce cas, MLN passe devant CRF. MLN encore une fois montre une large supériorité dans le cas de la classe Téléphone (80,1±23,56%) par rapport aux HMMs (49,8 ±35,8%) et à CRF (51,63±38,74%)

Classe	HMM	CRF	MLN	MLN Naïf
Nettoyage	70,8% (72,4 ±13,9 %)	79,82% (82,65±14,22%)	**81,85%(84,08±14,84%)**	74,84% (75,99±11%)
Habillage/Déshabillage	2,7% (7,1 ±26,7%)	**50%(52,78±26,66%)**	25,64%(23,89±28,67%)	20,51% (17,22 ±28%)
Repas	77,4% (76,7 ±23,9%)	**83,57%(87,30±15,86%)**	77,35%(79,13±22,65%)	76,63% (76,9±18%)
Hygiène	68,1% (65,8 ±25%)	**87,88%(77,32±25,29%)**	78,85%(77,4±21,61%)	79,81% (80,3±21%)
Téléphone	50% (49,8 ±35,8%)	54,34% (51,63±38,74%)	**79,55%(80,1±23,56%)**	68,18% (68,21±30%)
Lecture/ordinateur/radio	**92,1% (86,6 ±19,8%)**	91,66%(90,88±7,86%)	87,71%(90,16±10,34%)	90,86% (91,64 ±1%)
Repos	66,7% (67,2 ±26,6%)	**89,62%(87,56±13,87%)**	84,69%(85,25±14,4%)	85,2% (86,54±14%)
Autre	24,6% (23,2 ±18,8%)	**59,42%(50,37±36,85%)**	21,88%(18,96±23,36%)	18,75% (16,31±16%)
Global	72,45%	**83,57**	78,11%	76,73%
Classe	56,55 %	**74,54** %	67,19%	64,35 %

TABLE 4.7 – Résultats de classification pour l'ensemble des données Sweet-Home avec classe *« Autre »* (Exp 2) : forme résultats (moyenne par participant $\pm EcType$).

4.5.3.4 Résultats pour l'ensemble des données HIS

Des résultats détaillés par classe pour l'ensemble HIS sans et avec classe autre sont donnés respectivement dans les tableaux 4.8 et 4.9. Sans la classe *« Autre »*, les HMMs obtiennent les meilleurs résultats en global (77,3%) et par classe (71.0%), légèrement meilleurs que les MLN (75,95% globalement and 68,99% en moyenne par classe) et que CRF (75,85%

globalement et 66,71% par classe). Les tests statistiques n'ont cependant pas révélé de différence significative entre les modèles. De plus, les meilleures performances pour chaque classe sont assez bien distribuées entre les méthodes.

Classe	HMM	CRF	MLN	MLN Naive
Habillage/déshabillage	**46,2 % (30,8 ±40%)**	38,46%(23,73±39,97%)	30,77%(26,98±39,55%)	38,46%(28,63±39,52%)
Repas	90,6 % (90,2 ±17,7%)	**95,31%(94,81±9,5%)**	93,23%(93,43±12,86%)	93,75%(93,20±12,32%)
Aller aux toilettes	**85,7 % (65,5 ±46,7%)**	64,29%(46,61±43,35%)	48,21%(33,72±44,45%)	69,64%(49,48±40,47%)
Hygiènes	36,5 % (35,6 ±45,1%)	36,51%(41,44±36,28%)	**73,02%(52,27±50,56%)**	22,22%(26,52±42,63%)
Téléphone	83,8 % (76,6 ±36,3%)	81,08%(72,09±34,8%)	**91,89%(83,93±31,65%)**	83,78%(77,16±30,69%)
Lecture/Ordinateur/Radio	75,9 % (73,8 ±38%)	77,16%(75,4±35,64%)	75,93%(68,96±39,53%)	**77,16%(75,26±35,61%)**
Repos	**78,4 % (61,4 ±41,5%)**	74,13%(64,77±32,35%)	69,88%(65,13±37,93%)	75,29%(69,72±37,71%)
Global	**77,3%**	75,85%	75,95%	75,45%
Classe	**71,0 %**	66,71 %	68,99 %	68,7%

TABLE 4.8 – Résultats de classification pour l'ensemble des données HIS sans classe *« Autre »* (Exp 3) : forme résultats (moyenne par participant $\pm EcType$)

Lorsque la classe *« Autre »* est introduite, les résultats changent légèrement. CRF obtient les meilleurs résultats globaux (69,29%), mais pas par classe (59,07% contre 60,2% pour HMM). De même les résultats sont bien distribués entre les méthodes. Les bonnes performances de CRFs (et son meilleur classement entre sans et avec la classe *« Autre »*) sont dues à une meilleure reconnaissance de la classe *« Autre »* que les autres méthodes, classe qui représente pas loin d'un cinquième de l'ensemble des données.

Class	HMM	CRF	MLN	MLN Naive
Habillage/déshabillage	26,9 % (10,1 ±18,8%)	15,38%(4,25±9,74%)	**26,92%(18,36±32,73%)**	26,92%(17,89±31,82%)
Repas	88 % (87,6 ±20%)	**89,58%(87,37±20,25%)**	85,94%(84,52±19,05%)	88,54%(87,30±19,56%)
Aller aux toilettes	**76,8 % (61 ±46,9%)**	62,5%(42,07±38,89%)	57,14%(47,69±45,89%)	67,86%(48,73±39,64%)
Hygiènes	25,4 % (32,3 ±45,9%)	20,63%(26,36±40,56%)	**34,92%(34,55±48,24%)**	20,63%(27,73±43,09%)
Téléphone	77 % (69,8 ±32,4%)	**78,38%(69,68±35,54%)**	60,81%(55,65±36,62%)	60,81%(55,62±29,66%)
Lecture/Ordinateur/Radio	74,7 % (72,6 ±37,4%)	**75,93%(73,04±38,06%)**	69,44%(61,46±44,44%)	74,69%(72,88±38,18%)
Repos	**79,9 % (68,1 ±37,7%)**	70,66%(65,44±36,59%)	71,81%(64,2±39,11%)	71,04%(68,32±37,34%)
Unknown	32,9 % (33,9 ±7,8%)	**59,47%(60,33±14,72%)**	52,86%(53,43±14%)	51,98%(52,93±12,14%)
Global	67,1%	**69,29%**	65,82%	66,81%
Classe	**60,2 %**	59,07%	57,48%	51,08%

TABLE 4.9 – Résultats de classification pour l'ensemble des données HIS avec classe *« Autre »* (Exp 4) : forme résultats (moyenne par participant $\pm EcType$)

4.5.3.5 Performance des CRFs pour reconnaître les activités

Certaines classes sont très difficiles à distinguer des autres. Afin de faire une analyse plus fine, les tableaux 4.10 et 4.11 donnent les matrices de confusion [4] dans le cas des deux ensembles de données considérés avec la classe *« Autre »*.
Sans surprise, dans les deux corpus, la classe *« Autre »* se répartit entre les différentes autres classes dans la matrice de confusion, avec des conséquences plus marquées dans l'ensemble des données HIS, vu la place prise par cette classe.
Pour Sweet-Home, nous pouvons ajouter que *Repas* et *Nettoyage* se confondent. À noter que ces deux activités étaient couramment faites dans la même pièce. La classe *Lecture* a une spécificité faible avec des confusions fréquentes avec *Téléphone*, *Habillage* et *Repos*. Ce n'est pas surprenant, car cette classe est déjà composée de différentes sous-classes qui ont des comportements en commun avec les trois classes avec lesquelles elle est confondue.
Pour les données issues du HIS, *Aller aux toilettes* et *Hygiène* sont confondus de nouveau, car ils partagent la même pièce. En ce qui concerne la classe *Lecture*, de la même manière que dans Sweet-Home, cette classe est confondue

4. Dans ces tableaux nous avons également ajouté la sensibilité et la spécificité. Pour rappel, considérons le taux de vrais positifs **TP**, de vrais négatifs **TN**, de faux positifs **FP** et de faux négatifs **FN**. Sensibilité = $\frac{TP}{TP+FN}$ et Spécificité = $\frac{TN}{TN+FP}$

notamment avec *Repos*. Les deux activités impliquent d'être assis ou allongé, donc entraînent peu de mouvements et peu de sons. Enfin, pour la classe *Habillage*, la sensibilité est très faible. À noter que cette activité était courte, et réalisée entre le salon et la chambre (aucun mur n'étant présent). Le faible nombre d'exemples explique aussi la difficulté à apprendre cette activité.

	Nettoyage	Habillage	Repas	Hygiène	Téléphone	Lecture, etc,	Repos	Autre
Nettoyage	255	2	39	1	0	4	1	6
Habillage	0	20	1	1	0	3	2	2
Repas	48	1	364	8	0	1	0	6
Hygiène	3	3	3	80	0	1	2	4
Téléphone	1	0	0	0	23	3	0	0
Lecture, etc.	4	9	1	9	19	319	14	8
Repos	1	2	6	4	1	16	175	2
Autre	2	2	1	1	1	3	2	36
Sensibilité	81,21%	51,28%	87,71%	76,92%	52,27%	91,14%	89,29%	56,25%
Spécificité	95,63%	99,40%	94,24%	98,88%	99,73%	94,56%	97,59%	99,18%

TABLE 4.10 – Matrice de confusion pour CRF avec l'ensemble des données Sweet-Home

	Habillage	Repas	Toilettes	Hygiène	Téléphone	Lecture, etc.	Repos	Autre
Habillage	4	0	0	0	3	1	7	8
Repas	1	172	0	0	0	3	2	23
Toilettes	0	0	35	37	0	0	0	7
Hygiène	0	0	18	13	0	1	1	7
Téléphone	0	0	0	0	58	1	0	7
Lecture, etc.	5	1	0	4	4	246	48	23
Repos	7	2	0	0	0	59	183	17
Autre	9	17	3	9	9	13	18	135
Sensibilité	15.38%	89.58%	62.50%	20.64%	78.38%	75.93%	70.66%	59.47%
Spécificité	98.41%	97.18%	96.22%	97.67%	99.30%	90.52%	91.16%	92.15%

TABLE 4.11 – Matrice de confusion pour CRF avec l'ensemble de données HIS

4.5.3.6 Analyse complémentaire

Pour évaluer l'impact de l'inclusion de la classe *« Autre »* sur le processus d'apprentissage, une analyse complémentaire a été faite sur les données HIS, en prenant le meilleur et le moins bon classifieur (CRF et RF). Dans cette analyse, nous avons fait varier le pourcentage d'exemples de la classe *« Autre »* dans l'ensemble des données. La figure 4.5 montre une décroissance rapide de RF au-delà de 8% des données dans la classe *« Autre »* alors que pour CRF, une chute rapide au début est constatée (même avec 1%) puis une stabilisation ensuite jusqu'à environ 10%. Il semble donc que CRF soit plus robuste à l'inclusion de cette classe perturbant l'apprentissage. Cela rejoint les constatations de SEGAL (2004) sur les performances de RF en milieu bruité.

4.5.4 Discussion

Afin de pouvoir réaliser une reconnaissance en ligne, des méthodes séquentielles (HMM, CRF, MLN) ont été utilisées en utilisant les données de manière causale (donc en ayant uniquement connaissance du passé). De plus, pour se mettre dans des conditions réalistes, nous avons utilisé la classe *« Autre »* (KRISHNAN et al., 2014) et donc pris en compte les transitions entre activités. Cela différencie ces travaux de nombreux autres qui sont faits « hors lignes » (FLEURY, VACHER et NOURY, 2010)(NAZERFARD et al., 2010; VELIK, 2014). Cela comprend aussi mes travaux de thèses, pendant laquelle j'ai acquis l'ensemble des données HIS, et sur lesquels les SVMs donnaient de bien meilleurs résultats (la classe autre n'était pas prise en compte et les données ont été réindexées et retraitées en partie). En hors-ligne, il est également possible d'appliquer des prétraitements sur l'ensemble des données d'apprentissage, ce qui n'est pas le cas dans ces travaux.

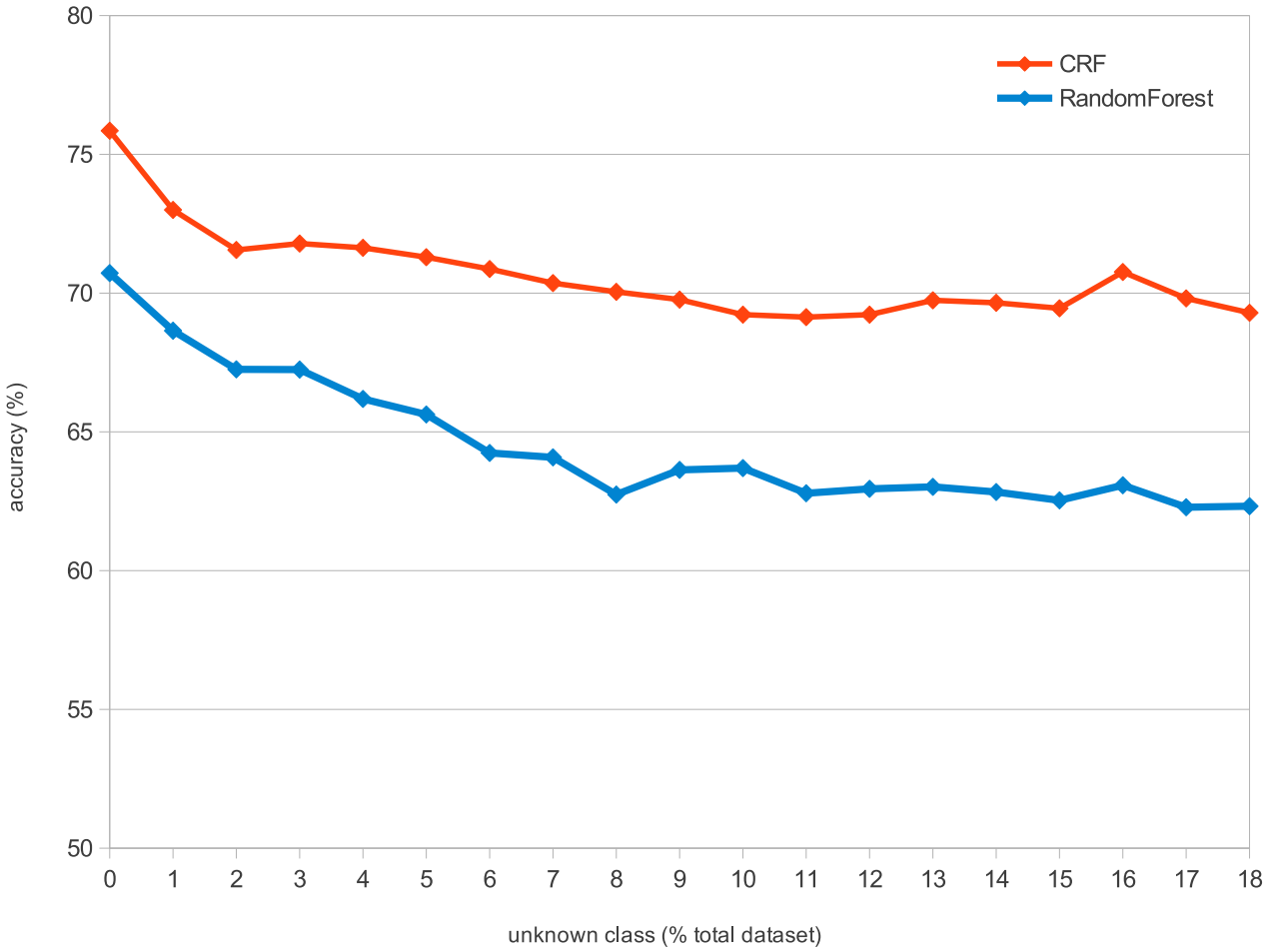


FIGURE 4.5 – Précision vs. pourcentage de classe *« Autre »* dans l'ensemble des données pour CRF et RF sur la base de données HIS.

Les résultats montrent que les méthodes séquentielles (CRF, MLN, HMM) sont supérieures mais pas largement au-dessus des méthodes non séquentielles (SVM, RF), statistiquement parlant. Cela peut s'expliquer par deux raisons. La première est que la reconnaissance d'activités est fortement dépendante de la localisation de la personne. La présence de cette information seule dans la fenêtre permet une bonne reconnaissance dans les modèles basés sur des instances. En seconde raison, la manière dont nous avons réalisé les fenêtres temporelles (inspirée de mes travaux de thèse) permet de mettre un peu d'historique dans celle-ci. Par exemple, le temps passé par la personne dans une pièce s'accumule d'une fenêtre à une autre si la personne ne change pas de pièce.
Cependant, nous pouvons souligner qu'un modèle séquentiel est toujours classé en premier dans toutes les conditions (que ce soit CRF ou HMM). CRF est généralement le meilleur algorithme pour faire de l'apprentissage en ligne d'activités à partir de capteurs non visuels simples. Il avait déjà été montré que CRF donnait de bons résultats de classification (VAIL et al., 2007 ; T. v. KASTEREN et al., 2010) mais celui-ci n'avait jamais été confronté à des modèles non séquentiels. Les différences de performances entre CRF et HMM/MLN peuvent s'expliquer par leur nature discriminative ou générative. CRF est entrainé en maximisant la vraisemblance sur l'ensemble entier des données alors que HMM et MLN vont maximiser la vraisemblance par classe. CRF biaise donc son apprentissage vers les classes dominantes comme le font les modèles statistiques non séquentiels (SVM, RF). HMM et MLN au contraire vont modéliser indépendamment les classes et ceci explique pourquoi les performances sur certaines classes moins représentées sont meilleures dans ce cas.
Le fait de prendre en compte la classe *« Autre »* réduit les performances de quasiment tous les modèles. Cependant, les approches génératives sont les plus touchées. En effet, MLN et HMM montrent la plus forte diminution des performances du fait de leur incapacité à modéliser cette classe. De plus, dans notre cas, comme rappelé en introduction, cette classe *« Autre »* contient surtout des éléments de transitions et donc des éléments qui pourraient être dans les

différentes classes, d'où le fait qu'elle soit difficile à modéliser. Ces effets sont encore plus notables pour le HIS où les données de cette classe représentent 18% du corpus. Les SVMs par exemple ont du mal à trouver des exemples cohérents sans exclure d'autres exemples d'autres classes, mais trop similaires. Les CRFs eux arrivent à capturer des dépendances complexes dans le vecteur d'attribut pour faire une meilleure classification de cette classe.

Même si les deux ensembles de données considérés sont comparables et de même nature, ils n'ont ni été acquis avec les mêmes types de participants ni dans le même habitat. La plus grosse différence réside dans la quantité d'information disponible dans chacun des habitats. Le HIS est bien moins informatif (beaucoup moins de capteurs) que Sweet-home. Ceci explique pourquoi les modèles non séquentiels ont été meilleurs sur Sweet-Home en profitant de ces informations. Dans le HIS au contraire vu le faible nombre d'informations les modèles non séquentiels ont plus de mal que les modèles séquentiels à construire des classifieurs cohérents et performants. Pour le HIS, les modèles séquentiels ont bénéficié de l'historique dans le cas de HMM et de CRF et pour MLN il a bénéficié de sa forte expressivité.

Pour la classe particulière *Téléphone*, MLN a été capable de comprendre que lorsque quelqu'un parle beaucoup dans une pièce où il y a le téléphone, alors il est très probable qu'il/elle soit en train d'effectuer cette activité (en confirmant l'hypothèse d'une seule personne dans l'habitat et donc peu d'exemple avec beaucoup de paroles sans cette activité en cours). Malgré les relativement basses performances de MLN, ce modèle semble être un bon candidat pour représenter également des activités de la vie quotidienne avec des objets (iADL (LAWTON et al., 1969)), car il est capable d'exprimer une relation sémantique complexe que les modèles purement probabilistes ne peuvent considérer.

Lorsque nous analysons les attributs retenus par la sélection, nous nous apercevons que les attributs les plus significatifs sont la localisation et le nombre de sons/paroles. Cela peut expliquer l'importance du son dans la décision prise au final pour la classification. Même si les activités étaient principalement localisées dans un nombre restreint de pièces pour chacune, afin de retirer toute ambiguïté le son était souvent utilisé. L'agitation totale, qui était dépendante également du nombre de sons, était très informative pour différencier par exemple repas et nettoyage. Dans ce cas, l'agitation produite par les capteurs de présences et les contacteurs de porte était similaire, le son était donc le moyen de différencier les activités. Idem pour lecture et téléphone toutes deux réalisées dans le bureau. De plus, dans le modèle MLN, le poids des règles portant sur des attributs de sons était en général élevé.

4.5.5 Conclusion sur ces travaux

Ces travaux avaient pour but d'approfondir les connaissances de l'état de l'art sur la reconnaissance d'activités humaines par la classification avec différents modèles. L'objectif était de faire un point sur les différentes méthodes et de proposer une nouvelle méthodologie complète pour faire cette reconnaissance, de l'acquisition d'information à la classification. Ces travaux montrent la difficulté de généraliser ce type d'approches dans des conditions quasi réelles en ayant toujours de bonnes performances de reconnaissance. Il en ressort que l'habitat utilisé a une grande importance et que le type de modélisation par contre n'est pas si influent que cela (si l'on a les données suffisantes pour correctement apprendre).

4.6 Utilisation de la reconnaissance d'activités et création de profils dans les habitats

La dernière partie de ce chapitre concerne les automatismes dans les habitats équipés de capteurs et l'adaptation de ces automatismes à l'utilisateur. Les travaux précédemment présentés visaient à obtenir une indication sur l'activité en cours effectuée par la personne avec pour but de vérifier l'évolution de cette personne (par exemple âgée) et de son autonomie dans sa vie quotidienne. En collaboration avec le Dr. Abir-Béatrice Karami-Lozenguez, nous avons aussi travaillé sur un autre aspect, celui de la création de profils utilisateurs afin d'adapter le comportement de l'habitat à l'utilisateur. Dans cette partie, la reconnaissance d'activités sert à prendre une décision à un moment donné.

4.6.1 Problématique

Les habitats intelligents ont vu ces dernières années des utilisations en terme de santé, de sécurité, d'économie d'énergie ou de bien être (comme nous l'avons décrit dans (GHAFFARIANHOSEINI et al., 2016)). Cependant, dès ses débuts, l'intérêt visé par ces technologies était l'amélioration du confort de la personne. VISCHER (2007) définit le confort à l'aide d'une pyramide à trois niveaux. Le premier niveau, la base du triangle, est le confort « physique ». C'est la base pour l'habitabilité. Pour combler ce besoin, il faut prendre en compte uniquement les normes de bonne réalisation des bâtiments. Le second niveau est le niveau fonctionnel, ou comment l'installation va aider l'utilisateur à vivre sa vie. Enfin, au plus haut de la pyramide, nous trouvons le niveau psychologique. ALLAMEH et al. (2011) présente ce dernier niveau comme étant relatif aux besoins humains et à son style de vie. Ce confort est plus relatif aux conditions d'intégration de la technologie en étant compatible avec la vie de tous les jours de la personne, ici le fait d'imposer la technologie est remplacé par une approche centrée utilisateur. Ces auteurs revisitent le travail fait autour des habitats intelligents avec le spectre de cette pyramide en voyant ce que la technologie peut apporter à l'utilisateur. Le but final de notre travail est de s'intéresser à ce niveau afin d'accroître l'expérience utilisateur de la personne dans son habitat. Il faut aussi prendre en compte l'aspect simplicité et utilisabilité des technologies. WILSON et al. (2014) reviennent sur la littérature à travers trois thématiques relatives aux habitats intelligents. Le premier est l'aspect fonctionnel, instrumental et socio-technologique. Le second est le fait que l'utilisateur auquel s'adresse la technologie soit capable d'interagir avec la technologie intégrée à l'habitat, et enfin le dernier est l'aspect qui s'occupe du matériel et du logiciel gérant l'automatisation. Sur le second point, les auteurs, à travers leur analyse, font ressortir une absence de recherche centrée sur l'utilisateur : les caractéristiques des utilisateurs potentiels, comment ils vont interagir avec la technologie et comment la conceptualisation de l'habitat et de ses technologies peut affecter l'utilisabilité et l'acceptabilité.
Ces travaux sont de premiers résultats, publiés dans (KARAMI et FLEURY, 2016; KARAMI, FLEURY et al., 2016) et qui seront ici succinctement présentés. Afin de répondre aux défis identifiés, ce travail se concentre sur les habitats intelligents « adaptatifs » et sur la manière dont les personnes pourront interagir avec l'habitat pour en changer le comportement.

4.6.2 Architecture globale de la solution proposée

La figure 4.6 présente les différents composants de l'architecture proposée et montre les flux de données de l'acquisition à la prise de décision. L'architecture repose sur cinq bases de données et trois processus. La première base, contenant les données brutes, recueille les informations de l'habitat (lecture des données capteurs). Elle contient aussi les informations sur les utilisateurs (profil). Ce profil peut être rempli pour chaque utilisateur et complété/modifié par le système automatiquement.
Le processus de traitement de données est en charge de :

1. Reconnaissance d'activités (cf. la section précédente)
2. Détection des retours implicites et exportation des préférences venant de l'analyse des retours explicites
3. Mise à jour des profils utilisateurs en utilisant les nouvelles habitudes détectées.

Le processus de reconnaissance d'activités, comme dans la partie précédente, reconnaît les activités en cours dans l'habitat. Les données brutes sont aussi analysées pour détecter les préférences sous forme d'un triplet (situation, action, retour). Une préférence potentielle est retenue lorsque :

1. Une action automatisée de l'habitat suivie d'un retour explicite est présente (exemple, la personne a éteint la lumière qui avait tout juste été allumée par l'automatisation).
2. Un retour implicite est détecté (une interface utilisateur permet de donner son avis sur une action effectuée).

Les processus décrits vont se charger de remplir deux bases de données nommées « préférences potentielles » et « activités détectées ». Ces bases de données vont être les entrées pour les algorithmes de traitement de données et les processus d'apprentissage en charge de mettre à jour un processus de récompense basé sur des Processus de Markov. Pendant ce processus d'apprentissage, la fonction récompense est généralisée pour couvrir des situations inconnues et de nouveaux utilisateurs en déterminant les attributs clés pour chaque action possible. Le dernier processus est

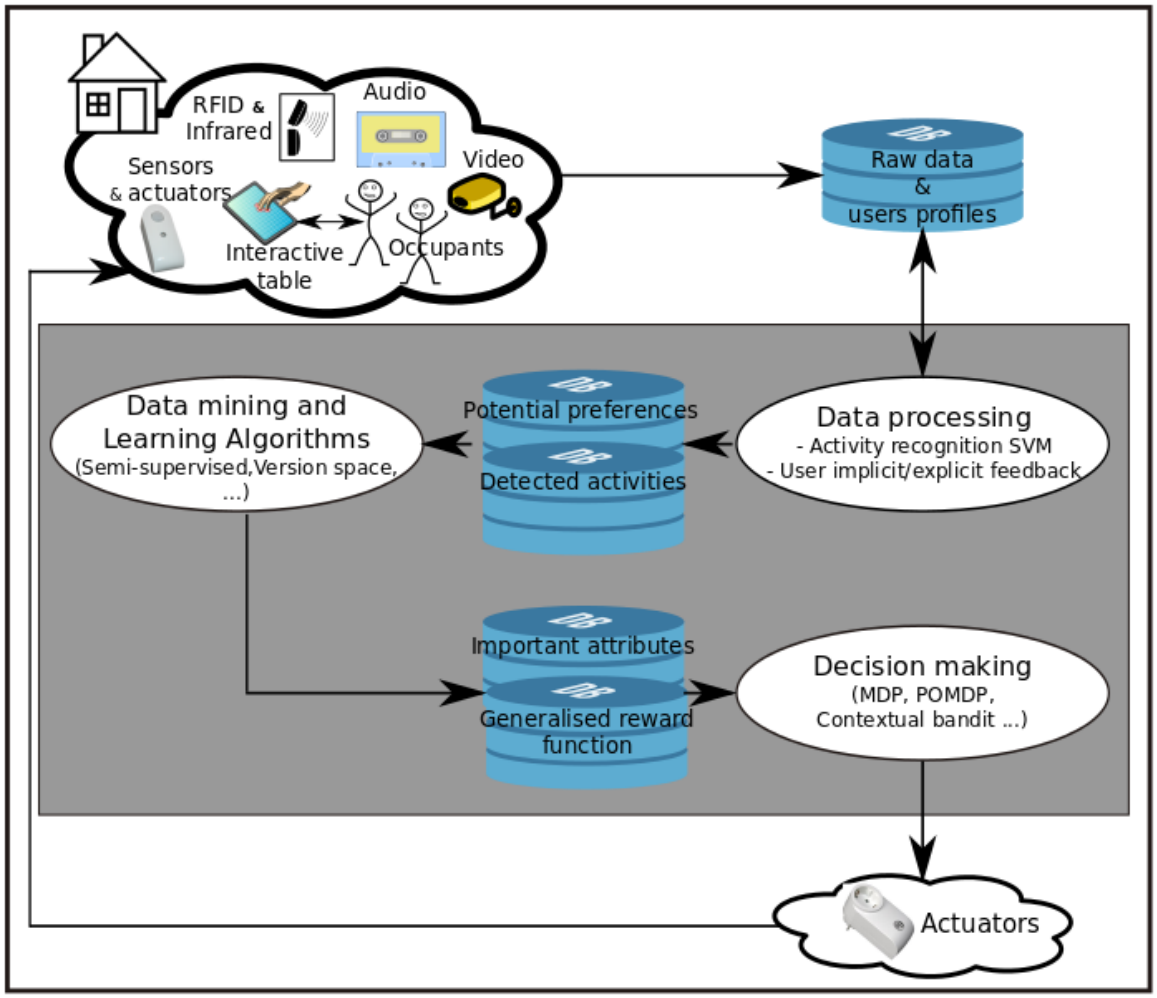


FIGURE 4.6 – Architecture globale du système proposé

en charge de décider l'action et contrôler les actionneurs pour modifier l'état actuel de l'environnement en utilisant les actions choisies par le processus de Markov. Les composants et leurs fonctionnalités sont décrits en détail dans la prochaine partie.

Les préférences potentielles sont sauvegardées dans une base de données et utilisées pour apprendre ou mettre à jour les fonctions de récompenses généralisées (règles adaptatives généralisées). Les méthodes de généralisation utilisées sont basées sur la détection d'attributs importants pour chacune des décisions. Le processus de prise de décision utilise la fonction de récompense généralisée pour prévoir les décisions optimales adaptées à la situation courante (situation courante en termes de profil utilisateur et d'état courant des capteurs dans l'habitat). Par exemple, si l'on considère le cas de l'allumage d'une lumière. Une telle action dépend du niveau de luminosité dans la pièce (l'attribut important qui sera détecté dans le cas de cette action). Cependant, une baisse de luminosité dans une pièce peut être détectée lorsque la personne est en train de dormir ou bien en train de lire. Dans un cas, allumer la lumière n'est pas une action adaptée à prendre, alors que dans le second cela peut être très utile à la personne. Dans les deux cas, les préférences apprises de l'utilisateur vont nous apprendre le fait que celui-ci voudrait, dans le cas de faible luminosité, allumer la lumière. Le couplage avec le système de reconnaissance d'activité va empêcher l'action de se faire dans l'un des deux cas. Le système de reconnaissance d'activités de la partie précédente sera utilisé ici comme une information supplémentaire dans le processus de décision.

4.6.3 Plateforme de test Eco-Confort d'IMT Lille Douai et retours utilisateurs

La figure 4.7 montre une vue générale de l'installation disponible dans le local Eco-Confort d'IMT Lille Douai, un living lab avec une double pièce et qui est équipé pour être une salle de détente/repos. Les éléments sont les suivants :

— Des prises électriques instrumentées dans les deux pièces,

— Des capteurs de luminosité, humidité et température à différentes positions dans l'habitat et en dehors de celui-ci,

— Des capteurs pour connaître la position des volets, ainsi que les ouvertures/fermetures des portes et fenêtres

— Des capteurs infrarouges, positionnés à des endroits stratégiques des deux pièces, pour savoir où la personne se situe,

— Des actionneurs pour commander les différents radiateurs,

— Des actionneurs pour commander les lumières,

— Une commande disponible pour activer/désactiver chacune des prises électriques,

— Des actionneurs pour commander la position des volets,

— Un lecteur de carte permettant d'identifier toute personne entrant dans l'habitat afin de la reconnaître et d'associer son profil s'il existe déjà.

Pour les besoins des expérimentations, certains mobiliers ont été installés :

— Une table et des chaises permettant de s'installer confortablement,

— Une grande télévision intelligente avec accès à différents médias (musique, TV, radio, chaines Youtube, Internet, etc.).

— Un Babyfoot afin de jouer (à plusieurs),

— Une table tactile sous Microsoft Windows qui permet aux utilisateurs de naviguer sur internet, lire, etc., mais aussi de donner des retours et de voir l'état de l'habitat.

La figure 4.8 illustre cette installation de capteurs dans l'habitat avec des photographies.

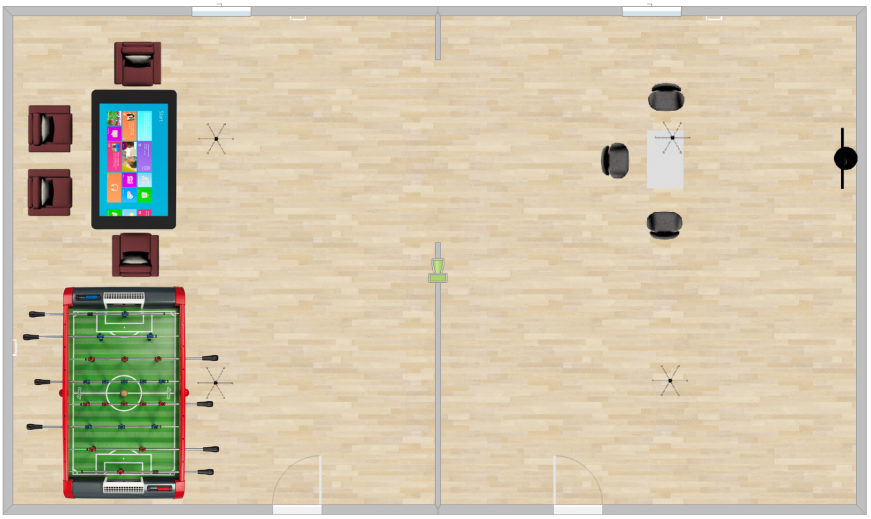


FIGURE 4.7 – Local Eco-Confort, IMT Lille Douai.

La table interactive (voir figure 4.9) possède une interface web (développée dans le cadre de projets étudiants à IMT Lille Douai) et qui permet à la personne de réaliser les actions suivantes :

— Vérifier l'état des différents capteurs des deux pièces et obtenir ainsi des informations sur l'environnement de vie,

— Agir sur les actuateurs de la pièce comme par exemple allumer ou éteindre une lumière, démarrer une prise électrique, etc.

— Donner une opinion sur une action qui vient juste d'avoir lieu dans la pièce (en fait sur la dernière action prise par le système). Par exemple, si l'habitat allume une lumière suite à la détection d'une luminosité trop faible et que l'utilisateur n'est pas d'accord avec cette décision, il pourra le signifier en venant appuyer sur la table sur un pouce vers le bas (et pouce vers le haut pour une approbation).

FIGURE 4.8 – Illustration de l'installation.

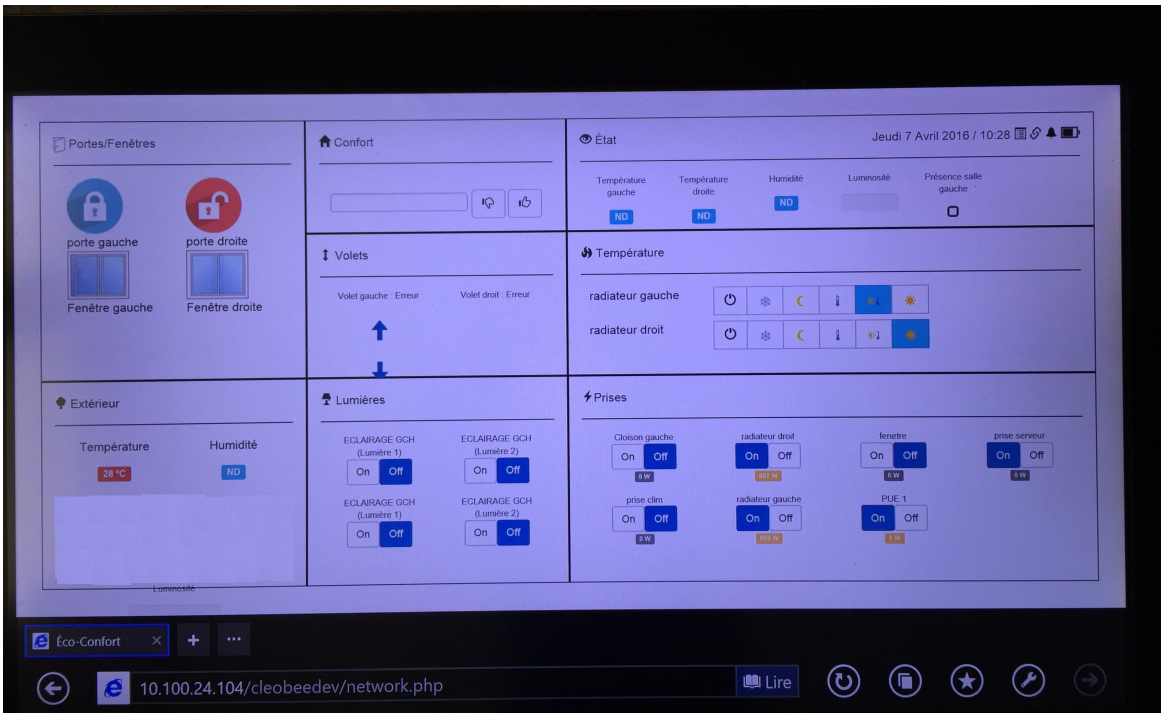


FIGURE 4.9 – Table interactive et application web permettant de gérer l'habitat.

Toutes les actions réalisées à travers la table sont enregistrées comme des éléments dans les bases de données gérant l'habitat. Une action sur un actionneur de la pièce par l'intermédiaire de la table est considérée comme un retour explicite alors qu'une opinion exprimée par les boutons de satisfaction (pouce levé ou baissé) est un retour implicite. Au contraire des retours implicites, les retours explicites (exemple allumer une lumière juste quelques secondes après qu'elle ait été éteinte par l'habitat) sont plus faciles à connecter à l'action qui les a causés (même actionneur). Lorsqu'un tel retour est reçu, une récompense négative est associée à l'action faite par l'habitat avec l'état courant complet. Lorsqu'un retour implicite intervient, il n'est pas facile de l'associer à l'action.

4.6.4 Suite des expérimentations

Des expérimentations avec des volontaires doivent maintenant être lancées dans le living lab du laboratoire. L'objectif de ces expérimentations sera de tester la capacité de notre architecture à utiliser les retours de l'utilisateur et à apprendre leurs préférences.

Les articles (KARAMI et FLEURY, 2016; KARAMI, FLEURY et al., 2016) présentent actuellement des résultats sur la convergence de modèles qui seront utilisés par la suite dans ces travaux. Des données de simulation avec des actions tirées au hasard ont prouvé la convergence des politiques de décisions testées, il faut maintenant mettre tout ceci en œuvre, conformément à l'architecture donnée précédemment, dans le cadre de l'habitat installé à IMT Lille Douai.

Dans une première expérimentation, nous nous concentrerons sur le bien-être et la satisfaction de l'utilisateur au

regard de l'éclairage de la pièce ainsi que de la température de la pièce, la position du volet (ouvert/fermé) et l'utilisation des médias (TV et radio) en utilisant la liste des capteurs décrite précédemment. Les retours utilisateurs peuvent avoir lieu avec une reconnaissance vocale, l'application web sur la table interactive ou encore les actions de l'utilisateur sur les actionneurs ou sur la table). Une première phase d'expérimentation va générer des données étiquetées (le système enregistre tout en continu depuis un long moment) dans lesquelles l'habitat sera principalement passif afin d'observer les retours explicites de l'utilisateur. Ce sera aussi le moment nous pourrons créer des modèles d'activités adaptés à notre environnement (et reconnaissant un certain nombre d'activités que nous aurons définies comme compatibles avec notre installation).
Dans une seconde phase, les préférences potentielles obtenues à partir des données collectées dans la première phase seront utilisées pour apprendre un modèle générique de fonction de récompense (avec GVS) et l'habitat pourra ainsi devenir plus actif tout en observant toujours les retours implicites comme explicites de l'utilisateur. Les utilisateurs, dans la seconde phase, devront remplir des questionnaires de satisfaction basée sur une échelle type Likert (LIKERT, 1932) qui analyse la satisfaction d'une personne par rapport à l'automatisation du système et par rapport à l'interaction qu'il a avec.

4.7 Conclusions

Ces travaux visaient à montrer les capacités de reconnaissance des algorithmes séquentiels et non séquentiels dans le cas des activités de la vie quotidienne dans un habitat intelligent. Pour cela, une étude sur les attributs et leur sélection a d'abord été réalisée, puis la classification par plusieurs méthodes (CRF, MLN et HMM en séquentiels, SVM et RF en non-séquentiels). Cette étude a montré que les algorithmes séquentiels sont souvent meilleurs que les autres, cependant la différence n'est pas non plus significative.
Une étude plus poussée a montré l'influence sur le classifieur d'inclure la classe *«Autre»* (et donc la diminution des performances engendrée) qui est une classe contenant tout ce qui n'est pas étiqueté dans l'ensemble d'apprentissage. Cette étude est très importante car cette classe est souvent majoritaire dans les ensembles de données qui permettent de construire les modèles d'activités. De plus, étant donné que par la suite nous sommes en train d'utiliser des algorithmes d'apprentissage incrémentaux pour ces ensembles de données, lors de l'apprentissage et de la mise à jour de ces modèles cette classe pourra également être présente.
La suite de ces travaux a été de proposer une architecture pour un habitat intelligent adapté aux besoins des occupants prenant en compte cette donnée d'activité et capable de proposer un réglage plus fin et adapté à la personne des automatismes de son habitat, dépendant d'une part de son activité, mais aussi de ses souhaits. Ainsi, la méthodologie propose de reprendre les retours de la personne, qu'ils soient implicites ou explicites. En effet, la personne peut soit proposer de noter en bien ou en mal l'action qui vient d'être réalisée, soit faire une action contraire par la suite montrant ainsi son mécontentement. Dans la publication tirée de ces premières réflexions méthodologiques, des données de simulations ont été créées afin de s'assurer de la convergence de tels modèles.
Les deux suites proposées de ces travaux sont en cours et devraient aboutir à de nouveaux résultats. En collaboration avec l'Université d'Alger, des travaux sur l'une des bases de données de CASAS (d'une université américaine) sont en cours pour proposer de nouveaux algorithmes de classification incrémental pour ces données. Ces algorithmes se basent sur des critères de mise à jour des modèles peu fréquents avec une première étape de clustering. La base de donnée considérée étant très importante (en taille), l'utilisation de celle-ci nous montre les limites des modèles et nous permet d'améliorer les algorithmes proposés afin que ceux-ci puissent un jour être utilisés en situation réelle. Les seconds travaux en cours exploitent la plateforme Eco-confort afin de proposer des modèles de profil pour des utilisateurs multiples dans un habitat intelligent en se basant sur leurs retours utilisateur. De plus amples expérimentations doivent se faire afin de s'assurer de la bonne sélection des modèles et également de vérifier l'influence des erreurs que nous pourrions avoir sur la première étape de classification d'activités.
Le chapitre suivant, concluant ces trois quatre premiers chapitres va revenir sur ces résultats avant de conclure sur les perspectives qui s'offrent suite à ces travaux.

CHAPITRE 5

Conclusions et perspectives des travaux de recherche

Sommaire

4.1 Introduction et verrous scientifiques à lever . . . 59

4.2 État de l'art sur l'extraction d'informations dans les habitats intelligents . . . 61

4.2.1 Segmentation des données . . . 61

4.2.2 Extraction des caractéristiques par fenêtrage . . . 62

4.3 Ensemble de données utilisés dans ce chapitre . . . 63

4.3.1 HIS . . . 64

4.3.2 Domus . . . 65

4.4 Sélection d'attributs dans les signaux issus des environnements intelligents . . . 67

4.4.1 Descripteurs . . . 67

4.4.2 Sélection des attributs . . . 67

4.4.3 Temps du fenêtrage et son influence . . . 69

4.4.4 Conclusions . . . 69

4.5 Un aperçu de la classification d'activités dans les habitats intelligents avec des méthodes en ligne . . . 70

4.5.1 Introduction . . . 70

4.5.2 Différentes méthodes de modélisation d'activités . . . 71

4.5.3 Expérimentations . . . 72

4.5.4 Discussion . . . 76

4.5.5 Conclusion sur ces travaux . . . 78

4.6 Utilisation de la reconnaissance d'activités et création de profils dans les habitats . . . 78

4.6.1 Problématique . . . 79

4.6.2 Architecture globale de la solution proposée . . . 79

4.6.3 Plateforme de test Eco-Confort d'IMT Lille Douai et retours utilisateurs . . . 80

4.6.4 Suite des expérimentations . . . 82

4.7 Conclusions . . . 83

Les travaux présentés dans les trois précédents chapitres couvrent des activités de recherche qui portent sur les différentes étapes de la chaîne du traitement des données, à savoir l'extraction et la mise en forme de caractéristiques du signal, puis leur interprétation et leur classification, sans oublier une dernière phase de prise de décision. Tous ces travaux s'inscrivent dans un champ de recherche spécifique lié à la compréhension de l'humain et de son comportement.

Par l'intermédiaire de futurs directions de thèses ainsi que par le dépôt de projets de recherche permettant d'encadrer des étudiants de renforcer les moyens consacrés au développement de ma thématique. Les prochaines années seront donc dévouées à approfondir ces travaux tout en visant la même finalité : **améliorer les méthodes et techniques de traitement du signal et de traitement de la donnée pour une amélioration des conditions de vie d'une personne, que ce soit en termes de santé connectée, de sécurité ou bien de confort**.

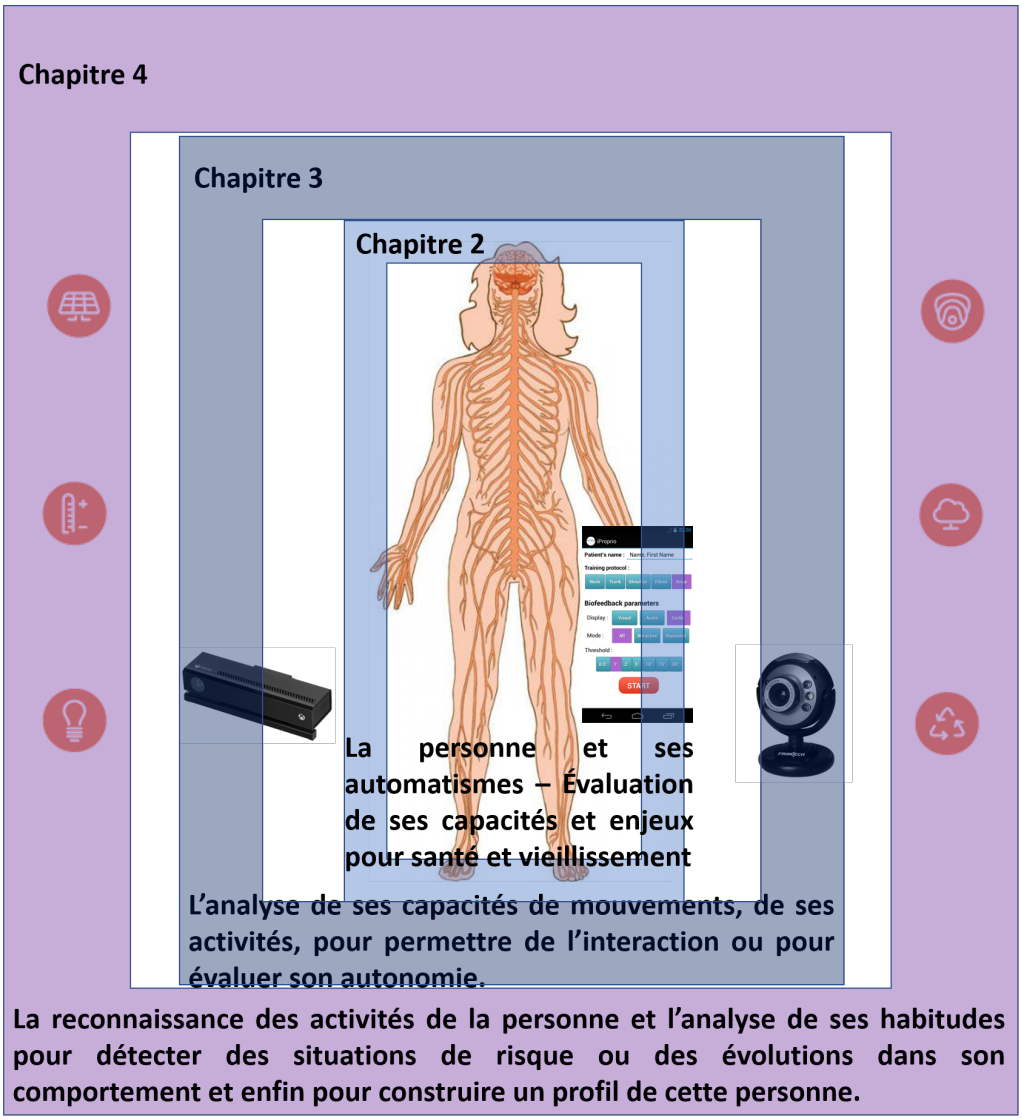


FIGURE 5.1 – Résumé synthétique des thématiques présentées dans chacun des trois chapitres de ce manuscrit, dans l'ordre croissant du plus intérieur au plus extérieur.

5.1 Conclusions des travaux présentés

La figure 5.1 présentée dans l'introduction de ce manuscrit (et ci-après rappelée) illustrait la structuration des travaux réalisés, notamment sous la forme des trois chapitres principaux de ce manuscrit, de l'automatisme propre à la personne à ses interactions avec son environnement.

Tout d'abord, nous nous sommes concentrés sur la compréhension de certains mécanismes inhérents à l'humain et à ses automatismes. Ceci a amené des travaux relatifs au traitement des signaux qu'ils soient physiologiques (le rythme cardiaque et la respiration) ou cinématiques/sensoriels (la posture donnée par une centrale inertielle ou par une plateforme de force), avec des méthodes non linéaires afin d'en extraire des informations sur leur contenu et d'en déduire une relation entre ce contenu et les automatismes présent dans le corps humain, et surtout le bon fonctionnement de ces automatismes (FRANCO, GUMÉRY et al., 2012; FRANCO, GUMERY et al., 2014; FRANCO, FLEURY, DIOT et al., 2018). Des critères ressortent ainsi sur les évolutions des valeurs des indices créés en fonction de la bonne gestion de l'équilibre par exemple. La suite des travaux, avec cette fois-ci des centrales inertielles, a porté sur l'utilisation du smartphone pour gérer l'exécution d'exercices précis pour la personne et l'aider dans le maintien de sa forme ou dans sa rééducation (FRANCO, FLEURY, GUMÉRY et al., 2013; MOURCOU, FLEURY, DIOT, FRANCO et al., 2015; MOURCOU, FLEURY, FRANCO et al., 2015).

Le chapitre suivant a porté sur l'analyse du comportement de la personne dans sa globalité et la reconnaissance, à

partir de la vidéo, de la personne en elle-même ou de ses actions, dans un but d'interaction. Les premiers travaux ont porté sur la reconnaissance de la personne par des algorithmes de classification en ligne, avec une partie sélection de paramètres et une autre sur la classification incrémentale à partir de données incomplètes (Y. LU, BOUKHAROUBA et al., 2014; Y. LU, FLEURY et al., 2015). Cela nous a amené à trouver de nouveaux descripteurs à partir des données vidéos dont nous disposons, puis à construire des classifieurs efficaces avec ces données. Le résultat principal de cela est un descripteur et une méthode de classification permettant de reconnaitre efficacement et rapidement (avant la fin) une action de la personne, quelle que soit son orientation par rapport à la caméra (TALHA, HAMMOUCHE et al., 2018). Cependant, pour arriver à ce résultat, différents travaux et différents types de classifieurs ont été proposés (HAMMOUCHE et al., 2016; TALHA, FLEURY et AMBELLOUIS, 2017).

Le dernier chapitre avant cette conclusion a porté sur les habitats intelligents. Dans ces habitats, nous avons pour but premier de reconnaître les activités de plus haut niveau faites par la personne, afin de mieux connaître son quotidien. Ce chapitre décrit un ensemble de méthodologies et de résultats permettant de faire une reconnaissance (avec des algorithmes séquentiels ou non, avec une connaissance partielle ou non des activités etc.), en analysant également les principaux problèmes liés à ces types de travaux (déséquilibre des classes, données incomplètes ou erreurs de capteurs, faible nombre de données indexées, etc.) (CHAHUARA, FLEURY, PORTET et al., 2016). Cette reconnaissance repose également en grande partie sur la capacité de segmenter les différentes activités dans le signal. Le but de cette reconnaissance d'activité est de proposer des indicateurs de qualité de vie et d'autonomie en se rapprochant des échelles utilisées par les gériatres et se basant sur les activités de la vie quotidienne. La fin de ce chapitre décrit succinctement les travaux commencés sur la création de profils utilisateurs dans des habitats, utilisant ces activités pour savoir dans quel contexte se trouve la personne et faire agir l'automatisation de l'habitat en conséquence (KARAMI, FLEURY et al., 2016; KARAMI et FLEURY, 2016).

À partir de ces travaux, je vais dresser par la suite des perspectives de ces recherches à court et moyen terme.

5.2 Perspectives de recherche

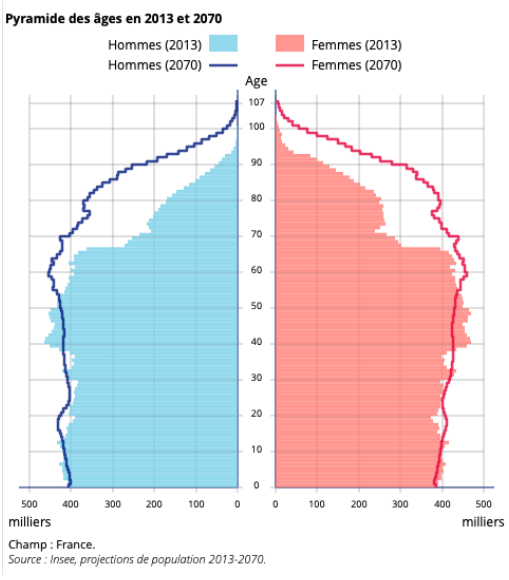


FIGURE 5.2 – Prévision d'évolution de démographie (source INSEE) en France

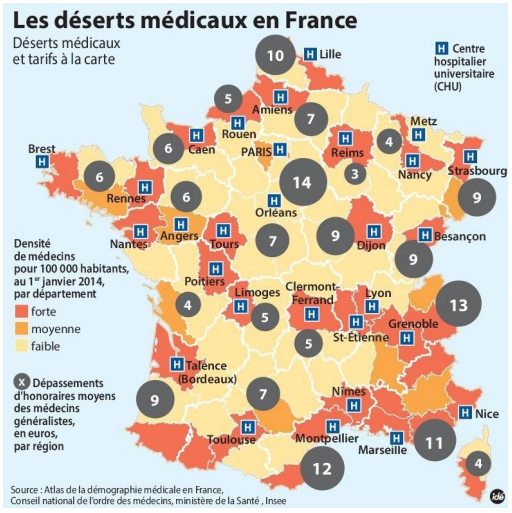


FIGURE 5.3 – Les déserts médicaux et leur répartition dans le pays

Les progrès de la médecine aidant, nous le savons depuis longtemps, les démographies de tous les pays développés évoluent fortement et les populations vieillissent. Dans son dernier rapport du 3 novembre 2016 [1], l'INSEE confirme que d'après ses projections, il y aura bien un vieillissement important de la population (cf. figure 5.2) dans les années à venir, et prévoit que les plus de 75 ans auront doublé à l'horizon 2070 par rapport à 2013.

1. https ://www.insee.fr/fr/statistiques/2496228?sommaire=2496793

Or, nous le savons, cette augmentation inclut aussi une recrudescence des maladies chroniques ou des maladies liées à l'âge. Qui dit augmentation de ces maladies dit aussi augmentation des personnes ne pouvant plus vivre au domicile de manière sécurisée. Le manque de place en institution est déjà à l'heure actuelle un problème de société et la situation ne devrait pas s'arranger. De plus, la situation est également difficile en termes d'offre de santé et de nombreuses zones (y compris les grandes villes) manquent cruellement de médecins. Comme le montrent des enquêtes récentes, les Hauts-de-France ne font pas exception et sont même dans les premiers concernés. La réforme des études de santé, si elle a l'effet escompté, ne le produira que dans de nombreuses années.

Tout cela évoque la pertinence et l'intérêt des travaux que nous effectuons sur l'analyse de comportement au domicile ou sur la création de solutions permettant d'aider les praticiens dans leur quotidien (cf. sections sur iBalance/iProprio) en permettant au patient de compléter la prise en charge au domicile. Dans le cadre de projets actuels (par exemple le projet Stic-AmSud) nous réfléchissons également sur l'utilisation des jeux sérieux dans le cadre de la rééducation de la personne, de la prévention des maladies chroniques ou pour aider à avoir une vie meilleure avec ces maladies. Ces travaux ne viseront jamais bien sûr le remplacement du personnel médical (même si certaines projections envisagent le contraire en surestimant le rôle de l'IA), mais complétera les actions de celui-ci et lui permettra de gérer plus simplement les patients.

Pour ceci, le principal prochain verrou scientifique à surmonter est de trouver une méthodologie efficace pour identifier au mieux les variables pertinentes et donnant les meilleurs indications permettant d'objectiver les grilles de scores connues des praticiens. En effet, pour que les technologies développées soient le plus efficace et le mieux diffusé, il est important de les rapprocher des pratiques quotidiennes des spécialistes et donc de leurs grilles d'évaluation. Cette méthodologie est pour l'instant empirique et fortement dépendante de l'application spécifique (lorsque nous trouvons), généraliser ceci serait un plus pour l'avancement des connaissances en interprétation automatique des données. Une sous-partie de ces travaux va continuer de faire avancer les connaissances sur les algorithmes de classification incrémentaux et en ligne afin de mieux en connaître les conditions aux limites et les conditions de convergence dans le cas de grands ensembles de données. Ce verrou scientifique est peu adressé pour l'instant car dans le cas de grandes bases de données, dans les pratiques actuelles, l'utilisation d'une base d'apprentissage suffisante se fait et les modèles n'évoluent pas.

L'une des perspectives ouverte par la fusion de deux entités pour donner IMT Lille Douai et le rapprochement avec certaines personnes du laboratoire CRIStAL qui ont des thématiques proches, est de confronter nos deux modèles d'analyse de données de mouvement (centrale inertielles vs. caméras de profondeur) afin d'en déduire de plus amples informations sur la personne et sur de possibles signes qui pourraient être manqués par l'une ou l'autre des méthodes. Cette analyse de motricité globale pourrait être appliquée à long terme tant à la personne âgée ou en possibilité d'entrée en dépendance qu'à la petite enfance pour déterminer des troubles du développement.

Enfin, cela n'a pas été décrit dans ce manuscrit, mais mes travaux de recherche actuels m'ont ouvert la porte de deux autres domaines applicatifs, toujours liés à l'humain, que sont la cyber-sécurité (VILLAIN et al., 2019) et l'apprentissage humain (VERMEULEN et al., 2015) (apprentissage au sens pédagogique du terme). Dans les deux cas, des projets en cours permettent d'étendre ce qui est fait actuellement à d'autres domaines et de confronter la méthodologie à ces nouvelles applications afin d'en connaître les limites de généralisation. Des dépôts de projets pourraient ainsi donner, dans les années à venir, des encadrements de thèses dans ces deux domaines. Le premier viserait à analyser finement les échanges sur un réseau afin de déterminer les modifications d'utilisation de celui-ci (de la même manière que nous détectons des changements d'habitude d'une personne) afin de détecter des attaques. Le second domaine se baserait sur l'analyse de traces d'utilisateurs dans le cadre d'enseignement pour vérifier à quel point une méthode pédagogique est efficace ou peut être améliorée.

5.3 Développements des recherches à court terme

Après des réflexions à moyen terme amenant à décrire des perspectives de recherche dans plusieurs années, cette section va tenter de donner quelques pistes de recherches dans la continuité immédiate des travaux menés, avec les résultats qui pourraient en être attendus.

5.3.1 Le projet iBalance/iProprio

Comme décrit dans la section 2.6 page 20, ce projet, relativement avancé et en discussion avec la SATT Nord pour une valorisation de la propriété intellectuelle, a encore quelques verrous scientifiques qui doivent être levés :

— Ajouter une gestion d'exercices liés à la démarche de la personne. Le verrou scientifique ici est l'implémentation et l'adaptation à la plateforme des algorithmes ainsi que la sélection de variables pertinentes pour décrire sur le long terme l'évolution de la personne dans ces exercices. Nous avons déjà proposé un ajout, dans l'application, d'exercices liés à la démarche dans lesquels une personne va se voir donner des consignes sur la pose des pieds avec des temps différents entre les deux pieds pour les étapes de marche. Ceci sert dans le cas d'une personne avec une marche très fortement déséquilibrée (temps sur les appuis très différent). Une amélioration sur ce point offrirait une meilleure gestion de la qualité de la posture et de l'équilibre de la personne dans le cadre de la démarche et ainsi améliorerait l'évaluation que nous sommes capables d'effectuer sur la personne au cours du temps pour évaluer son évolution.

— Améliorer fortement l'interface graphique, l'ergonomie et l'utilisabilité du prototype. Ce que nous proposons n'est qu'un prototype de recherche qui devra être amélioré pour répondre au mieux aux demandes des différentes personnes qui y auront accès. Le verrou scientifique ici est de savoir comment présenter et que présenter comme résultat avec la personne pour la faire évoluer et la motiver.

— Faire de cette application un vrai jeu sérieux afin de le proposer en préventif et d'utiliser les résultats discutés précédemment. Nous avons toujours pensé ce projet comme devant se présenter sous la forme d'un jeu sérieux afin d'être ludique et utilisé en prévention au cours de la vie. Pour ce faire, doit-on ajouter des capteurs ou peut-on rester avec un point de mesure unique ? Les jeux sérieux sont un domaine porteur et qui permettra aussi de prévenir des risques dans les années à venir, il semble important de finaliser la direction dans cette voie qui a déjà commencé à être réfléchie. Le verrou scientifique ici est le modèle d'adaptation de la demande des professionnels en terme d'exercices à la notion de jeu et à l'amusement de la personne.

Une thèse, en co-encadrement avec des spécialistes de l'ergonomie, pourrait alors être proposée sur le sujet *Proposition d'un jeu sérieux permettant à une personne de travailler/maintenir/optimiser sa démarche, ses mouvements ou sa posture. Application au cas de la personne âgée ou ayant subit un traumatisme engendrant un handicap.*

5.3.2 Profils utilisateurs – jeu sérieux dans les habitats

Les travaux sur la création de profils utilisateurs ont été commencés avec Dr. Abir-Béatrice Karami-Lozenguez qui a apporté sa compétence en modèles d'intelligence artificielle pour donner un niveau d'abstraction supplémentaire aux données que nous avons (données brutes de l'habitat et données de plus haut niveau comme celles issues de la reconnaissance d'activités). Ces travaux doivent cependant être continués et les objectifs à court terme sur ceux-ci sont les suivants :

— Perfectionner et pérenniser l'installation des capteurs de la plateforme Éco-Confort d'IMT Lille Douai. Afin de réaliser les travaux entrepris sur l'adaptation de l'habitat au profil de son occupant, il est important de perfectionner la plateforme Éco-Confort de l'IMT Lille Douai afin de la rendre pérenne et de sécuriser et fiabiliser l'acquisition aux données. Une réflexion de fond sur ce sujet permettra de conserver cette plateforme à jour avec les personnels présents et mobilisables sur le sujet.

— Finaliser les travaux sur les retours implicites et explicites des utilisateurs puis inclure le modèle de décision proposé. Les parties sur les retours utilisateurs dans cet habitat ont été prévues, mais doivent être maintenant finalisées et vérifiées minutieusement pour en contrôler le bon respect des règles de détection données dans les travaux déjà réalisés.

— Rendre indépendant la reconnaissance d'activités du scénario de collecte ainsi que de l'installation. Il parait important de travailler sur les algorithmes de sélection d'attributs et de classification afin que les résultats puissent devenir transposables d'un habitat à un autre, que les habitats diffèrent par leur disposition, leur nombre et diversité des capteurs ou les deux. Les résultats présentés tendent déjà à montrer une indépendance

de la personne effectuant l'activité mais les deux conditions citées précédemment, elles, impliquent une forte dépendance sur les résultats.

— Expérimenter en conditions réelles. Des expérimentations en conditions réelles doivent être menées pour vérifier les modèles que nous avons ainsi que leur pertinence dans des conditions où plusieurs utilisateurs utilisent soit en même temps soit séparément cette salle de pause.

Le verrou scientifique ici est la méthodologie complète pour correctement associer les données aux actions et aux personnes pour la création de ces profils. La recherche actuelle sur les habitats intelligents manque de ce point sur l'adaptabilité des algorithmes aux habitats pour être en capacité de se déployer à plus grande échelle.
Une thèse qui pourrait être lancée sur le sujet serait *Adaptation et adaptabilité dans les habitats intelligents ? Extension des modèles actuels pour proposer la création de profils multi-utilisateurs dans les habitats et adapter la domotique pour le confort et le bien-être de la personne.*
Une seconde partie sur les habitats pourrait concerner également l'inclusion de jeu sérieux dans le concept de celui-ci. L'idée d'un tel projet est d'aller plus loin dans l'utilisation des données de l'habitat intelligent et de se placer dans le contexte par exemple d'un EHPAD [2] ou d'un béguinage. Dans ce contexte, je voudrais proposer un jeu sérieux, dans lequel chaque personne âgée possède son avatar, et qui refléterait son état de santé et sa qualité de vie au travers d'un certain nombre de mesures. Le but est alors de motiver la personne elle-même à améliorer le score de son avatar, mais aussi de recréer du lien social en montrant aux autres l'avancement des autres avatars, ce qui permettrait par exemple, dans le cas où une personne a une période où elle va moins bien, de pouvoir être supporté par les autres. Cette idée peut être amenée à évoluer en fonction de discussions avec des spécialistes (gériatres) et avec des personnes spécialisées en psychologie, mais l'idée de base de proposer un jeu permettant aux personnes d'interagir entre eux et d'avoir un retour sur leur manière de vivre au quotidien semble intéressante.
Une seconde thèse sur les habitats intelligents qui pourrait être proposée aurait comme titre *SHoW Game : un jeu sérieux décrivant l'évolution des personnes vivant dans des habitats intelligents et permettant l'interaction et la motivation entre les personnes.*

5.3.3 Analyse des signaux smartphone dans la vie courante – Déduction d'indicateurs de qualité de vie ?

Pour terminer sur ce point, une thèse en cours co-financée par la région Hauts-de-France s'intéresse à l'analyse des signaux inertiels des smartphones sur le long terme, dans la vie courante. Sid Ahmed Walid Talha travaille sur ce sujet (début de troisième année de thèse en septembre 2018) et s'intéresse à la meilleure manière d'interpréter les signaux issus des centrales du smartphone, en considérant que nous n'avons que peu d'informations sur les conditions de port ou d'utilisation conjointe du smartphone. Un premier article, soumis (TALHA, FLEURY et LECOEUCHE, 2019), décrit les différentes étapes du traitement de ces données smartphone, en considérant l'état de l'art et en y ajoutant l'expertise apportée afin de les considérer au mieux, puis s'intéresse à la création de classifieurs pour reconnaître les transitions posturales et les postures de base dans ce signal, quelque soit la position du smartphone sur le corps. Ces travaux ouvrent la perspective, pour la troisième année de thèse, de travailler sur un indicateur décrivant l'évolution de la personne dans son quotidien et pourquoi pas obtenir, si la littérature nous donne des exemples validés de telles mesures, une évaluation de l'autonomie ou du bien-être de la personne ?
Ces différents travaux seraient des verrous scientifiques qui seraient adressables à court terme suite aux travaux présentés dans cette habilitation.

5.4 La place de la santé, de la donnée et des habitats intelligents dans l'IMT et l'IMT Lille Douai

La santé est l'une des thématiques phares de l'Institut Mines-Télécom depuis sa création et dans sa stratégie actuelle. Les applications en ingénierie pour la santé sont ainsi très présentes dans les différentes écoles et très diversifiées,

2. Établissement d'Hébergement pour Personnes Âgées Dépendantes

allant du traitement d'image ou du traitement du signal à des réalisations de matériaux spécifiques pour les prothèses par exemple. De par mon implication dans ce domaine, je représente déjà l'IMT Lille Douai dans les instances de l'institut liées à la santé. L'implication dans de telles tâches et continuer dans le futur est fortement intéressant car cela permet de participer au fait de faire vivre la thématique dans notre institut et dans mon établissement et de voir comment organiser au mieux la collaboration entre écoles et équipes pour améliorer la visibilité. L'organisation afférente de colloques au sein de l'institut sur cette thématique est également une perspective intéressante car elle permet d'améliorer à chaque fois nos connaissances par l'intermédiaire des personnes invitées.
Il sera intéressant qu'à terme cette thématique apparaisse distinctement dans l'école. Ainsi, la création d'un groupe thématique afin de mettre en commun les efforts faits sur les deux sites (Lille et Douai) et de répondre de manière plus unifiée et donc avec plus d'importance aux appels à projets. La santé apparait comme une thématique commune aux trois CERI (Centres d'Enseignement, Recherche et Innovation) actuellement en cours de création. Dans les années à venir, il faudra peut être dépasser les clivages des différentes thématiques et équipes pour imaginer un modèle différent afin de rassembler les chercheurs plus efficacement et de partager les avancées.
Toujours sur la nouvelle organisation de l'école et sur un rapprochement éventuel avec les UMR Lilloises, il me semblerait intéressant de m'impliquer d'avantage dans la gestion d'une thématique transversale liée au traitement de la donnée, thématique qui se retrouve de nos jours commune à tous les champs de recherche. Construire cette transversalité avec l'aide des collègues des différents centres (et principalement ceux du numérique pour en définir le contour) me semble une tâche primordiale par la suite pour faire progresser la visibilité de l'école et l'attractivité de sa recherche dans les années à venir. L'après HDR verra très certainement (et je le souhaite) une partie de mes activités liée à ce type de tâche que ce soit au niveau de l'institut ou au niveau de l'école.
Enfin, avec les travaux que nous avons déjà effectués avec les autres personnes dont les doctorants et les post-doctorants du laboratoire sur la plateforme Eco-Confort, nous voyons également qu'il faudrait être en mesure de pouvoir dédier des ressources à cette plateforme afin de pouvoir faire avancer les projets. Le défi pour faire ceci est de trouver un modèle de financement pérenne pour les-dites ressources.

5.5 Collaboration et communauté scientifique

Ces travaux ont été effectués au travers de collaborations nationales et internationales avec différents laboratoires. Que ce soit avec des laboratoires nationaux (Grenoble, Lyon, Toulouse, Brest, Île de France) ou internationaux (Algérie, Amérique du sud ou Océanie), ces collaborations ont apporté de nombreuses idées et fait progresser l'état de l'art. Il est donc important de continuer ces collaborations ainsi que d'en créer d'autres dans les années à venir, par l'intermédiaire notamment de montage de projets.
La forte implication dans le domaine des technologies pour la santé rend important un autre type de collaboration, celle avec le monde médical. Les travaux précédents se sont déjà déroulés dans le cadre de telles collaborations (avec les kinésithérapeutes ou les gériatres par exemple), cependant cela devra s'accroitre au fur et à mesure que les solutions se rapprocheront de hauts niveaux de TRL [3]. En effet, pour que celles-ci soient utilisables et utilisées, elles doivent se rapprocher au plus des outils actuels et s'adapter au médecin et non le contraire. L'inclusion dans la Structure Fédérative de Recherche Technologies pour la Santé et Médicament [4] portée par l'Université de Lille représente également un atout permettant de pousser ces collaborations plus loin.
Les discussions avec la SATT Nord permettront peut-être par la suite de trouver un modèle économique viable et de transférer tout ou partie des technologies développées, ce qui est important en recherche. Le modèle économique pour le moment est difficile à appréhender car sur ce type d'application les revenus classiques ne sont pas utilisables (publicité etc.). Dans le cadre de tels développement il faut donc trouver des partenariats avec des mutuelles ou avec des financeurs de la santé, et le modèle actuel tend également sur la proposition d'abonnement avec un service spécifique offert. La valorisation de toutes ces activités de recherche sera un point abordé plus en détail dans les années à venir.
Enfin, et pour conclure, les annexes suivantes vont présenter une vue succincte de mon parcours et de mes activités, en y ajoutant toutes les activités d'enseignement, mais également une forte implication dans la communauté scientifique

3. Technology Readiness Level
4. http ://sfr-tsm.ec-lille.fr

avec des éditions de numéros spéciaux de journaux, des participations à des organisations de manifestations ainsi que des comités éditoriaux. Ces annexes vont aussi lister les travaux publiés et les projets de recherche financés passés et en cours.

Bibliographie

ABBATE, S., M. AVVENUTI, F. BONATESTA, G. COLA, P. CORSINI et A. VECCHIO (2012). « A smartphone-based fall detection system ». In : *Pervasive and Mobile Computing* 8.6, p. 883-899.

AGGARWAL, J. K. et L. XIA (2014). « Human activity recognition from 3d data : a review ». In : *Pattern Recognition Letters* 48, p. 70-80.

ALLAMEH, E., M. HEIDARI JOZAM, B. d. VRIES, H. TIMMERMANS et J. BEETZ (2011). « Smart home as a smart real estate : a state of the art review. » In : *18th International Conference of European Real Estate Society, Eindhoven, The Netherlands Eindhoven : ERES*.

AMOUD, H., M. ABADI, D. J. HEWSON, V. MICHEL-PELLEGRINO, M. DOUSSOT et J. DUCHÊNE (2007). « Fractal time series analysis of postural stability in elderly and control subjects ». In : *Journal of neuroengineering and rehabilitation* 4.1, p. 12.

ANAS, E. M. A., S. Y. LEE et M. K. HASAN (2011). « Exploiting correlation of ecg with certain emd functions for discrimination of ventricular fibrillation ». In : *Computers in biology and medicine* 41.2, p. 110-114.

ANGUITA, D., A. GHIO, L. ONETO, X. PARRA et J. L. REYES-ORTIZ (2013). « A public domain dataset for human activity recognition using smartphones. » In : *ESANN*.

AL-ANI, T., F. CAZETTES, S. PALFI et J.-P. LEFAUCHEUR (2011). « Automatic removal of high-amplitude stimulus artefact from neuronal signal recorded in the subthalamic nucleus ». In : *Journal of neuroscience methods* 198.1, p. 135-146.

ANIRUDH, R., P. TURAGA, J. SU et A. SRIVASTAVA (2015). « Elastic functional coding of human actions : from vector-fields to latent variables ». In : *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, p. 3147-3155.

BAO, L. et S. S. INTILLE (2004). « Activity recognition from user-annotated acceleration data ». In : *Pervasive Computing : Second International Conference (PERVASIVE 2004), Linz/Vienna, Austria, April 21-23*. Springer-Verlag. ISBN : 978-3-540-21835-7.

BEZDEK, J. C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Norwell, MA, USA : Kluwer Academic Publishers. ISBN : 0306406713.

BORG, F. G. et G. LAXABACK (juil. 2010). « Entropy of balance - some recent results. » eng. In : *J Neuroeng Rehabil* 7.1, p. 38. DOI : 10.1186/1743-0003-7-38.

BOUBACAR, H. A., S. LECOEUCHE et S. MAOUCHE (2008). « Sakm : self-adaptive kernel machine a kernel-based algorithm for online clustering ». In : *Neural Networks* 21.9, p. 1287-1301.

BOUISSET, S. et B. MATON (1995). « Muscles, posture et mouvement ». In : *Bases et applications de la méthode électromyographique. Paris : Hermann* 26.

BOUKHAROUBA, K., L. BAKO et S. LECOEUCHE (2009). « Incremental and decremental multi-category classification by support vector machines ». In : *Machine Learning and Applications, 2009. ICMLA'09. International Conference on*. IEEE, p. 294-300.

BROSSEAU, L., M. TOUSIGNANT, J. BUDD, N. CHARTIER, L. DUCIAUME, S. PLAMONDON, J. P. O'SULLIVAN, S. O'DONOGHUE et S. BALMER (1997). « Intratester and intertester reliability and criterion validity of the parallelogram and universal goniometers for active knee flexion in healthy subjects ». In : *Physiotherapy Research International* 2.3, p. 150-166.

CAMPBELL, A. et T. CHOUDHURY (2012). « From smart to cognitive phones ». In : *IEEE Pervasive Computing* 11.3, p. 7-11.

CHAARAOUI, A. A., J. R. PADILLA-LÓPEZ, P. CLIMENT-PÉREZ et F. FLÓREZ-REVUELTA (2014). « Evolutionary joint selection to improve human action recognition with rgb-d devices ». In : *Expert systems with applications* 41.3, p. 786-794.

CHAHUARA, P., A. FLEURY, F. PORTET et M. VACHER (juil. 2016). « On-line human activity recognition from audio and home automation sensors : comparison of sequential and non-sequential models in realistic smart homes ». In : *Journal of Ambient Intelligence and Smart Environments* 8.4, p. 399-422. DOI : 10.3233/AIS-160386.

CHAHUARA, P., A. FLEURY, M. VACHER et F. PORTET (2012). « Méthodes SVM et MLN pour la reconnaissance automatique d'activités humaines dans les habitats perceptifs : tests et perspectives ». In : *RFIA 2012 (Reconnaissance des Formes et Intelligence Artificielle)*, p. 978-2.

CHANG, Y.-J., S.-F. CHEN et J.-D. HUANG (2011). « A kinect-based system for physical rehabilitation : a pilot study for young adults with motor disabilities ». In : *Research in developmental disabilities* 32.6, p. 2566-2570.

CHEN, C., K. LIU et N. KEHTARNAVAZ (2016). « Real-time human action recognition based on depth motion maps ». In : *Journal of real-time image processing* 12.1, p. 155-163.

CIPPITELLI, E., S. GASPARRINI, E. GAMBI et S. SPINSANTE (2016). « A human activity recognition system using skeleton data from rgbd sensors ». In : *Computational intelligence and neuroscience* 2016, p. 21.

CLARK, R. A., Y.-H. PUA, K. FORTIN, C. RITCHIE, K. E. WEBSTER, L. DENEHY et A. L. BRYANT (2012). « Validity of the microsoft kinect for assessment of postural control ». In : *Gait & posture* 36.3, p. 372-377.

COOK, D. J. et S. K. DAS (2012). « Pervasive computing at scale : transforming the state of the art ». In : *Pervasive and Mobile Computing* 8.1, p. 22-35.

COOK, D., M. SCHMITTER-EDGECOMBE, A. CRANDALL, C. SANDERS et B. THOMAS (2009). « Collecting and disseminating smart home sensor data in the casas project ». In : *Proceedings of the CHI Workshop on Developing Shared Home Behavior Datasets to Advance HCI and Ubiquitous Computing Research*.

CUTTI, A. G., A. GIOVANARDI, L. ROCCHI, A. DAVALLI et R. SACCHETTI (2008). « Ambulatory measurement of shoulder and elbow kinematics through inertial and magnetic sensors ». In : *Medical & biological engineering & computing* 46.2, p. 169-178.

DADASHI, F., B. MARIANI, S. ROCHAT, C. J. BÜLA, B. SANTOS-EGGIMANN et K. AMINIAN (2013). « Gait and foot clearance parameters obtained using shoe-worn inertial sensors in a large-population sample of older adults ». In : *Sensors* 14.1, p. 443-457.

DAUBECHIES, I., J. LU et H.-T. WU (2011). « Synchrosqueezed wavelet transforms : an empirical mode decomposition-like tool ». In : *Applied and Computational Harmonic Analysis* 30.2, p. 243-261. ISSN : 1063-5203. DOI : 10.1016/j.acha.2010.08.002.

DAUBECHIES, I. et S. MAES (1996). « A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models ». In : *Wavelets in medicine and biology*, p. 527-546.

DEVANNE, M., H. WANNOUS, S. BERRETTI, P. PALA, M. DAOUDI et A. DEL BIMBO (2015). « 3-d human action recognition by shape analysis of motion trajectories on riemannian manifold ». In : *IEEE transactions on cybernetics* 45.7, p. 1340-1352.

DIAZ, M., G. JUAN, O. LUCAS et A. RYUGA (2012). « Big data on the internet of things : an example for the e-health ». In : *Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2012 Sixth International Conference on*. IEEE, p. 898-900.

DONKER, S. F., M. RŒRDINK, A. J. GREVEN et P. J. BEEK (juil. 2007). « Regularity of center-of-pressure trajectories depends on the amount of attention invested in postural control. » eng. In : *Experimental Brain Research* 181.1, p. 1-11. DOI : 10.1007/s00221-007-0905-4.

DOZZA, M., L. CHIARI et F. B. HORAK (2005). « Audio-biofeedback improves balance in patients with bilateral vestibular loss ». In : *Archives of physical medicine and rehabilitation* 86.7, p. 1401-1403.

DU, Y., W. WANG et L. WANG (2015). « Hierarchical recurrent neural network for skeleton based action recognition ». In : *Proceedings of the IEEE conference on computer vision and pattern recognition*, p. 1110-1118.

DUCHÊNE, J. et D. J. HEWSON (oct. 2011). « Longitudinal evaluation of balance quality using a modified bathroom scale : usability and acceptability. » eng. In : *Journal of Telemedicine and Telecare* 17.7. DOI : 10.1258/jtt.2011.110307.

DUNN, J. C. (1973). « A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters ». In : *Journal of Cybernetics* 3.3, p. 32-57. DOI : 10.1080/01969727308546046.

DUONG, T., D. PHUNG, H. BUI et S. VENKATESH (mai 2009). « Efficient duration and hierarchical modeling for human activity recognition ». In : *Artificial Intelligence* 173.7-8, p. 830-856.

EUSTON, M., P. COOTE, R. MAHONY, J. KIM et T. HAMEL (2008). « A complementary filter for attitude estimation of a fixed-wing uav ». In : *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*. IEEE, p. 340-345.

FINLAYSON, G., B. SCHIELE et J. CROWLEY (1998). « Comprehensive colour image normalization ». In : *ECCV'98 Fifth European Conference on Computer Vision*, p. 475-490.

FLEURY, A., N. VUILLERME, B. DIOT et P. HLAVACKOVA (28 sept. 2017). « Procédé et système pour l'estimation d'une demande ou d'un coût attentionel associé à l'éxecution d'une tâche ou des stratégies de partage de l'attention développées par un individu ». FR 30 49371 / WO 2017/162789 A1.

FLEURY, A., N. VUILLERME, C. FRANCO, Q. MOURCOU et B. DIOT (27 mai 2016). « Procédé et système pour la mesure, le suivi, le contrôle et la correction d'un mouvement ou d'une posture d'un utilisateur ». FR 14 61233 / WO 2016/079452 A1 / US Patent 20170354843.

FLEURY, A., Q. MOURCOU, C. FRANCO, B. DIOT, J. DEMONGEOT et N. VUILLERME (2013). « Evaluation of a smartphone-based audio-biofeedback system for improving balance in older adults-a pilot study ». In : *Engineering in Medicine and Biology Society, 2013. EMBS 2013. 35th Annual International Conference of the IEEE*. IEEE, p. 1198-1201. DOI : 10.1109/EMBC.2013.6609721.

FLEURY, A., Q. MOURCOU, C. FRANCO et N. VUILLERME (oct. 2018). « Implementation and evaluation of a smartphone app to assess attentional demand for balance control ». In : *Engineering in Medicine and Biology Society, 2018. EMBS 2018. 40th Annual International Conference of the IEEE*, p. 5598-5601. DOI : 10.1109/EMBC.2018.8513540.

FLEURY, A., M. VACHER et N. NOURY (2010). « SVM-based multimodal classification of activities of daily living in health smart homes : sensors, algorithms, and first experimental results ». In : *Information Technology in Biomedicine, IEEE Transactions on* 14.2, p. 274-283.

FLEURY, A., M. VACHER, F. PORTET, P. CHAHUARA et N. NOURY (2013). « A french corpus of audio and multimodal interactions in a health smart home ». In : *Journal on Multimodal User Interfaces* 7.1-2, p. 93-109.

FLORES, M., G. GLUSMAN, K. BROGAARD, N. D. PRICE et L. HOOD (2013). « P4 medicine : how systems medicine will transform the healthcare sector and society ». In : *Personalized medicine* 10.6, p. 565-576.

FRANCO, C., J. FONTECAVE-JALLON, N. VUILLERME et P. GUMÉRY (2011). « Towards a suitable time-scale representation of cardio-respiratory signals through empirical mode decomposition algorithms : a simulation and validation tool ». In : *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*. IEEE, p. 802-805.

FRANCO, C., A. FLEURY, B. DIOT et N. VUILLERME (oct. 2018). « Applying entropy to human center of foot pressure data to assess attention investment in balance control ». In : *Engineering in Medicine and Biology Society, 2018. EMBS 2018. 40th Annual International Conference of the IEEE*. (***Accepté***), p. 5586-5589. DOI : 10.1109/EMBC.2018.8513533.

FRANCO, C., A. FLEURY, P.-Y. GUMÉRY, B. DIOT, J. DEMONGEOT et N. VUILLERME (2013). « iBalance-ABF : a smartphone-based audio-biofeedback balance system ». In : *Biomedical Engineering, IEEE Transactions on* 60.1, p. 211-215.

FRANCO, C., P.-Y. GUMERY, A. FLEURY et N. VUILLERME (2014). « Analyzing postural signals from a synchrosqueezing transform-based approach of wavelet entropy ». In : *4th International Congress on Complex Systems in Sports & Healthy Ageing*.

FRANCO, C., P.-Y. GUMÉRY, N. VUILLERME, A. FLEURY et J. FONTECAVE-JALLON (2012). « Synchrosqueezing to investigate cardio-respiratory interactions within simulated volumetric signals ». In : *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*. IEEE, p. 939-943.

FREEMAN, J. (2017). « An investigation into the correlation between step width variability and balance deficits in patients with chronic obstructive pulmonary disease ». In :

GAJDOSIK, R. L. et R. W. BOHANNON (1987). « Clinical measurement of range of motion : review of goniometry emphasizing reliability and validity ». In : *Physical therapy* 67.12, p. 1867-1872.

GALLISSOT, M., J. CAELEN, F. JAMBON et B. MEILLON (2013). « Une plate-forme usage pour l'intégration de l'informatique ambiante dans l'habitat : Domus ». In : *Technique et Science Informatiques (TSI)* 32/5, p. 547-574.

GHAFFARIANHOSEINI, A., A. GHAFFARIANHOSEINI, J. TOOKEY, H. OMRANY, A. FLEURY, N. NAISMITH et M. GHAFFARIANHOSEINI (mar. 2016). « Handbook of research on creative technologies for multidisciplinary applications ». In : sous la dir. d'A. CONNOR et S. MARKS. IGI Global. Chap. The Essence of Smart Homes : Application of Intelligent Technologies towards Smarter Urban Future, Ch. 14, Ch. 14. DOI : 10.4018/978-1-5225-0016-2.ch014.

GHORBEL, E. (2017). « Fast and accurate human action recognition using RGB-D cameras ». Thèse de doct. Normandie Université.

GIDDENS, L., E. GONZALEZ et D. LEIDNER (2016). *I track, therefore I Am : exploring the impact of wearable fitness devices on employee identity and well-being*. Rapp. tech.

GLEICK, J. (2011). *Chaos : Making a new science (Enhanced edition)*. Open Road Media.

GOLDBERGER, A. L. (1996). « Non-linear dynamics for clinicians : chaos theory, fractals, and complexity at the bedside ». In : *The Lancet* 347.9011, p. 1312-1314.

GOVINDAN, R., J. WILSON, H. ESWARAN, C. LOWERY et H. PREISSL (2007). « Revisiting sample entropy analysis ». In : *Physica A : Statistical Mechanics and its Applications* 376, p. 158-164. ISSN : 0378-4371. DOI : 10.1016/j.physa.2006.10.077.

GU, T., Z. WU, X. TAO, H. K. PUNG et J. LU (mar. 2009). « EpSICAR : an emerging patterns based approach to sequential, interleaved and concurrent activity recognition ». In : *Pervasive Computing and Communications, 2009. PerCom 2009. IEEE International Conference on*, p. 1-9. DOI : 10.1109/PERCOM.2009.4912776.

GUBBI, J., R. BUYYA, S. MARUSIC et M. PALANISWAMI (2013). « Internet of things (iot) : a vision, architectural elements, and future directions ». In : *Future generation computer systems* 29.7, p. 1645-1660.

GUVENSAN, M. A., A. O. KANSIZ, N. C. CAMGOZ, H. TURKMEN, A. G. YAVUZ et M. E. KARSLIGIL (2017). « An energy-efficient multi-tier architecture for fall detection using smartphones ». In : *Sensors* 17.7, p. 1487.

HABIB, M. A., M. S. MOHKTAR, S. B. KAMARUZZAMAN, K. S. LIM, T. M. PIN et F. IBRAHIM (2014). « Smartphone-based solutions for fall detection and prevention : challenges and open issues ». In : *Sensors* 14.4, p. 7181-7208.

HALL, M. (1999). « Correlation-based feature selection for machine learning ». Thèse de doct. University of Waikato, New-Zealand.

HAMACHER, D., D. BERTRAM, C. FÖLSCH et L. SCHEGA (2012). « Evaluation of a visual feedback system in gait retraining : a pilot study ». In : *Gait & posture* 36.2, p. 182-186.

HAMMOUCHE, M., E. GHORBEL, A. FLEURY et S. AMBELLOUIS (fév. 2016). « Toward a real time view-invariant 3d action recognition ». In : *11th International Conference on Computer Vision Theory and Applications (VISAPP'16)*.

HAN, J., L. SHAO, D. XU et J. SHOTTON (2013). « Enhanced computer vision with microsoft kinect sensor : a review ». In : *IEEE transactions on cybernetics* 43.5, p. 1318-1334.

HÄNDEL, P., J. OHLSSON, M. OHLSSON, I. SKOG et E. NYGREN (2014). « Smartphone-based measurement systems for road vehicle traffic monitoring and usage-based insurance ». In : *IEEE Systems Journal* 8.4, p. 1238-1248.

HARALICK, R., K. SHANMUGAM et I. DINSTEIN (1973). « Textural features for image classification ». In : *Systems, Man and Cybernetics, IEEE Transactions on* 3.6, p. 610-621. ISSN : 0018-9472.

HILLIARD, M. J., K. M. MARTINEZ, I. JANSSEN, B. EDWARDS, M.-L. MILLE, Y. ZHANG et M. W. ROGERS (2008). « Lateral balance factors predict future falls in community-living older adults ». In : *Archives of physical medicine and rehabilitation* 89.9, p. 1708-1713.

HÖRSTER, E., J. LUX et R. LIENHART (2007). « Recognizing persons in images by learning from videos ». In : *Proceedings of SPIE*. T. 6506, p. 65060D.1-65060D.9.

HUANG, N. E., Z. SHEN, S. R. LONG, M. C. WU, H. H. SHIH, Q. ZHENG, N.-C. YEN, C. C. TUNG et H. H. LIU (1998). « The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis ». In : *Proceedings of the Royal Society of London A : mathematical, physical and engineering sciences*. T. 454. 1971. The Royal Society, p. 903-995.

HUSNJAK, S., D. PERAKOVIĆ, I. FORENBACHER et M. MUMDZIEV (2015). « Telematics system in usage based motor insurance ». In : *Procedia Engineering* 100, p. 816-825.

INTILLE, S. S., K. LARSON, E. M. TAPIA, J. S. BEAUDIN, P. KAUSHIK, J. NAWYN et R. ROCKINSON (2006). « Using a live-in laboratory for ubiquitous computing research ». In : *Pervasive Computing*. Springer, p. 349-365.

ISABLEU, B. M., P. HLAVACKOVA, B. DIOT et N. VUILLERME (2017). « Regularity of center of pressure trajectories in expert gymnasts during bipedal closed-eyes quiet standing ». In : *Frontiers in Human Neuroscience* 11, p. 317.

JALLON, J. F., P.-Y. GUMÉRY, P. CALABRESE, R. BRIOT et P. BACONNIER (2013). « A wearable technology revisited for cardio-respiratory functional exploration : stroke volume estimation from respiratory inductive plethysmography ». In : *International Journal of eHealth and Medical Communications* 4.1, p. 12-22.

JENNY, J.-Y. (2013). « Measurement of the knee flexion angle with a smartphone-application is precise and accurate ». In : *The Journal of arthroplasty* 28.5, p. 784-787.

JOHN, G., R. KOHAVI et K. PFLEGER (1994). « Irrelevant features and the subset selection problem ». In : *Proceedings of the eleventh international conference on machine learning*. T. 129, p. 121-129.

KAHNEMAN, D. (1973). *Attention and effort*. Sous la dir. de N. ENGLEWOOD CLIFFS. Prentice-Hall Inc., 246p.

KARAGIANNIS, A., P. CONSTANTINOU et D. VOUYIOUKAS (2011). « Biomedical time series processing and analysis methods : the case of empirical mode decomposition ». In : *Advanced Biomedical Engineering*. InTech.

KARAMI, A. B. et A. FLEURY (juin 2016). « Using feedback in adaptive and user-dependent one-step decision making ». In : *IJCAI-16, 25th International Joint Conference on Artificial Intelligence, Workshop Interactive Machine Learning*, p. ID 5.

KARAMI, A. B., A. FLEURY, J. BOONAERT et S. LECOEUCHE (juin 2016). « User in the loop : adaptive smart homes exploiting user feedback — state of the art and future directions ». In : *Information* 7.2, p. 35. ISSN : 2078-2489.

KASTEREN, T. van, G. ENGLEBIENNE et B. KRÖSE (2010). « An activity monitoring system for elderly care using generative and discriminative models ». In : *Personal and Ubiquitous Computing* 14.6, p. 489-498.

KASTEREN, T. L. M. van, G. ENGLEBIENNE et B. J. A. KRÖSE (2011). « Hierarchical activity recognition using automatically clustered actions ». In : *Proceedings of the Second international conference on Ambient Intelligence*. AmI'11. Berlin, Heidelberg : Springer-Verlag, p. 82-91.

KATZ, S. (1983). « Assessing self-maintenance : activities of daily living, mobility, and instrumental activities of daily living ». In : *Journal of the American Geriatrics Society* 31.12, p. 721-727.

KAZISKA, D. et A. SRIVASTAVA (2007). « Gait-based human recognition by classification of cyclostationary processes on nonlinear shape manifolds ». In : *Journal of the American Statistical Association* 102.480, p. 1114-1124.

KEJONEN, P. (2004). « Body movements during postural stabilization. Measurements with a motion analysis system ». Thèse de doct. Oulun Yliopisto (Finland).

KIM, H.-C., S. PANG, H.-M. JE, D. KIM et S.-Y. BANG (2002). « Support vector machine ensemble with bagging ». In : *Pattern recognition with support vector machines*, p. 397-408.

KNOOP, J., M. STEULTJENS, M. VAN DER LEEDEN, M. VAN DER ESCH, C. THORSTENSSON, L. ROORDA, W. LEMS et J. DEKKER (2011). « Proprioception in knee osteoarthritis : a narrative review ». In : *Osteoarthritis and Cartilage* 19.4, p. 381-388.

KOSSE, N. M., S. CALJOUW, D. VERVOORT, N. VUILLERME et C. J. LAMOTH (2015). « Validity and reliability of gait and postural control analysis using the tri-axial accelerometer of the ipod touch ». In : *Annals of biomedical engineering* 43.8, p. 1935-1946.

KRISHNAN, N. C. et D. J. COOK (2014). « Activity recognition on streaming sensor data ». In : *Pervasive and mobile computing* 10, p. 138-154.

LAFFERTY, J. D., A. MCCALLUM et F. C. N. PEREIRA (2001). « Conditional random fields : probabilistic models for segmenting and labeling sequence data ». In : *International Conference on Machine Learning,ICML'01*, p. 282-289.

LAMOTH, C. J., F. J. van DEUDEKOM, J. P. van CAMPEN, B. A. APPELS, O. J. de VRIES et M. PIJNAPPELS (jan. 2011). « Gait stability and variability measures show effects of impaired cognition and dual tasking in frail people. » eng. In : *Journal of NeuroEngineering and Rehabilitation* 8.1, p. 2. DOI : 10.1186/1743-0003-8-2.

LAWTON, M. et E. BRODY (1969). « Assessment of older people : self-maintaining and instrumental activities of daily living ». In : *Gerontologist* 9, p. 179-186.

LEBEL, K., P. BOISSY, M. HAMEL et C. DUVAL (2015). « Inertial measures of motion for clinical biomechanics : comparative assessment of accuracy under controlled conditions–changes in accuracy over time ». In : *PloS one* 10.3, e0118361.

LEE, J., B. A. REYES, D. D. MCMANUS, O. MAITAS et K. H. CHON (2013). « Atrial fibrillation detection using an iphone 4s ». In : *IEEE Transactions on Biomedical Engineering* 60.1, p. 203-206.

LI, W., Z. ZHANG et Z. LIU (2010). « Action recognition based on a bag of 3d points ». In : *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*. IEEE, p. 9-14.

LI, Y., J. GAO, C. MENG, Q. LI, L. SU, B. ZHAO, W. FAN et J. HAN (2015). « A survey on truth discovery ». In : *arXiv preprint arXiv :1505.02463*.

LIKERT, R. (1932). « A technique for the measurement of attitudes. » In : *Archives of psychology*.

LIU, W., Y. ZHANG, S. TANG, J. TANG, R. HONG et J. LI (2013). « Accurate estimation of human body orientation from rgb-d sensors ». In : *IEEE Transactions on Cybernetics* 43.5, p. 1442-1452.

LIU, Z., C. ZHANG et Y. TIAN (2016). « 3d-based deep convolutional neural network for action recognition with depth sequences ». In : *Image and Vision Computing* 55, p. 93-100.

LU, Y., K. BOUKHAROUBA, J. BOONÆRT, A. FLEURY et S. LECOEUCHE (2014). « Application of an incremental SVM algorithm for on-line human recognition from video surveillance using texture and color features ». In : *Neurocomputing* 126, p. 132-140.

LU, Y., A. FLEURY, J. BOONAERT, S. LECOEUCHE et S. AMBELLOUIS (2015). « Online person identification and new person discovery using appearance features ». In : *Evolving and Adaptive Intelligent Systems (EAIS), 2015 IEEE International Conference on*. IEEE, p. 1-8.

MADELEINE, P., M. NIELSEN et L. ARENDT-NIELSEN (avr. 2011). « Characterization of postural control deficit in whiplash patients by means of linear and nonlinear analyses - a pilot study. » eng. In : *Journal of Electromyography and Kinesiology* 21.2, p. 291-297. DOI : 10.1016/j.jelekin.2010.05.006.

MADGWICK, S. (2010). « An efficient orientation filter for inertial and inertial/magnetic sensor arrays ». In : *Report x-io and University of Bristol (UK)* 25.

MAHONY, R., T. HAMEL et J.-M. PFLIMLIN (2008). « Nonlinear complementary filters on the special orthogonal group ». In : *IEEE Transactions on automatic control* 53.5, p. 1203-1218.

MAKRISA, D., N. DOULAMISC et S. MIDDLETOND (mar. 2009). « Vision-Based Production of Personalized Video ». In : *Signal Processing : Image Computation* 24.5, p. 158-176.

MILANESE, S., S. GORDON, P. BUETTNER, C. FLAVELL, S. RUSTON, D. COE, W. O'SULLIVAN et S. MCCORMACK (2014). « Reliability and concurrent validity of knee angle measurement : smart phone app versus universal goniometer used by experienced and novice clinicians ». In : *Manual therapy* 19.6, p. 569-574.

MILANI, P., C. A. COCCETTA, A. RABINI, T. SCIARRA, G. MASSAZZA et G. FERRIERO (2014). « Mobile smartphone applications for body position measurement in rehabilitation : a review of goniometric tools ». In : *PM&R* 6.11, p. 1038-1043.

MIRANDA, L., T. VIEIRA, D. MARTÍNEZ, T. LEWINER, A. W. VIEIRA et M. F. CAMPOS (2014). « Online gesture recognition from pose kernel learning and decision forests ». In : *Pattern Recognition Letters* 39, p. 65-73.

MITCHELL, K., S. B. GUTIERREZ, S. SUTTON, S. MORTON et A. MORGENTHALER (2014). « Reliability and validity of goniometric iphone applications for the assessment of active shoulder external rotation ». In : *Physiotherapy theory and practice* 30.7, p. 521-525.

MOLINA, L. C., L. BELANCHE et À. NEBOT (2002). « Feature selection algorithms : a survey and experimental evaluation ». In : *Proceedings of the IEEE International Conference on Data Mining (ICDM 2002)*, p. 306-313.

MOURCOU, Q., A. FLEURY, B. DIOT et N. VUILLERME (août 2016). « iProprio : a smartphone-based system to measure and improve proprioceptive function ». In : *Engineering in Medicine and Biology Society, 2016. EMBS 2016. 38th Annual International Conference of the IEEE*, p. 2622-2625. DOI : 10.1109/EMBC.2016.7591268.

MOURCOU, Q., A. FLEURY, B. DIOT, C. FRANCO et N. VUILLERME (2015). « Mobile phone-based joint angle measurement for functional assessment and rehabilitation of proprioception ». In : *BioMed Research International* 2015.2015, Article 328142, 15 p.

MOURCOU, Q., A. FLEURY, C. FRANCO, F. KLOPCIC et N. VUILLERME (2015). « Performance evaluation of smartphone inertial sensors measurement for range of motion ». In : *Sensors* 15.9, p. 23168-23187.

MURRAY, A. M., J. H. RYU, J. SPROULE, A. P. TURNER, P. GRAHAM-SMITH et M. CARDINALE (2017). « A pilot study using entropy as a non-invasive assessment of running ». In : *International journal of sports physiology and performance*, p. 1-13.

NAFUS, D. et J. SHERMAN (2014). « Big data, big questions| this one does not go up to 11 : the quantified self movement as an alternative big data practice ». In : *International journal of communication* 8, p. 11.

NAZERFARD, E., B. DAS, L. B. HOLDER et D. J. COOK (2010). « Conditional random fields for activity recognition in smart environments ». In : *Proceedings of the 1st ACM International Health Informatics Symposium*. IHI '10. Arlington, Virginia, USA : ACM, p. 282-286. ISBN : 978-1-4503-0030-8. DOI : 10.1145/1882992.1883032.

NILSSON, F. et al. (2016). *Intelligent network video : Understanding modern video surveillance systems*. CRC Press.

OHN-BAR, E. et M. M. TRIVEDI (2013). « Joint angles similarities and HOG2 for action recognition ». In : *Computer Vision and Pattern Recognition Workshops (CVPRW), 2013 IEEE Conference on*. IEEE, p. 465-470.

OREIFEJ, O. et Z. LIU (2013). « Hon4d : histogram of oriented 4d normals for activity recognition from depth sequences ». In : *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, p. 716-723.

PACHORI, R. B. et V. BAJAJ (2011). « Analysis of normal and epileptic seizure eeg signals using empirical mode decomposition ». In : *Computer methods and programs in biomedicine* 104.3, p. 373-381.

PASTERNACK, J. et D. ROTH (2010). « Knowing what to believe (when you already know something) ». In : *Proceedings of the 23rd International Conference on Computational Linguistics*, p. 877-885.

PATEL, M., P. A. FRANSSON, D. LUSH et S. GOMEZ (nov. 2008). « The effect of foam surface properties on postural stability assessment while standing. » eng. In : *Gait Posture* 28.4, p. 649-656. DOI : 10.1016/j.gaitpost.2008.04.018.

PEARSON, K. (1901). « On lines and planes of closest fit to systems of points in space ». In : *Philosophical Magazine* 2.6, p. 559-572.

PENG, C.-K., M. COSTA et A. L. GOLDBERGER (2009). « Adaptive data analysis of complex fluctuations in physiologic time series ». In : *Advances in adaptive data analysis* 1.01, p. 61-70.

PINCUS, S. M. (1991). « Approximate entropy as a measure of system complexity ». In : *Proceedings of the National Academy of Sciences of the United States of America* 88.6, p. 2297-2301. eprint : http://www.pnas.org/content/88/6/2297.full.pdf+html.

QUINLAN, J. R. (1996). « Bagging, boosting, and C4.5 ». In : *Proceedings of the 13th National Conference on Artificial Intelligence (AAAI-96)*. AAAI/MIT Press, p. 725-730.

RABINER, L. (1989). « A tutorial on Hidden Markov Models and selected applications in speech recognition ». In : *Proceedings of the IEEE* 77.2, p. 257-286.

RABINER, L. et B. JUANG (1986). « An introduction to hidden markov models ». In : *ieee assp magazine* 3.1, p. 4-16.

RAMDANI, S., B. SEIGLE, J. LAGARDE, F. BOUCHARA et P. L. BERNARD (2009). « On the use of sample entropy to analyze human postural sway data ». In : *Medical Engineering & Physics* 31.8, p. 1023-1031. ISSN : 1350-4533. DOI : DOI:10.1016/j.medengphy.2009.06.004.

RICHARDSON, M. et P. DOMINGOS (2006). « Markov logic networks ». In : *Machine Learning* 62.1-2, p. 107-136.

RICHMAN, J. S. et J. R. MOORMAN (2000). « Physiological time-series analysis using approximate entropy and sample entropy ». In : *Am J Physiol Heart Circ Physiol* 278.6, H2039-2049. eprint : http://ajpheart.physiology.org/cgi/reprint/278/6/H2039.pdf.

RILEY, M. A., A. A. BAKER, J. M. SCHMIT et E. WEAVER (juil. 2005). « Effects of visual and auditory short-term memory tasks on the spatiotemporal dynamics and variability of postural sway. » eng. In : *J Mot Behav* 37.4, p. 311-324. DOI : 10.3200/JMBR.37.4.311-324.

RILLING, G. et P. FLANDRIN (2008). « One or two frequencies? the empirical mode decomposition answers ». In : *IEEE transactions on signal processing* 56.1, p. 85-95.

ROARK, D., A. O'TOOLE et H. ABDI (2003). « Human recognition of familiar and unfamiliar people in naturalistic video ». In :

ROERDINK, M., P. HLAVACKOVA et N. VUILLERME (juin 2010). « Center-of-pressure regularity as a marker for attentional investment in postural control : a comparison between sitting and standing postures. » eng. In : *Hum Mov Sci*. DOI : 10.1016/j.humov.2010.04.005.

ROETENBERG, D., H. LUINGE et P. SLYCKE (2009). *Xsens MVN : full 6DOF human motion tracking using miniature inertial sensors*. Rapp. tech. Xsens Motion Technologies BV, Tech. Rep.

ROSSO, O. A., S. BLANCO, J. YORDANOVA, V. KOLEV, A. FIGLIOLA, M. SCHÜRMANN et E. BAŞAR (2001). « Wavelet entropy : a new tool for analysis of short duration brain electrical signals ». In : *Journal of neuroscience methods* 105.1, p. 65-75.

SAEYS, Y., I. INZA et P. LARRAÑAGA (2007). « A review of feature selection techniques in bioinformatics ». In : *Bioinformatics* 23.19, p. 2507. ISSN : 1367-4803.

SCHMIT, J. M., D. I. REGIS et M. A. RILEY (juin 2005). « Dynamic patterns of postural sway in ballet dancers and track athletes. » eng. In : *Exp Brain Res* 163.3, p. 370-378. DOI : 10.1007/s00221-004-2185-6.

SEGAL, M. (2004). *Machine learning benchmarks and random forest regression*. Rapp. tech. San Francisco, CA, USA : Center for Bioinformartics et Molecular Biostatistics, University of California.

SEIDENARI, L., V. VARANO, S. BERRETTI, A. BIMBO et P. PALA (2013). « Recognizing actions from depth cameras as weakly aligned multi-part bag-of-poses ». In : *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, p. 479-485.

SEVRIN, L. (2016). « Mesure et suivi d'activité de plusieurs personnes dans un Living Lab en vue de l'extraction d'indicateurs de santé et de bien-être ». Thèse de doct. Université de Lyon.

SHINMURA, F., D. DEGUCHI, I. IDE, H. MURASE et H. FUJIYOSHI (2015). « Estimation of human orientation using coaxial rgb-depth images ». In : *International Conference on Computer Vision Theory and Applications (VISAPP)*, p. 113-120.

SHOTTON, J., A. FITZGIBBON, A. BLAKE, A. KIPMAN, M. FINOCCHIO, B. MOORE et T. SHARP (juin 2011). « Real-time human pose recognition in parts from a single depth image ». In : *Computer Vision and Pattern Recognition (CVPR)*. IEEE. URL : https://www.microsoft.com/en-us/research/publication/real-time-human-pose-recognition-in-parts-from-a-single-depth-image/.

SINGHALA, P., D. N. SHAH et B. PATEL (jan. 2014). « Temperature control using fuzzy logic ». In : *International Journal of Instrumentation and Control Systems* 4.1, p. 1-10. ISSN : 2249-1147.

STINS, J., M. MICHIELSEN, M. RŒRDINK et P. BEEK (2009). « Sway regularity reflects attentional involvement in postural control : effects of expertise, vision and cognition ». In : *Gait & Posture* 30.1, p. 106-109. ISSN : 0966-6362. DOI : DOI:10.1016/j.gaitpost.2009.04.001.

STONE, E. E. et M. SKUBIC (2015). « Fall detection in homes of older adults using the microsoft kinect ». In : *IEEE journal of biomedical and health informatics* 19.1, p. 290-301.

STRANG, A. J., J. HAWORTH, M. HIERONYMUS, M. WALSH et L. J. SMART (déc. 2010). « Structural changes in postural sway lend insight into effects of balance training, vision, and support surface on postural control in a healthy population. » eng. In : *Eur J Appl Physiol*. DOI : 10.1007/s00421-010-1770-6.

SWAN, M. (2012). « Sensor mania! the internet of things, wearable computing, objective metrics, and the quantified self 2.0 ». In : *Journal of Sensor and Actuator Networks* 1.3, p. 217-253.

— (2013). « The quantified self : fundamental disruption in big data science and biological discovery ». In : *Big Data* 1.2, p. 85-99.

TALHA, S. A. W., A. FLEURY et S. LECOEUCHE (2019). « Position-invariant recognition of human activities and postural transitions using smartphones ». In : *Sensors*. (***Soumis***).

TALHA, S. A. W., A. FLEURY et S. AMBELLOUIS (2017). « Human action recognition from body-part directional velocity using hidden markov models ». In : *16th IEEE International Conference On Machine Learning And Applications*.

TALHA, S. A. W., M. HAMMOUCHE, E. GHORBEL, A. FLEURY et S. AMBELLOUIS (déc. 2018). « Features and classification schemes for view-invariant and real-time human action recognition ». In : *IEEE Transaction on Cognitive and Developmental Systems* 10.4. DOI : 10.1109/TCDS.2018.2844279.

TAPIA, E. M., S. S. INTILLE et K. LARSON (2004). « Activity recognition in the home using simple and ubiquitous sensors ». In : *International Conference on Pervasive Computing*. Springer, p. 158-175.

TRUONG CONG, D., L. KHOUDOUR, C. ACHARD, C. MEURIE et O. LEZORAY (2010). « People re-identification by spectral classification of silhouettes ». In : *Signal Processing* 90.8, p. 2362-2374. ISSN : 0165-1684.

TRUONG CONG, D., L. KHOUDOUR, C. ACHARD et L. DOUADI (2010). « People Detection and Re-Identification in Complex Environments ». In : *IEICE Transactions on Information and Systems* 93.7, p. 1761-1772. ISSN : 1745-1361.

TSAI, C.-W., C.-F. LAI, M.-C. CHIANG, L. T. YANG et al. (2014). « Data mining for internet of things : a survey. » In : *IEEE Communications Surveys and Tutorials* 16.1, p. 77-97.

VACHER, M., A. FLEURY, F. PORTET, J.-F. SERIGNAT et N. NOURY (fév. 2010). « Complete Sound and Speech Recognition System for Health Smart Homes : Application to the Recognition of Activities of Daily Living ». In : *New Developments in Biomedical Engineering*. Sous la dir. de D. CAMPOLO. ISBN : 978-953-7619-57-2. In-Tech, pp. 645 -673.

VACHER, M., B. LECOUTEUX, P. CHAHUARA, F. PORTET, B. MEILLON et N. BONNEFOND (2014). « The Sweet-Home speech and multimodal corpus for home automation interaction ». In : *The 9th edition of the Language Resources and Evaluation Conference (LREC)*. Reykjavik, Iceland, p. 4499-4506.

VAIL, D. L., M. M. VELOSO et J. D. LAFFERTY (2007). « Conditional random fields for activity recognition ». In : *Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems*. AAMAS '07. New York, NY, USA, 235 :1-235 :8.

VARNFIELD, M., M. KARUNANITHI, C.-K. LEE, E. HONEYMAN, D. ARNOLD, H. DING, C. SMITH et D. L. WALTERS (2014). « Smartphone-based home care model improved use of cardiac rehabilitation in postmyocardial infarction patients : results from a randomised controlled trial ». In : *Heart*, heartjnl-2014.

VELIK, R. (jan. 2014). « A brain-inspired multimodal data mining approach for human activity recognition in elderly homes ». In : *Journal of Ambient Intelligence and Smart Environments* 6.4, p. 447-468. DOI : 10.3233/AIS-140266.

VEMULAPALLI, R., F. ARRATE et R. CHELLAPPA (2014). « Human action recognition by representing 3D skeletons as points in a Lie group ». In : *Proceedings of the IEEE conference on computer vision and pattern recognition*. IEEE, p. 588-595.

VENKATARAMAN, V., P. TURAGA, N. LEHRER, M. BARAN, T. RIKAKIS et S. WOLF (2013). « Attractor-shape for dynamical analysis of human movement : applications in stroke rehabilitation and action recognition ». In : *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, p. 514-520.

VERMEULEN, M., A. FLEURY, K. FRONTON et J. LAVAL (2015). « Les alpes : approches agiles pour l'enseignement supérieur ». In : *Colloque Questions de Pédagogie pour l'Enseignement Supérieur (QPES 2015)*.

VILLAIN, J., V. DENIAU, A. FLEURY, P. E. SIMON, C. GRANSART et R. KOUSRI (2019). « EM monitoring and classification of IEMI and protocol-based attacks on IEEE 802.11n communication networks ». In : *IEEE Transactions on Electromagnetic Compatibility*. ***[IF Courant : 1.52]***.

VISCHER, J. C. (2007). « The concept of environmental comfort in workplace performance ». In : *Ambiente Construído* 7.1, p. 21-34.

VOHRALIK, S. L., A. R. BOWEN, J. BURNS, C. E. HILLER et E. J. NIGHTINGALE (2015). « Reliability and validity of a smartphone app to measure joint range ». In : *American journal of physical medicine & rehabilitation* 94.4, p. 325-330.

WANG, L., T. GU, X. TAO et J. LU (2012). « A hierarchical approach to real-time activity recognition in body sensor networks ». In : *Pervasive and Mobile Computing* 8.1, p. 115-130.

WARD, A. F., K. DUKE, A. GNEEZY et M. W. BOS (2017). « Brain drain : the mere presence of one's own smartphone reduces available cognitive capacity ». In : *Journal of the Association for Consumer Research* 2.2, p. 140-154. DOI : 10.1086/691462.

WELSH, G. et G. BISHOP (1995). « An introduction to the kalman filter ». In : *University of North Carolina at Chapel Hill Chapel Hill NC* 95, p. 95-041.

WICKENS, C. (1984). « Varieties of attention ». In : sous la dir. de R. PARASURAMEN et D. DAVIES. Orlando : Academic. Chap. Processing resources in attention. P. 63-101.

WILLIAMSON, R. et B. ANDREWS (2001). « Detecting absolute human knee angle and angular velocity using accelerometers and rate gyroscopes ». In : *Medical and Biological Engineering and Computing* 39.3, p. 294-302.

WILSON, C., T. HARGREAVES et R. HAUXWELL-BALDWIN (2014). « Smart homes and their users : a systematic analysis and key challenges ». In : *Personal and Ubiquitous Computing* 19.2, p. 463-476.

WINTER, J. S. (2014). « Surveillance in ubiquitous network societies : normative conflicts related to the consumer in-store supermarket experience in the context of the internet of things ». In : *Ethics and Information Technology* 16.1, p. 27.

WORRINGHAM, C., A. ROJEK et I. STEWART (2011). « Development and feasibility of a smartphone, ecg and gps based system for remotely monitoring exercise in cardiac rehabilitation ». In : *PloS one* 6.2, e14669.

WU, H.-T., P. FLANDRIN et I. DAUBECHIES (2011). « One or two frequencies ? the synchrosqueezing answers ». In : *Advances in Adaptive Data Analysis* 3.01n02, p. 29-39.

XIA, L., C.-C. CHEN et J. K. AGGARWAL (2011). « Human detection using depth information by kinect ». In : *Computer Vision and Pattern Recognition Workshops (CVPRW), 2011 IEEE Computer Society Conference on*. IEEE, p. 15-22.

XIA, L., C.-C. CHEN et J. AGGARWAL (2012). « View invariant human action recognition using histograms of 3d joints ». In : *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on*. IEEE, p. 20-27.

XU, Z., C. HU et L. MEI (2016). « Video structured description technology based intelligence analysis of surveillance videos for public security applications ». In : *Multimedia Tools and Applications* 75.19, p. 12155-12172.

YALA, N., B. FERGANI et A. FLEURY (avr. 2017). « Towards improving feature extraction and classification for activity recognition on streaming data ». In : *Journal of Ambient Intelligence and Humanized Computing* 8.2, p. 177-189. DOI : 10.1007/s12652-016-0412-1.

YANG, X., C. ZHANG et Y. TIAN (2012). « Recognizing actions using depth motion maps-based histograms of oriented gradients ». In : *Proc. of the 20th ACM international conference on Multimedia*, p. 1057-1060.

YOON, K., D. HARWOOD et L. DAVIS (2006). « Appearance-based person recognition using color/path-length profile ». In : *Journal of Visual Communication and Image Representation* 17.3, p. 605-622. ISSN : 1047-3203.

YOUNG, W., S. FERGUSON, S. BRAULT et C. CRAIG (fév. 2011). « Assessing and training standing balance in older adults : a novel approach using the 'nintendo wii' balance board. » eng. In : *Gait Posture* 33.2, p. 303-305. DOI : 10.1016/j.gaitpost.2010.10.089.

ZHOU, X. et B. BHANU (2007). « Integrating face and gait for human recognition at a distance in video ». In : *Systems, Man, and Cybernetics, Part B : Cybernetics, IEEE Transactions on* 37.5, p. 1119-1137. ISSN : 1083-4419.

ZHOU, X. et B. BHANU (2008). « Feature fusion of side face and gait for video-based human identification ». In : *Pattern Recognition* 41.3, p. 778-795. ISSN : 0031-3203.

ZHU, N., T. DIETHE, M. CAMPLANI, L. TAO, A. BURROWS, N. TWOMEY, D. KALESHI, M. MIRMEHDI, P. FLACH et I. CRADDOCK (2015). « Bridging e-health and the internet of things : the sphere project ». In : *IEEE Intelligent Systems* 30.4, p. 39-46.

Annexes
Curriculum Vitæ détaillé

Annexe A

Curriculum Vitæ

Anthony FLEURY

IMT Lille Douai, Département Informatique et Automatique
941 Rue Charles Bourseul, BP 10838, 59508 Douai Cedex
Tel. : 06 43 96 79 85
E-mail : Anthony.Fleury@imt-lille-douai.fr
Né le 30 mai 1982 à Harfleur – Nationalité Française

Maître Assistant de l'Institut Mines Télécom
IMT Lille Douai

Formation Universitaire

Jan. 2009 et 2014 : Qualification aux fonctions de Maître de conférences par les sections CNU 27 et 61.

2005 – 2008 : Doctorat en Traitement du Signal à l'école doctorale EEATS (Électronique, Électrotechnique, Automatique et Traitement du Signal) à Grenoble.

2004 – 2005 : Master 2 Recherche Signal, Image, Parole, Télécommunications à l'Institut National Polytechnique de Grenoble. Mention Bien.

2002 – 2005 : Ingénieur en Informatique Industrielle et Instrumentation (3I) à Polytech'Grenoble (spécialité de dernière année : Signal et Images). Mention Bien.

2000 – 2002 : Classe préparatoire Physique Technologie (PT) au Lycée Robert Schuman du Havre.

2000 : Baccalauréat Scientifique option Physique-Chimie au Lycée François 1er du Havre.

Expérience Professionnelle

Depuis Avril 2011 : Maître assistant à IMT Lille Douai. Enseignement de l'informatique en tronc commun (Programmation et Bases de données relationnelles). Membre de l'Unité de Recherche Informatique et Automatique d'IMT Lille Douai.

Sept. 2009 - Mars 2011 : Enseignant-Chercheur statut ARMINES à l'école des Mines de Douai.

Sept. 2008 - Sept. 2009 : Contrat post-doctoral au Laboratoire de Mesure et d'Analyse du Mouvement (LMAM) à l'École Polytechnique Fédérale de Lausanne (Suisse), sous la direction du Professeur Kamiar Aminian.

2005 – 2008 : Thèse de doctorat de l'Université Joseph Fourier de Grenoble, spécialité Traitement du Signal à l'école doctorale EEATS (Électronique, Électrotechnique, Automatique et Traitement du Signal) à Grenoble. Allocation de Recherche du Ministère de l'Éducation Nationale et de la Recherche Technologique, sous la direction du Professeur Norbert Noury (TIMC-IMAG) et du Dr. Michel Vacher (LIG).

Compétences Techniques

— Programmation en C (Standard et POSIX), C++, Java, Objective-C, Swift, Shell Unix (bash), VB.

— Programmation de systèmes embarqués et d'objets connectés.

— SQL, Bases de données en MySQL et Access.

— Matlab, R, Weka.

Langues

— Anglais : Lecture/écriture d'articles scientifiques, présentation dans plusieurs conférences internationales, développement de collaboration internationales.

Annexe B

Activités d'enseignement

B.1 Résumé des activités d'enseignement

Année	Établissement	Niveau	Effectif	Description	Volume Horaire
Depuis 2017–2018 :	IMT Lille Douai	M1		Mesure et traitement de données physiologiques chez l'humain.	12h
Depuis 2017–2018 :	IMT Lille Douai	M1		Création et encadrement de projets sur la validation de mesures physiologiques avec des objets connectés.	16h
Octobre 2016 – Mars 2017 :	IMT Lille Douai	M2		Co-encadrement (avec Dr. Abir-Béatrice Karami Lozenguez) d'une élève de filière recherche (2 jours par semaine). Sujet : interfacer les données de la salle de pause intelligente d'IMT Lille Douai avec les algorithmes de décision.	
2016–2017 :	IMT Lille Douai	M1		Encadrement technique d'un étudiant en filière Étudiant-Entrepreneur de l'Université d'Artois (étudiant de la majeure ISIC d'IMT Lille Douai)	10h
Mars – Juin 2016 :	IMT Lille Douai			Encadrement d'un stage d'un élève CMSI (formation pour chercheur d'emploi) sur la réalisation d'un circuit de mesure embarqué sur Raspberry Pi pour le département Génie Civil.	
Octobre 2015 – Mars 2016 :	IMT Lille Douai	M2		Co-encadrement (avec Dr. Abir-Béatrice Karami Lozenguez) de deux élèves en filière recherche de l'école (2 jours par semaine) sur la réalisation d'une interface graphique pour gérer la salle de pause de Mines Douai et indiquer dans la base les actions utilisateurs.	
Depuis 2015–2016 :	IMT Lille Douai	L3		Encadrement de projets ouverts sur le développement mobile. Projets de première année, travaillés par un groupe d'environ 5-6 élèves tout au long de l'année et pour un client extérieur. 3 projets en 2015, 3 en 2016, 2 en 2017.	
Depuis 2014–2015 :	IMT Lille Douai	L3	20	Cours de Swift (programmation iDevices Apple)	24h
Depuis 2013–2014 :	IMT Lille Douai	M1	20	Langage C	20h
Depuis 2012–2013 :	IMT Lille Douai	M1	25	Data mining	16h
	IMT Lille Douai	M1	25	Programmation embarquée sur plateforme iOS en Objective-C.	36h
2011–2012 :	IMT Lille Douai	M1	30/Grp	TD de Programmation Orientée Objet	24h
Depuis 2010–2011 :	IMT Lille Douai	L3	30/Grp	Informatique en tronc commun de la formation d'ingénieurs (Cours, TD, TP d'algorithmie et programmation en Java et enseignement de Merise/Bases de données relationnelles et SQL)	130h
	IMT Lille Douai	M1	3/Grp	Projet de découverte de la recherche, 80h élèves (option Informatique et option Automatique)	4 projets/an
2009–2010 :	IMT Lille Douai	L3	30/Grp	Informatique en tronc commun de la formation d'ingénieurs (Td, TP d'algorithmie et programmation en Visual Basic et enseignement de Merise/Bases de données relationnelles et SQL	144h

Année	Établissement	Niveau	Effectif	Description	Volume Horaire
2007–2008 :	Polytech'Grenoble			Installation et maintenance de la salle informatique Linux de Polytech'Grenoble	20 h
	Polytech'Grenoble	E2I1	16	Langage C et méthodologie de la programmation	49h
	Polytech'Grenoble	E2I3	14	BE d'Architecture des Processeurs (MIPS) – utilisation d'un simulateur sur machine SUN	4h
2006–2007 :	IUP GSI	L3	35	TD d'automatique	18h
	Polytech'Grenoble	E2I1	12	Langage C et méthodologie de la programmation	51h
2005–2006 :	IUP GSI	L3	10	Tutorat d'automatique	6h

B.2 Implication dans les activités administratives liées à la pédagogie d'IMT Lille Douai

Depuis Sept. 2017 : Création, avec Arnaud Doniec responsable du parcours informatique, et animation d'une mineure (90h d'enseignement) en Technologie pour la Santé. Ouverture rentrée 2017 avec pour la première année 10 étudiants et cette seconde année 10 étudiants aussi (le maximum défini en fonction du matériel disponible).

2015 – 2016 : Participation active aux groupes de travail sur la fusion entre Mines Douai et Télécom Lille sur les aspects formation et recherche. Dans la première phase définition du modèle, et dans la seconde phase, organisation par lot pour couvrir toute la formation. Participation principalement au lot sur le tronc commun.

Depuis Oct. 2013 : Membre élu (actuellement deuxième mandat) du comité de l'enseignement et du jury des études et de la commission pédagogique d'IMT Lille Douai.

Depuis Sept. 2011 : Organisation des intervenants et emplois du temps de l'Unité d'Enseignement Mathématiques et Informatique pour la partie Informatique (couvrant le tronc commun de Licence 3).

B.3 Enseignements réalisés et en cours

Le cursus à l'école des Mines a fortement évolué lors de la rentrée 2012-2013. D'un enseignement tronc commun (1 an) puis Option (2 ans), le cursus est passé à Tronc Commun (1 an - L3) - Mineure (1 semestre - le premier de M1) - Majeure (1 semestre - le second de M1) - Option (1 an - M2). La mineure est un enseignement d'un semestre, permettant de donner une coloration spécialisée qui ne doit pas être un pré-requis à la majeure (les parcours étant souples, une même mineure mène à plusieurs majeures) mais qui serait en relation avec celle-ci. Ainsi, une mineure sera commune à plusieurs majeures/options et permettra à des élèves qui n'ont pas eu l'option voulue d'avoir tout de même des enseignements en rapport avec leur souhait (pour les majeures/options les plus demandées et dont l'affectation se fait alors sur le classement scientifique de première année). Leur parcours sera ainsi plus à la carte.

B.3.1 Tronc commun IMT Lille Douai

Lors de mon arrivée à l'école des Mines, dans une école généraliste, j'ai été affecté sur les cours de tronc commun et ce afin que nous les fassions évoluer. Les cours de tronc commun se font devant environ deux cents élèves dans la promotion, élèves qui pour une grande partie d'entre-eux ne se destinent pas à une carrière informatique (l'option la plus demandée étant le génie énergétique suivi de génie civil) et donc qui ne perçoivent pas tout de suite l'intérêt et la place de l'informatique dans leur futur métier.

Ainsi, nous avons défini un socle minimum de connaissances à acquérir en algorithmique de base (notion de boucles, de conditions, de récursivité, de tri, etc.). Le langage support du cours est Java, sans cependant énormément insister

sur la notion de classe (nous leur présentons les notions de méthode de classes et d'instances, les portées des variables etc. mais nous ne présentons pas par exemple l'héritage ou les exceptions, vus en deuxième année). La difficulté de ce cours (4h de cours en classe inversée, 12h de TD et 24h de TP par élèves, les TP et TD étant en groupe d'une trentaine d'élèves) réside justement en l'apprentissage des bases de l'algorithmique en même temps que l'apprentissage du langage support dans un temps très court. Cet enseignement a été l'occasion de créer nombre de supports de cours, de TD, de TP, ou encore des examens. Toujours en tronc commun, j'ai également pris part aux enseignements de base de données (cours, TD et TP), en enseignant la méthode Merise et l'utilisation du logiciel Microsoft Access (création de formulaires, de tables, de relations, etc. puis utilisation du SQL). Cet enseignement a ensuite été modifié pour utiliser des méthodes de gestion de projet AGILES d'une part et pour utiliser MySQL d'autre part.

B.3.2 Cours électifs et cours de mineures

Lors de l'évolution du cursus évoquée précédemment, en accord avec les responsables de majeures, nous avons créé deux mineures, l'une sur le Développement Mobile, dans laquelle je réalise l'ensemble du module Développement iOS, et l'autre sur l'ingénierie de la décision, dans laquelle j'enseigne un module de Data Mining. Cependant ces cours sont une introduction, le volume horaire n'étant pas suffisant pour approfondir les notions à ce moment du cursus. Lors de la rentrée 2013, nous avons ajouté, à la mineure Développement Mobile, un cours de langage C. Ce cours de C est utile tant pour iOS que pour un cours de Systèmes embarqués. J'ai été responsable de créer et animer l'ensemble de ce module de C (Cours, TD et TP) en plus du cours de développement iOS.
La formation permet également aux étudiants de sélectionner des cours dits électifs, en première année. Ces cours sont au choix de l'apprenant qui va pouvoir choisir trois modules différents parmi tous ceux qui sont proposés à la carte par l'équipe pédagogique. En septembre 2014, Apple a mis à disposition un nouveau langage de programmation, Swift. Devant l'engouement certain pour les applications mobiles, j'ai donc créé, en janvier 2015, un cours électif sur ce langage récent. J'anime ce cours depuis, en le faisant évoluer au rythme des évolutions du langage.

B.4 Domaine de la Santé

En accord avec la direction des études et de la formation, nous avons créé, pour cette rentrée 2017, une nouvelle mineure sur la santé connectée et les technologies pour la Santé. Cette mineure va inclure un cours de développements mobiles (pour aborder un minimum le développement avec des accessoires tels que les montres connectées), des cours de développement et test de solutions embarquées pour l'acquisition et le traitement de données physiologiques, des cours de traitement de données (le cours de développement mobile et celui de traitement de données seront mutualisé avec les autres parcours) et enfin des séminaires de professionnels du monde des données de santé pour aborder les questions de sécurité, partage et accès à ces données sensibles ainsi que leur anonymisation et toutes les normes en cours (4 conférences de 4h). Un projet permettant d'implémenter des solutions de mesures de telles données sera effectué par les étudiants pour évaluer ce parcours.

B.5 Nouvelles pratiques et innovations pédagogiques

Comme pour tout enseignant du supérieur à l'heure actuelle, confronté à ce que les sociologues appellent la génération Z, les méthodes d'enseignement classiques ne sont plus adaptées. En effet, cette génération, ayant toujours connu l'informatique et Internet, a un accès facilité aux médias et à l'information. Ils ont également tendance à prendre uniquement ce qui les intéressent dans un contenu ou dans une formation et sont ainsi très exigeants. Il est alors plus difficile de leur dispenser des cours dans un modèle très classique tel qu'il l'était avant et les méthodes pédagogiques doivent ainsi évoluer.
C'est pourquoi dans la formation d'IMT Lille Douai, en collaboration avec les enseignants des différentes matières de tronc commun notamment, des méthodes d'enseignement différentes sont expérimentées. Elles sont présentées rapidement dans les sections suivantes.

B.5.1 Méthode ALPES

Avec notre collègue Jannik Laval, l'ensemble des intervenants du cours de SGBD de tronc commun et la cellule TICE de l'école, nous avons retravaillé en profondeur ce cours afin de lui donner un nouveau souffle et une nouvelle forme. À la place de travailler, comme précédemment, sur un mode magistral avec imitation, des méthodes de gestion de projet Agiles ont été adoptées à nos pratiques pédagogiques afin de motiver les étudiants et de les faire travailler à leur rythme, en autonomie, pour une meilleure assimilation. Cette méthode d'enseignement, baptisée Approches ALPES (Approches Agiles Pour l'Enseignement Supérieur), offre une toute nouvelle gestion du temps et de la promotion [1]. Cette méthode nous oblige à repenser totalement la manière dont le cours est dispensé et à tout découper de manière cohérente et compatible avec une construction incrémentale de la connaissance. Elle a permis de faciliter la tâche des intervenants lors du cours et tout en permettant d'obtenir une forte adhésion des étudiants. Un article de retour d'expérience dans la conférence QPES (Questions de pédagogies pour l'enseignement supérieur) a montré l'évolution positive suite à la mise en place de cette méthode [NC1].
Celle-ci est également en partie appliquée pour le cours de développement mobile en Swift (les contraintes de groupes étant différentes et le cours étant un choix de l'élève, tous les outils pédagogiques de la méthode ALPES ne sont pas repris dans ce cours).
Il est à noter que les travaux sur cette nouvelle méthode d'enseignement ont donné lieu au montage d'un projet (projet APACHES) qui a été accepté pour un lancement en janvier 2019. Ce projet, financé par l'i-Site ULNE (Université Lille-Nord-Europe), prévoit d'étudier, au travers des données qu'il est possible de collecter, l'impact de cette méthode et la meilleure mise en place possible, notamment dans des formations comme le doctorat.

B.5.2 Classes inversées

Deux heures de cours magistraux pour présenter les notions fondamentales d'algorithmique et ce quatre fois de suite sont difficiles à appréhender par les étudiants. Présenter ainsi les notions de récursivité ou de tri est compliqué. Il a donc été décidé, il y a deux ans, en accord avec les autres enseignants de TD d'algorithmique et programmation, de mettre en ligne le cours magistral, afin que les étudiants le travaillent par eux-même, puis de faire une séance de travail dans laquelle je réponds à toutes leurs questions sur ce cours. Le temps gagné sur les séances de cours magistraux est ainsi converti en heures de TD afin que l'assimilation des concepts soit mieux ancré par la pratique.
Cette formule est ainsi bien acceptée d'après les retours des étudiants (sondages chaque année à la fin des différents cours).

B.5.3 Notation automatisée et basée sur le jeu

Enfin, avec mon collègue Jérémie Humeau, intervenant de TP algorithmique, nous construisons depuis 3 ans maintenant des évaluations différentes des TP notés classiques pour le cours d'algorithmiques. Nous nous basons à chaque fois sur un jeu de société connu, que nous re-codons en partie pour les élèves, et à qui sur un TP nous faisons acquérir des bases de celui-ci en leur faisant écrire certaines parties de ce jeu (typiquement le score, les fins de parties etc.).
Ensuite, pour le TP noté, en dernière séance, les élèves ont une consigne simple. Réaliser une stratégie de jeu automatisée (une intelligence artificielle) pour ce jeu. Nous fournissons un certains nombre de stratégies paliers (de la plus bête à la plus sophistiquée) qui permettent de séparer les étudiants en groupes. Le classement obtenu alors en faisant s'affronter tous les binômes de chacun des groupes déterminés, nous permettant d'attribuer automatiquement des notes.
Les tentatives de tricheries sont éliminées par deux méthodes. La première est que nous détectons les stratégies ayant un comportement similaire (les exécutions aléatoires sur un grand nombre de parties seront à quasiment 50%-50%). Ensuite, nous vérifions avec un outil de détection de plagiat donnant les taux de recopie de code (cela détecte y compris les changements de variables ou les maigres changements d'indices ou de méthodes de boucles).
Cette notation est stimulante pour les étudiants qui nous racontent se faire des « concours » dans la résidence avec ce jeu, en étant motivés par la compétition. Le temps de préparation est par contre pour nous important car nous

1. http://approchealpes.info

devons re-coder le jeu (avec son interface et toutes ses règles) et qu'il nous faut en changer chaque année (pour chaque promotion).
À noter également que, du fait de l'intérêt fort des étudiants pour ce types d'application, le cursus créé pour la nouvelle école inclus un cours de création de stratégies et d'intelligences artificielles appliquées aux jeux (cours électif en première année – L3).

B.6 Initiation à la recherche des élèves ingénieurs

Les élèves d'école d'ingénieurs se destinent, pour la plupart d'entre eux, à des métiers techniques et à du management. Cependant, nous leur offrons la possibilité de se sensibiliser à l'innovation et à la recherche, lors de leur formation, par l'intermédiaire de projets. Cela se fait par deux types de projets. Le premier est le Projet de Découverte de la Recherche. Ce projet, de M1, est prévu sur 80h à l'emploi du temps, avec des sujets donnés par les enseignants-chercheurs et choisis ensuite parmi les sujets disponibles par les étudiants en groupes de 2 ou 3. Cela leur permet de découvrir des sujets un peu moins dirigés et structurés que des sujets classiques et d'apprendre à réaliser une étude bibliographique (et donc l'habituent à un format de lecture différent d'une documentation). Chaque enseignant-chercheur encadre plusieurs projets par an. Les sujets que j'ai proposés portent principalement sur mes activités de recherche (smartphone, habitat intelligent, etc.). Par exemple, nous avons fait travailler des élèves sur l'habitat intelligent pour inclure de nouvelles données dans le living lab, sur les smartphones pour les connecter à une plateforme de force connectée, etc.
La seconde initiation se fait, au choix des étudiants, en dernière année (niveau M2). Lors de leur dernière année, les étudiants choisissent une filière (logistique, entrepreneuriat, recherche, etc.) sur laquelle ils vont travailler environ deux jours par semaine. Ces deux dernières années, des étudiants en informatique ont choisi ce parcours. Nous avons donc eu l'occasion, avec Abir Karami, d'encadrer deux étudiants en 2015 et une qui commence en 2016, sur les thématiques actuelles liées à l'habitat intelligent. Cette opportunité montre aux étudiants ce qu'est vraiment notre domaine avec des sujets plus poussés et plus longs et leur permet aussi d'acquérir de nouvelles méthodes de travail.

B.7 Évolution des activités d'enseignement dans le nouveau programme d'IMT Lille Douai

Suite à la création de l'école IMT Lille Douai, née de la fusion de Mines Douai et de Télécom Lille, au 1er janvier 2017, un nouveau cursus pédagogique a été certifié par la commission des titres d'ingénieurs et les années L1 et L3 de ce cursus ont été démarrées en septembre 2018. C'est une refonte majeure de la formation avec deux changements importants. Le premier est le format des cours et des interventions, beaucoup plus basé sur la notion d'UV, avec disparition des mineures, majeures et options. Le second changement est la taille des promotions qui va doubler à l'horizon 2020 et ainsi nous devrons nous adapter pour tout ce qui est organisation des cours et évaluations (aller encore plus vers des notations automatisées). En plus du nombre d'élèves, une évolution importante est la diversité des entrées dans l'école avec une admission en post-bac en plus des admissions en L3 et M1.
Ce cursus est également, du fait de l'orientation initiale de Télécom Lille et en accord avec les besoins actuels des entreprises, beaucoup plus orienté vers le numérique que précédemment. En plus d'être une spécialisation comme actuellement, les sciences du numérique seront aussi un socle fondamental « transversal » à toutes les spécialisations. Cette réforme est l'occasion de monter de nouveaux programmes de cours et de définir un nouveau socle de connaissances indispensables dans la société telle que nous la connaissons actuellement. Pour ceci, la participation active aux groupes de travail sur cette fusion a été importante. Par exemple, le tronc commun possède maintenant un seul cours d'algorithmique et programmation orienté objet (reprenant et combinant intelligemment les cours de L3 et de M1 faits précédemment). Les modules électifs et mineures seront transformées en UV à la rentrée 2019.

Annexe C

Activités de recherche académique et appliquée

Sommaire

B.1 Résumé des activités d'enseignement . . . 107
B.2 Implication dans les activités administratives liées à la pédagogie d'IMT Lille Douai . . . 108
B.3 Enseignements réalisés et en cours . . . 108
B.3.1 Tronc commun IMT Lille Douai . . . 108
B.3.2 Cours électifs et cours de mineures . . . 109
B.4 Domaine de la Santé . . . 109
B.5 Nouvelles pratiques et innovations pédagogiques . . . 109
B.5.1 Méthode ALPES . . . 110
B.5.2 Classes inversées . . . 110
B.5.3 Notation automatisée et basée sur le jeu . . . 110
B.6 Initiation à la recherche des élèves ingénieurs . . . 111
B.7 Évolution des activités d'enseignement dans le nouveau programme d'IMT Lille Douai . . . 111

C.1 Mots clés

C.1.1 Recherche [1]

— PE6_1 Computer architecture, pervasive computing, ubiquitous computing

— PE6_2 Computer systems, parallel/distributed systems, sensor networks, embedded systems, cyber-physical systems

— PE6_11 Machine learning, statistical data processing and applications using signal processing (e.g. speech, image, video)

— PE6_12 Scientific computing, simulation and modelling tools

— PE7_4 Systems engineering, sensorics, actorics, automation

— PE7_7 Signal processing

— PE7_8 Networks (communication networks, sensor networks, networks of robots...)

— PE8_10 Production technology, process engineering

— LS7_1 Medical engineering and technology

C.1.2 Applications

— Santé personnalisée

— Autonomie et qualité de vie

— Bâtiments intelligents

— Sécurité et cyber-sécurité

— Analyse de données pour l'usine du futur

1. Basés sur les mots clés donnés par l'European Research Council et utilisés cette année également par l'Agence Nationale de la Recherche

C.2 Résumé des activités

C.2.1 Formation à la recherche

2005 – 2008 : Thèse de doctorat de l'Université Joseph Fourier de Grenoble, spécialité Traitement du Signal à l'école doctorale EEATS (Électronique, Électrotechnique, Automatique et Traitement du Signal) à Grenoble. Allocation de Recherche du Ministère de l'Éducation Nationale et de la Recherche Technologique.
Directeurs de thèse : Pr. Norbert Noury (TIMC-IMAG, équipe AFIRM), et Dr. Michel Vacher (LIG, équipe GETALP)
Sujet : 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 télémédecine.
Date de soutenance : 24 octobre 2008
Jury :
Présidente : **Michèle ROMBAUT** (Pr., Grenoble)
Rapporteurs : **Olivier BASSET** (Pr., Lyon) et **François CHARPILLET** (D.R. INRIA, Nancy)
Examinateurs : **Jacques DUCHÊNE** (Pr., Troyes) et **Kamiar AMINIAN** (Pr., Lausanne)
Directeurs : **Norbert NOURY** (Pr., Lyon) et **Michel VACHER** (I.R. CNRS, Grenoble)
Invités : **Nicolas VUILLERME** (MCU, Grenoble) et **Vincent RIALLE** (MCU, Grenoble)

Avril – Sept. 2005 : Stage Recherche de Master 2 au laboratoire TIMC-IMAG, équipe AFIRM, sous la direction du Pr. Norbert Noury.
Titre : *Exploitation des Signaux de Capteurs Magnétométriques Embarqués. Application à la Détection automatique des Activités de Vie Quotidienne d'une Personne.*

C.2.2 Formation par la recherche

Depuis Novembre 2018 : Encadrement du post-doctorat du Dr. Jonathan Villain, projet SECOURT (ELSAT2020) financé par Région/FEDER. Sujet : Classification d'attaques sur des réseaux wifi à partir de données des couches hautes (suite).

Sept. 2017 - Novembre 2018 : Encadrement du post-doctorat du Dr. Jonathan Villain, projet X2Rail-1 financé par l'IRT Railenium. Sujet : Classification d'attaques sur des réseaux wifi à partir de données des couches hautes.

Depuis Oct. 2016 : Co-Encadrement (Directeur : Pr. Stéphane Lecoeuche) de la thèse de Sid Ahmed Walid Talha, avec co-financement obtenu auprès de la région Hauts de France. Sujet : Apprentissage et segmentation pour le suivi de comportements à domicile. **Taux d'encadrement : 0,5**.

Juin 2016 - Mai 2017 : Co-encadrement (avec Dr. Jacques Boonaert et Pr. Moamar Sayed-Mouchaweh) du post-doctorat du Dr. Houari Toubakh, projet REPAR. Sujet : Apprentissage dynamique sur des données Sparses.

Mars - Sept. 2016 : Co-encadrement (avec Dr. Sébastien Ambellouis) du stage de Master 2 de Sid Ahmed Walid Talha (Université Toulouse 3). Sujet : Analyse évolutive de l'interaction humaine perçue par un système de vision.

Mars - Sept. 2015 : Co-encadrement (avec Dr. Sébastien Ambellouis) du stage de Master 2 de Mounir Hammouche (Université de Paris-Est Créteil). Sujet : Human action recognition for the intervention of a proactive assistance robot.

Jan. 2013 - Jan. 2016 : Co-encadrement (Directeur : Dr. Nicolas Vuillerme) de la thèse de Quentin Mourcou, avec co-financement Université de Grenoble Alpes. Sujet : Conception et validation d'un système ubiquitaire centré autour du Smartphone et dédié à l'amélioration de l'autonomie de la personne sur son lieu de vie. **Taux d'encadrement : 0,5**. Situation actuelle : Co-fondateur et Directeur Recherche de la société Wegoto.

Fev. - Juin 2012 : Co-encadrement (avec Pr. Yvonne Delevoye) du stage de master 1 de Manon Lecomte (Université Lille 3). Sujet : Extraction d'informations et classification à partir de tests et de questionnaires.

Fev. 2010 – Sept. 2014 : Co-encadrement (Directeur : Pr. Stéphane Lecoeuche, co-encadrant avec Jacques Boonaert) de la thèse de Yanyun Lu, avec financement gouvernement Chinois. Sujet : Online classification and clustering of persons using appearance-based features from video images : application to person discovery and re-identification in multicamera environments. **Taux d'encadrement : 0,3**. Situation actuelle : en rapprochement familial, sans emploi.

Sept. 2009 – Mars 2014 : Co-encadrement (Directeurs : Pr. Pierre-Yves Gumery et Dr. Nicolas Vuillerme) de la thèse de Céline Franco, financement Université de Grenoble. Sujet : Maintient à domicile de la personne âgée fragile : fusion de données physiologiques et environnementales pour la détection de dérives comportementales. **Taux d'encadrement : 0,3**. Situation actuelle : après un CDD au CEA, a travaillé chez Sublimed (gestion de la douleur) puis est maintenant en CDI chez Diabeloop.

C.2.3 Principales participations à des projets de recherche

À partir de Jan. 2019 : Projet APACHES (Apprentissage des conduites de Projets Agiles et Centrés Humain dans l'Enseignement Supérieur). Sujet : Proposer aux enseignants un cadre théorique et des outils pour former par la pratique les étudiants sur des méthodes de conduite de projets et de recherche centrées humain (approche par compétences). Partenaires : IMT Lille Douai, ED SPI (Université de Lille), IUT Lyon 2 Lumières, équipe MOCAH (LIP6, Sorbonne Université), IUT 2 Grenobles Alpes, Ecole Doctorale MSTII (Université Grenoble Alpes). Financement : I-Site ULNE.

Mai 2017 – Oct. 2017 : PoC (proof of concept) Teralab e.l.m. Leblanc. Sujet : Analyse des données de production de chaudières en vue de la prédiction de possibles retours en SAV dans la période de garantie. Partenaires : IMT Lille Douai, PIA Teralab, e.l.m Leblanc. Financement : e.l.m. Leblanc.

Mai 2017 – Déc. 2017 : Projet COMRADES (Coordinated Multi-Robot Assistance Deployment in Smart Spaces). Sujet : Flottes de robots mobiles pour l'assistance à la personne. Partenaires : IMT Lille Douai, IMT Atlantique. Financement : Institut Mines Télécom.

Mai 2017 – Déc. 2017 : Projet SolSens (Sol Sensitif pour Analyse Cognitive et Actimétrique). Sujet : Analyse des données d'un sol sensitif pour un jeu sérieux en santé. Partenaires : IMT Mines Alès, Télécom Sud Paris et IMT Atlantique. Financement : Institut Mines Télécom.

Déc. 2016 – Déc. 2018 : Projet e-Monitor (Chronic DIsease : Ambient Assisted Living and vital teleMONITORing for e-health). Sujet : Cette collaboration entre France et Amérique du Sud vise à partager nos travaux sur la reconnaissance d'activités, la détection de situations de détresse et les jeux sérieux en Santé. Partenaires : Institut Mines télécom (TSP/IMT Lille Douai), UTC, SAMU-92, Universidade Federal do Ceará (Brésil), Universidad de Valparaíso (Chilie), ICT4V (Uruguay), Instituto Federal do Espírito Santo (Brésil). Financement : STIC-AmSud.

Depuis 2016 : Projet SECOURT (Cyber-SEcurité dans les systèmes COmmUnicants pour les Transports). Sujet : Sécurisation des communications Inter-Véhicules et Véhicules-Infrastructures pour les véhicules connectés par analyse des utilisations anormales. Partenaires : IFSTTAR, LaMiH, IEMN-Telice, IEMN-DOAE, IMT Lille Douai, CRIStAL. Financement : CPER/FEDER.

Depuis 2016 : Projet Shift2Rail, WP8 (Cyber-Sécurité). Sujet : Analyse de l'utilisation des infrastructures informatiques dans les systèmes de transport ferroviaire pour détecter des utilisations anormales et des attaques du réseau. Partenaires : IRT Railenium (dont fait partie IMT Lille Douai), Alsthom, AZD, ASTS. Financement : H2020 (Europe).

Depuis Septembre 2015 : Projet Analyse Inertielle Smartphone. Sujet : Analyse de signaux inertiels issus d'un smartphone pour la détermination d'indicateurs sur une personne. Partenaires : Orange Labs, Grenoble, INL, LAAS, IRIT, IMT Lille Douai. Financement : fonds propres des laboratoires.

Mai 2014 – Mai 2017 : Projet REPAR (Représentation Parcimonieuse et Apprentissage Dynamique pour le Signal et l'Image). Sujet : Représentation et classification dans le cadre de données parcimonieuses. Partenaires : LISIC, CRIStAL, IMT Lille Douai, GEMTEX. Financement : région Hauts de France.

Depuis 2011 : Projets iBalance/iProprio. Sujet : Application et jeux sérieux pour l'entrainement et la rééducation de capacités posturales et proprioceptives. Partenaire : Université de Grenoble Alpes. Financement : fonds propres.

C.3 Jury de thèse

17 oct. 2018 : Examinateur de la thèse d'Halim Tannous (Université de Technologie de Compiège). Sujet : Interactive and connected rehabilitation systems for e-health. Direction : Dr. Dan Istrate et Dr. Tien Tuan Dao, UTC et Pr. Aziz Benlarbi-Delai, Université Pierre et Marie Curie.

14 Déc. 2017 : Examinateur de la thèse d'Alexis Brenon (Université Grenoble Alpes). Sujet : Modèle profond pour le contrôle vocal adaptatif d'un habitat intelligent. Direction : Dr. Michel Vacher, LIG et Dr. François Portet, LIG.

20 Sept. 2016 : Examinateur de la thèse de Loïc Sévrin (Université de Lyon 1). Sujet : Mesure et suivi de l'activité de plusieurs personnes dans un Living Lab en vue de l'extraction d'indicateurs de santé et de bien-être. Direction : Pr. Nacer Abouchi, CPE Lyon et Pr. Norbert Noury, INSA Lyon.

C.4 Distinctions

Janvier 2016 : Relecteur de l'année 2015 pour la revue IEEE JBHI, de la société EMBS.

Octobre 2015 : Obtention du prix Silver Talent de la région Lorraine pour les travaux de recherche iBalance.

C.5 Activités de recherche

Mes activités de recherche s'étendent de l'analyse du signal jusqu'à la classification des données en passant par l'extraction et la sélection de paramètres. Les modèles créés, qu'ils soient statiques ou dynamiques, sont utilisés dans diverses applications : l'amélioration de la qualité de vie et du quotidien dans un habitat, la surveillance de personne à domicile et la continuité de certains soins dans cet environnement, la sécurité et plus précisément la cyber-sécurité en utilisant des méthodes de classification ou encore l'analyse de données pour l'usine du futur. Ces travaux sont résumés dans les sous-sections suivantes.

Avant toute application, l'enjeu majeur de ces travaux de recherche est d'améliorer notre connaissance sur les différentes phases du traitement de la donnée, que ce soit en terme d'efficacité et de temps de calcul sur des volumes importants, de généralisation des algorithmes à partir de l'apprentissage sur peu de données ou encore de pertinence des résultats

obtenus en fonction de standards déjà existants. Chaque domaine d'application se pose ensuite des problèmes précis, que nous abordons lors des différents projets de recherche présentés précédemment et nos algorithmes sont alors améliorés et adaptés à ces données afin d'évaluer le gain apporté par l'utilisation de nos techniques.

C.5.1 Verrous scientifiques

Les verrous scientifiques qui sont adressés dans ces travaux sont les suivants :

— Extraire une information pertinente à partir de signaux multi-modaux et fortement non stationnaires, afin d'en déduire des informations ou des indices sur une application précise que nous adressons. Ces informations devront être multi-échelles afin de récupérer toute la richesse du signal de départ. Les attributs ou indices que nous obtenons devront être simples et donner de bonnes performances ensuite par rapport à un problème donné.

— Construire des modèles adaptés et adaptables pour des phénomènes complexes, et ce afin de déterminer des informations sur des comportements suivis.

— Création ou modification d'algorithmes de classification afin de les rendre le plus efficace possible pour modéliser ces phénomènes non stationnaires.

— Être capable d'utiliser les données extraites ou les classification effectuées afin de récupérer des informations sur le plus long terme et d'adapter des comportements ou des réponses dans certaines applications.

Ces verrous sont abordés sur plusieurs problèmes avec des applications précises et en lien avec des problèmes qui se posent actuellement.

C.5.2 Extraction et traitement de signaux : application à la médecine personnelle

Avant toute classification, il nous faut obtenir les signaux. Pour ceci, un travail d'implémentation et d'adaptation au matériel est nécessaire. Ce travail est fait par exemple dans le cadre de la mesure et la rééducation de personnes à l'aide de Smartphones. Dans ce travail, un ensemble de signaux est acquis en fonction du matériel et traité afin d'en obtenir la position d'une partie du corps humain [JIF11 ; JIF7]. Ce calcul doit être robuste (validation à l'aide d'une référence absolue dont l'erreur est connue) [JIF8] et rapide (le matériel est de plus en plus puissant mais doit tout de même suivre une fréquence d'échantillonnage suffisante afin d'avoir des résultats satisfaisants). Les traitements sont ensuite de type statistiques afin de pouvoir identifier la personne dans une population et lui proposer des exercices adaptés. Les résultats de ces travaux donnent une bonne identification des groupes de population dans le cadre d'exercices posturaux et une bonne utilisation du retour sensoriel créé sur le téléphone grâce à ces mesures [IC18]. En plus de l'application à la rééducation ou à l'entrainement du bon maintien de la posture, des travaux ont été faits dans le cadre de la thèse de Quentin Mourcou sur l'apprentissage et la rééducation de certains mouvements en toute autonomie (genoux, épaule, etc.) [IC8] et des travaux dans le cadre de la marche sont en cours.
D'autres traitements, axés temps-échelles, sont également utilisés dans les travaux réalisés ces dernières années, afin toujours d'identifier des problèmes sur une personne ou de catégoriser une personne par rapport à une population et ainsi détecter de manière précoce de possibles symptômes. Ces travaux, qui ont été réalisés dans le cadre de la thèse de Céline Franco, ont porté sur des techniques telles que la décomposition empirique modale ou le synchrosqueezing, afin d'analyser finement les signaux multidimensionnels obtenus [IC22]. Des analyses d'entropie ont également été menées dans cette thèse notamment sur les signaux posturaux, afin d'analyser les performances de la personnes dans des tâches diverses [IC2 ; NC2].
Pour compléter ces travaux et se diriger de plus en plus vers l'évaluation, par un indice pertinent, de la qualité de vie d'une personne dans son environnement et de manière la moins intrusive possible, la thèse de Sid Ahmed Walid Talha, débutée en octobre 2016, reprend les travaux de Céline Franco et les continue. Le but est d'évaluer un indicateur, sur les données smartphones recueillies durant la journée, de la qualité de vie de la personne et de son « état général ». Ses travaux se concentrent pour le moment sur une extraction et une compréhension plus fine des périodes d'activités/d'inactivités de la personne (détection des postures, caractérisation de la marche etc.), avant de se diriger par la suite vers une caractérisation de l'évolution de ces périodes.

C.5.3 Extraction, sélection d'attributs et classification statique/dynamique

C.5.3.1 Appliquée à l'habitat intelligent et à la reconnaissance d'activités

Ces travaux font suite à mes travaux de thèse. À partir d'un ensemble de capteurs plus ou moins informatifs lorsque considérés un à un, nous tentons d'inférer le comportement d'une personne à son domicile et ainsi d'obtenir les activités qu'elle effectue au cours de la journée. Ceci nous permet de suivre l'évolution de cette personne au cours du temps avec des applications dans deux domaines. Le premier est la domotique adaptée (lors du lancement d'une commande par l'utilisateur, le système répondra différemment en fonction de l'état et de l'activité courante de la personne et des préférences apprises au cours du temps de celle-ci) [IC9 ; IJ1], mais aussi et surtout la détection de symptômes apparaissant chez la personne âgée de manière précoce et l'évaluation en continue de son autonomie.
Ces travaux ont donné lieu à de nombreuses publications, que ce soit sur le corpus que nous avons diffusé [JIF10] ou sur les résultats de classification avec différentes méthodes (SVM, MLN, Naive Bayes, etc.) [JIF12 ; IJ5 ; IC21 ; JIF6]. Les résultats sont pour le moment encourageants et permettent de correctement identifier de nombreuses activités, certaines restant tout de même difficiles à différencier du fait des mesures erronées des capteurs pas toujours corrigées par d'autres capteurs ou bien de la manière pour la personne de réaliser ces activités. En effet, travailler sur les activités humaines est complexe du fait de la grande variabilité des exécutions. Avec la thèse de Nawel Yala, de l'université d'Alger et son encadrant le Professeur Belkacem Fergani, nous nous dirigeons vers une utilisation des algorithmes de classification incrémentaux pour la classification des activités de la vie quotidienne de manière plus fine et prenant en compte les évolutions dans la manière de réaliser l'activité [IC14 ; JIF5].
Enfin, dernièrement, dans le projet COMRADES [IC5 ; IC4] des travaux sur cet habitat et sur le moyen d'inclure de la robotique mobile à l'intérieur afin d'améliorer le confort de l'utilisateur ont été démarrés, en collaboration avec IMT Atlantique.

C.5.3.2 Applications à la sécurité et à l'usine du futur

Les travaux sur la modélisation dynamique des données sont utilisés sur d'une part la classification d'activités de la vie quotidienne, comme précédemment présenté mais en faisant évoluer cette fois les modèles de classes, mais aussi dans des applications différentes liées à la sécurité. Ainsi, la thèse de Yanyun Lu a porté sur l'analyse de vidéos émanant par exemple d'un système de vidéo surveillance, dans lequel plusieurs caméras avec des caractéristiques différentes se chargent de l'acquisition des images. Le but était alors de construire des modèles évolutifs pour l'identification [JIF9] et la ré-identification [IC11] de personnes dans de tels systèmes d'acquisition. D'autres travaux concernant l'analyse de modifications de comportement d'utilisation pour la sécurité ou sur l'évolution de paramètres mesurés pour la prédiction d'anomalies sont en cours ou démarreront prochainement.
Ces travaux sur la reconnaissance dans la vidéo ont été poursuivis avec les mémoires de masters de Mounir Hammouche et Sid Hamed Walid Talha [JIF4] qui se sont intéressés à la reconnaissance en ligne d'actions à partir des signaux issus de la Kinect avec des méthodes soit basées sur des classifieurs multiples [IC10] soit sur des chaînes de Markov en proposant un nouvel attribut [IC6].
Toujours sur un aspect sécurité, le travail dans les projets SECOURT et X2RAIL-1 appliquera et complètera nos travaux de recherche à la sécurité des réseaux de communication et à la détection d'utilisations frauduleuses de certains réseaux de communications sensibles (réseaux véhicules à véhicules pour gérer les situations d'urgence ou réseaux ferroviaires) [IC3]. Ces travaux s'effectueront notamment dans le cadre du postdoctorat de Jonathan Villain dans lequel nous tenterons d'appliquer nos techniques de reconnaissance de formes et de classification afin d'isoler des attaques spécifiques sur ces réseaux, le tout en s'intégrant à un framework d'analyse d'utilisation des réseaux de communications développé à l'IFSTTAR. Pour notre part nous travaillerons sur les couches hautes du modèle OSI (couches applications), l'IFSTTAR travaille déjà sur des couches plus basses (physiques).
Enfin, nos travaux peuvent aussi être utilisés (et le sont ou l'ont déjà été) dans le cadre de partenariat avec des entreprises ou des laboratoires d'autres domaines. Ainsi, nous avons identifié des groupes et des comportements de personnes, à partir de questionnaires et d'examens d'évaluations du code de la route, dans une population devant repasser son permis de conduire. Cela a donné lieu à un stage de M1 et une collaboration. Nous travaillons aussi en ce moment

avec e.l.m. Leblanc pour identifier des causes de retours en service après vente à partir des données de chaînes de production. Tous ces travaux font intervenir nos connaissances d'une part en extraction d'information et d'autre part en classification et création de modèles génériques.
Un sous-ensemble de ces travaux est présenté en détail dans la première partie de ce manuscrit.

Annexe D

Implication dans la communauté scientifique et éducative

Sommaire

C.1 Mots clés 113
C.1.1 Recherche 113
C.1.2 Applications 113
C.2 Résumé des activités 114
C.2.1 Formation à la recherche 114
C.2.2 Formation par la recherche 114
C.2.3 Principales participations à des projets de recherche 115
C.3 Jury de thèse 116
C.4 Distinctions 116
C.5 Activités de recherche 116
C.5.1 Verrous scientifiques 117
C.5.2 Extraction et traitement de signaux : application à la médecine personnelle 117
C.5.3 Extraction, sélection d'attributs et classification statique/dynamique 118

Les fonctions d'enseignant-chercheur amènent à participer activement à l'animation d'une communauté scientifique en organisant des évènements, en évaluant nos pairs, etc. Cette partie du curriculum vitæ va revenir sur différentes actions dans ce domaine..

D.1 Activités en lien avec la communauté scientifique

Décembre 2017 : Édition d'un numéro spécial pour *Multimodal Technologies and Interaction* (MDPI). Co-Editeur : Dr. Abir-Béatrice Karami Lozenguez. Titre : Multimodal Interactions and Smart Homes. Parution prévue courant 2018.

Octobre 2017 : Co-organisation de la journée de l'Institut Mines Télécom « Objet connectés en Santé ». Co-organisateurs : Juliette Lebreton et Estelle Fieve (DRI IMT), Christine Balagué (Télécom Ecole de Management), Bernadette Dorizzi (Télécom Sud Paris). Rôle : gestion d'une session et d'une table ronde.

Octobre 2016 : Vice-chair, Conférence IPAS 2016, Tunis.

Juillet 2016 : Évaluation d'un projet de recherche sur la substitution sensorielle et la robotique pour la mairie de Paris.

Janvier 2016 : Édition d'un numéro spécial pour Evolving Systems (Springer). Co-Editeurs : Pr. Moamar Sayed-Mouchaweh et Dr. Edwin Lughofer. Titre : Adaptive Intelligent Systems for Learning, Control and Optimization. Paru en février 2017.

Janvier 2016 : Évaluation de propositions de thèses région pour la région Rhônes-Alpes dans le domaine de la dépendance.

Décembre 2015 : Participation en tant que membre du comité d'organisation IEEE EAIS 2015 (Evolving and Adaptive Intelligent Systems). Rôles : responsable de l'organisation locale, co-editeur des actes de la conférence.

Mars 2015 : Évaluation de 5 projets de recherche pour le ministère de la santé Italien

Depuis Sept. 2013 : Membre de l'Editorial Board de la revue Journal of Ambient Intelligence and Smart Environement (IF 2015 : 1,063)

Depuis 2011 : Évaluation d'un ou deux projets de recherche par an pour le Conseil de recherches en sciences naturelles et en génie du Canada.

Depuis 2008 : Relecture d'articles pour des différentes revues internationales (IEEE Transactions on Information Technology in Biomedicine, IEEE Transactions on Reliability, Evolving Systems, ACM Computing Survey, IEEE System Man and Cybernetics part C, EURASIP Journal on Advances in Signal Processing, IEEE Transactions on Automation Science and Engineering, BioMedical Engineering OnLine, Sensors, etc.)

Depuis 2008 : Relecture d'articles pour des conférences internationales (e-Health, IEEE HealthCom, IEEE EMBC, pHealth, ICMLA, ISIAE, ICSPDM, IINTEC, etc.)

Depuis 2008 : Participation en tant que membre de différents comités de techniques programmes de conférences : eHalth, IEEE HealthCom, pHealth, RAMI, StreamEvolv-2016, SAI 2017, CICN 2018.

D.2 Appartenances à des groupements scientifiques

Depuis 2015 : Membre de l'axe transversal Santé, Autonomie et Qualité de vie de l'Institut Mines Télécom et au Réseau Thématique Santé (RT6) de l'institut. Représentant d'IMT Lille Douai dans les instances.

Depuis 2014 : Membre de la SFR Technologie pour la Santé et le Médicament, créée et supportée par le CIC-IT de Lille et l'Université de Lille 2.

Depuis 2009 : Membre du GDR ISIS.

Depuis 2005 : Membre IEEE.

D.3 Évaluation de projets

Que ce soit dans les thématiques liées à la santé et à l'autonomie ou dans les thématiques liées à la classification ou au traitement du signal, j'ai eu l'occasion de faire de nombreuses relectures de projets (pour le Canada - le CRNSG depuis 2014, la France - Région Rhônes-Alpes-Auvergne depuis 2016, l'Iltalie - Ministero de la Salute depuis 2016). Pour les projets, je participe à l'évaluation en moyenne de 6 à 8 projets par an.

D.4 Relectures, comités éditoriaux

Je suis également solicité fréquemment pour la relecture d'articles scientifiques pour de nombreuses revues.
Ceci m'a même conduit à être nommé relecteur de l'année 2015 (6 primés) de la société IEEE EMBS. Pour la relecture d'articles, j'en effectue entre 5 et 10 par an pour des revues et une dizaine pour des conférences.
Je suis également membre du comité éditorial de la revue Journal of Ambient Intelligence and Smart Environment (JAISE) depuis 2013, gérant entre 5 et 7 articles (avec leurs re-soumissions possibles) par an.

D.5 Organisations de conférences et de sessions

Récemment (Octobre 2017), nous avons organisé, avec l'Institut Mines Télécom, une journée technique sur les objets connectés en santé dans laquelle j'ai mis en place une session et une table ronde. Cette journée a recensé 200 inscrits (dont environ 80 industriels) pour environ 150 présents le jour même.
En 2015, nous avons organisé, à IMT Lille Douai, la conférence IEEE Evolving and Adaptive Intelligent Sytems (EAIS) pour laquelle j'étais organisateur local et pour laquelle j'ai également édité les actes (avec mise au format et création des supports pour la diffusion chez IEEE) .

J'ai également été membre de comité de programmes de plusieurs conférences (e-Health, pHealth, etc.) et suis l'un des Vice-Chair de la conférence IEEE IPAS 2016.

D.6 Participation à des réseaux scientifiques Santé

Pour ce qui est de l'axe santé, je représente IMT Lille Douai dans les instances de l'Institut Mines Télécom sur la thématique santé. Cela est valable dans le comité de pilotage de la thématique santé mais également dans le réseau thématique Santé Numérique de l'institut (RT6).
Le travail sur la santé m'a également amené à participer aux réunions de mise en place de la Structure Fédérative de Recherche Technologies pour la santé et Médicament portée par l'université de Lille 2.

D.7 Collaborations suivies nationales et internationales

Mes activités de recherche m'ont amené à collaborer de manière durable avec les équipes suivantes :

— Laboratoire d'Informatique de Grenoble, collaboration avec Dr Michel Vacher et Dr. François Portet.

— Laboratoire AGIM, Grenoble, collaboration avec Dr. Nicolas Vuillerme

— Laboratoire de Communication Parlée et Traitement de Signaux (LCPTS), Université d'Alger, collaboration avec Pr. Belkacem Fergani

— Orange Labs Grenoble, dans le cadre d'un projet inertiel, collaboration sur la durée avec également INL (Pr. Norbert Noury), LAAS (Pr. Eric Campo) et IRIT (Dr. Remy Bastide).

— Telecom Sud Paris, laboratoire SAMOVAR, avec Pr. Jérôme Boudy et Dr. Nesma Houmani.

D'autres collaborations, avec des laboratoires Allemand, Néérlandais, Néo-Zélandais ou d'Amérique du Sud se font de manière ponctuelle pour des projets ou des rédactions d'articles.

D.8 Vie de l'équipe et de l'école

Une dernière tâche consiste en la participation aux diverses activités de l'école et du laboratoire. Ainsi, nous avons tous déjà participé à la dernière évaluation AERES et la prochaine évaluation (HCERES) aura lieu en 2018. La rédaction du dossier est ainsi menée collectivement.
En ce qui concerne la vie de l'école, nous participons aussi chaque année aux recrutement des étudiants arrivant en L3 dans notre école, en plus de gérer les recrutements de nos doctorants et post-doctorants.

Annexe E

Indices de production et liste des publications

Sommaire

D.1 Activités en lien avec la communauté scientifique . . . 121
D.2 Appartenances à des groupements scientifiques . . . 122
D.3 Évaluation de projets . . . 122
D.4 Relectures, comités éditoriaux . . . 122
D.5 Organisations de conférences et de sessions . . . 122
D.6 Participation à des réseaux scientifiques Santé . . . 123
D.7 Collaborations suivies nationales et internationales . . . 123
D.8 Vie de l'équipe et de l'école . . . 123

E.1 Indices de production

E.1.1 Résumé des contributions

Type	Nombre
Brevets 2	
Dépôts APP	3
Journaux internationaux parus	24 dont 17 avec IF JCR
Chapitres d'ouvrages	4
Journaux nationaux	3
Conférences internationales	40 dont 3 rang A
Conférences nationales	12
Journées techniques	10
Séminaires invités	3
Article de vulgarisation	1
Mémoires	2
Éditions	3
Bases de données publiques	3
Masters encadrés	3
Thèses co-encadrées	4 dont 1 en cours [1]

E.1.2 Bibliométrie et citations

Cette partie présente des éléments de bibliométrie récupérés d'une part sur Google Scholar et aussi sur ISI Web of Knowledge. Ces éléments sont pris au 30 septembre 2018.

1. Antoine Ghorra avait débuté une thèse en octobre 2016 en plus de ceci. Cependant, il a décidé, ayant changé ses ambitions en terme de carrières, d'arrêter et de retourner dans une entreprise de développement informatique.

Mesure	Google Scholar	ISI Web of Knowledge
Nombre de citations	1887	629
H-Index	21	10

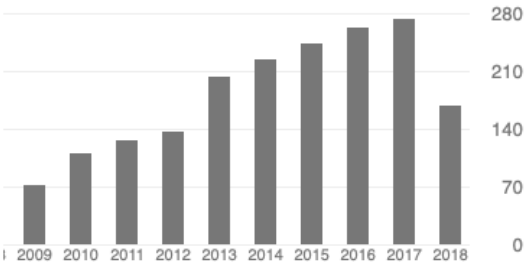


FIGURE E.1 – Nombre de citations pour les différentes années, Google Scholar

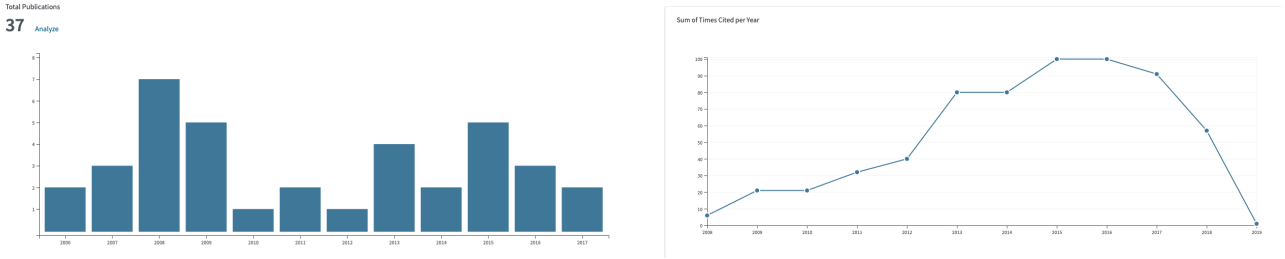


FIGURE E.2 – Nombre de publications référencées (gauche) et nombre de citations dans des publications référencées pour les différentes années (droite) sur ISI Web Of Knowledge

E.2 Liste des publications

Brevets

[P1] **Fleury, A.**, Vuillerme, N., Diot, B., Hlavackova, P., (28 sept. 2017). « Procédé et système pour l'estimation d'une demande ou d'un coût attentionel associé à l'éxecution d'une tâche ou des stratégies de partage de l'attention développées par un individu ». FR 30 49371 / WO 2017/162789 A1.

[P2] **Fleury, A.**, Vuillerme, N., Franco, C., Mourcou, Q., Diot, B., (27 mai 2016). « Procédé et système pour la mesure, le suivi, le contrôle et la correction d'un mouvement ou d'une posture d'un utilisateur ». FR 14 61233 / WO 2016/079452 A1 / US Patent 20170354843.

Dépôts Logiciels (APP)

[APP1] **Fleury, A.**, Franco, C., Vuillerme, N., (2018). « AVQ_PRED (Logiciel conçu afin de conduire de manière automatique une analyse prédictive d'une série chronologique d'AVQ selon un modèle n-grams utilisé en modélisation du langage) ».

[APP2] **Fleury, A.**, Mourcou, Q., Franco, C., Vuillerme, N., (2018a). « Clasp Your Age (Logiciel d'évaluation personnalisée des capacités de génération et de contrôle de la force de préhension palmaire) ».

[APP3] — (2018b). « IHMBalance (Interface graphique développée sous Matlab conçue afin de conduire de manière automatique une analyse quantitative, objective et fiable de différents signaux posturaux) ».

Journaux Internationaux (avec IF JCR)

[JIF1] Gomez Montenegro, C., Chessa, S., **Fleury, A.**, Roussos, G., Preuveneerse, D., (2019). « Internet of things : trends and comparison of communication technologies and solutions for enabling smart environments ». In : *Journal of Ambient Intelligence and Smart Environments* 11.1. DOI : 10.3233/AIS-180509. ***[IF Courant : 0.878]***.

[JIF2] Talha, S. A. W., **Fleury, A.**, Lecoeuche, S., (2019). « Position-invariant recognition of human activities and postural transitions using smartphones ». In : *Sensors*. (***Soumis***).

[JIF3] Villain, J., Deniau, V., **Fleury, A.**, Simon, P. E., Gransart, C., Kousri, R., (2019). « EM monitoring and classification of IEMI and protocol-based attacks on IEEE 802.11n communication networks ». In : *IEEE Transactions on Electromagnetic Compatibility*. DOI : 10.1109/TEMC.2019.2900262. ***[IF Courant : 1.52]***.

[JIF4] Talha, S. A. W., Hammouche, M., Ghorbel, E., **Fleury, A.**, Ambellouis, S., (déc. 2018). « Features and classification schemes for view-invariant and real-time human action recognition ». In : *IEEE Transaction on Cognitive and Developmental Systems* 10.4, p. 894-902. DOI : 10.1109/TCDS.2018.2844279. ***[IF Courant : 1.952]***.

[JIF5] Yala, N., Fergani, B., **Fleury, A.,** (avr. 2017). « Towards improving feature extraction and classification for activity recognition on streaming data ». In : *Journal of Ambient Intelligence and Humanized Computing* 8.2, p. 177-189. DOI : 10.1007/s12652-016-0412-1. ***[IF Courant : 1.588]***.

[JIF6] Chahuara, P., **Fleury, A.**, Portet, F., Vacher, M., (juil. 2016). « On-line human activity recognition from audio and home automation sensors : comparison of sequential and non-sequential models in realistic smart homes ». In : *Journal of Ambient Intelligence and Smart Environments* 8.4, p. 399-422. DOI : 10.3233/AIS-160386. ***[IF 2016 : 0,707 ; IF Courant : 0.878]***.

[JIF7] Mourcou, Q., **Fleury, A.**, Diot, B., Franco, C., Vuillerme, N., (2015). « Mobile phone-based joint angle measurement for functional assessment and rehabilitation of proprioception ». In : *BioMed Research International* 2015.2015, Article 328142, 15 p. DOI : 10.1155/2015/328142. ***[IF 2015 : 2.134 ; IF Courant : 2.476]***.

[JIF8] Mourcou, Q., **Fleury, A.**, Franco, C., Klopcic, F., Vuillerme, N., (2015). « Performance evaluation of smartphone inertial sensors measurement for range of motion ». In : *Sensors* 15.9, p. 23168-23187. DOI : 10.3390/s150923168. ***[IF 2015 : 2.033 ; IF Courant : 2.677]***.

[JIF9] Lu, Y., Boukharouba, K., Boonært, J., **Fleury, A.**, Lecoeuche, S., (2014). « Application of an incremental SVM algorithm for on-line human recognition from video surveillance using texture and color features ». In : *Neurocomputing* 126, p. 132-140. DOI : 10.1016/j.neucom.2012.08.071. ***[IF 2014 : 2.083 ; IF Courant : 2.392]***.

[JIF10] **Fleury, A.**, Vacher, M., Portet, F., Chahuara, P., Noury, N., (2013). « A french corpus of audio and multimodal interactions in a health smart home ». In : *Journal on Multimodal User Interfaces* 7.1-2, p. 93-109. DOI : 10.1007/s12193-012-0104-x. ***[IF 2013 : 0.462 ; IF Courant : 1.423]***.

[JIF11] Franco, C., **Fleury, A.**, Guméry, P.-Y., Diot, B., Demongeot, J., Vuillerme, N., (2013). « iBalance-ABF : a smartphone-based audio-biofeedback balance system ». In : *Biomedical Engineering, IEEE Transactions on* 60.1, p. 211-215. DOI : 10.1109/TBME.2012.2222640. ***[IF 2013 : 2.233 ; IF Courant : 4.288]***.

[JIF12] **Fleury, A.**, Vacher, M., Noury, N., (2010). « SVM-based multimodal classification of activities of daily living in health smart homes : sensors, algorithms, and first experimental results ». In : *Information Technology in Biomedicine, IEEE Transactions on* 14.2, p. 274-283. DOI : 10.1109/TITB.2009.2037317. ***[IF 2010 : 1.707 ; IF Courant : - (devenu IEEE JHBI IF: 3.451)]***.

[JIF13] Noury, N., **Fleury, A.**, Nocua, R., Poujaud, J., Gehin, C., Dittmar, A., Delhomme, G., Demongeot, J., McAdam, E., (2009). « Ehealth sensors. biomedical sensors, algorithms and sensors networks ». In : *IRBM* 30.3, p. 93-103. ***[IF 2009 : 0.295 ; IF Courant : 0.628]***.

[JIF14] Vuillerme, N., Pinsault, N., Chenu, O., **Fleury, A.**, Payan, Y., Demongeot, J., (2009). « A wireless embedded tongue tactile biofeedback system for balance control ». In : *Pervasive and Mobile Computing* 5.3, p. 268-275. DOI : 10.1016/j.pmcj.2008.04.001. ***[IF 2009 : - ; IF Courant : 1.719]***.

[JIF15] Vuillerme, N., Chenu, O., Pinsault, N., **Fleury, A.**, Demongeot, J., Payan, Y., (2008). « Can a plantar pressure–based tongue-placed electrotactile biofeedback improve postural control under altered vestibular and neck proprioceptive conditions ? » In : *Neuroscience* 155.1, p. 291-296. DOI : 10.1016/j.neuroscience.2008.05.018. ***[IF 2008 : 3.556 ; IF Courant : 3.231]***.

[JIF16] Vuillerme, N., Pinsault, N., Chenu, O., **Fleury, A.**, Payan, Y., Demongeot, J., (2008a). « Postural destabilization induced by trunk extensor muscles fatigue is suppressed by use of a plantar pressure-based electro-tactile biofeedback ». In : *European journal of applied physiology* 104.1, p. 119-125. DOI : 10.1007/s00421-008-0768-9. ***[IF 2008 : 1.931 ; IF Courant : 2.328]***.

[JIF17] Vuillerme, N., Pinsault, N., **Fleury, A.**, Chenu, O., Demongeot, J., Payan, Y., Pavan, P., (2008). « Effectiveness of an electro-tactile vestibular substitution system in improving upright postural control in unilateral vestibular-defective patients ». In : *Gait & posture* 28.4, p. 711-715. DOI : 10.1016/j.gaitpost.2008.05.017. ***[IF 2008 : 2.2743 ; IF Courant : 2.286]***.

[JIF18] Vuillerme, N., Chenu, O., Pinsault, N., **Fleury, A.**, Demongeot, J., Payan, Y., (2007). « A plantar-pressure based tongue-placed tactile biofeedback system for balance improvement ». In : *Computer Methods in Biomechanics and Biomedical Engineering* 10.sup1, p. 63-64. DOI : 10.1080/10255840701480113. ***[IF 2007 : 0.572 ; IF Courant : 1.85]***.

Journaux Internationaux

[IJ1] Karami, A. B., **Fleury, A.**, Boonaert, J., Lecoeuche, S., (juin 2016). « User in the loop : adaptive smart homes exploiting user feedback — state of the art and future directions ». In : *Information* 7.2, p. 35. ISSN : 2078-2489. DOI : 10.3390/info7020035.

[IJ2] Bilal Abidine, M., Fergani, L., Fergani, B., **Fleury, A.,** (2015). « Improving human activity recognition in smart homes ». In : *International Journal of E-Health and Medical Communications (IJEHMC)* 6.3, p. 19-37.

[IJ3] Mourcou, Q., **Fleury, A.**, Franco, C., Vuillerme, N., (2015). « Wegoto : a smartphone application to assess the accessibility of roads and public spaces for people with disabilities – a proof of concept study ». In : *Modelling C, Association for the advancement of modeling and simulation techniques in entreprises (AMSE)* 75.2, p. 165-176.

[IJ4] Noury, N., Barralon, P., Vuillerme, N., **Fleury, A.,** (2012). « Fusion of multiple sensors sources in a smart home to detect scenarios of activities in ambient assisted living ». In : *International Journal of E-Health and Medical Communications (IJEHMC)* 3.3, p. 29-44. DOI : 10.4018/jehmc.2012070103.

[IJ5] **Fleury, A.**, Noury, N., Vacher, M., (2011). « Improving supervised classification of activities of daily living using prior knowledge ». In : *International Journal of E-Health and Medical Communications (IJEHMC)* 2.1, p. 17-34. DOI : 10.4018/978-1-4666-2794-9.ch008.

[IJ6] Vacher, M., Portet, F., **Fleury, A.**, Noury, N., (2011). « Development of audio sensing technology for ambient assisted living : applications and challenges ». In : *International Journal of E-Health and Medical Communications (IJEHMC)* 2.1, p. 35-54. DOI : 10.4018/978-1-4666-2794-9.ch009.

[IJ7] **Fleury, A.**, Vacher, M., Glasson, H., Noury, N., Serignat, J.-F., (2008). « Data fusion in health smart home : primary results of individual evaluation of two sensors ». In : *Gerontechnology* 7.2, p. 105. DOI : 10.4017/gt.2008.07.02.042.00.

[IJ8] Pinsault, N., **Fleury, A.**, Virone, G., Bouvier, B., Vaillant, J., Vuillerme, N., (2008). « Test-retest reliability of cervicocephalic relocation test to neutral head position ». In : *Physiotherapy theory and practice* 24.5, p. 380-391. DOI : 10.1080/09593980701884824.

Chapitres d'ouvrages

[BC1] GhaffarianHoseini, A., GhaffarianHoseini, A., Tookey, J., Omrany, H., **Fleury, A.**, Naismith, N., Ghaffarianhoseini, M., (mar. 2016). « Handbook of research on creative technologies for multidisciplinary applications ». In : sous la dir. d'A. CONNOR et S. MARKS. IGI Global. Chap. The Essence of Smart Homes : Application of Intelligent Technologies towards Smarter Urban Future, 334-376 (ch. 14).

[BC2] Franco, C., Vuillerme, N., Diot, B., Demongeot, J., **Fleury, A.,** (2015). « Multisensor data fusion ». In : sous la dir. de H. FOURATI. CRC Press. Chap. Data Fusion for Telemonitoring : Application to Health and Autonomy, p. 535-548.

[BC3] Noury, N., Poujaud, J., **Fleury, A.**, Nocua, R., Haddidi, T., Rumeau, P., (2011). « Activity recognition in pervasive intelligent environments ». In : sous la dir. de L. CHEN, C. D. NUGENT, J. BISWAS et J. HOEY. Atlantis Press. Chap. Smart Sweet Home... A Pervasive Environment for Sensing our Daily Activity?, p. 187-208.

[BC4] Vacher, M., **Fleury, A.**, Portet, F., Serignat, J.-F., Noury, N., (2010). « New developments in biomedical engineering ». In : sous la dir. de D. CAMPOLO. In-Tech. Chap. Complete sound and speech recognition system for health smart homes : application to the recognition of activities of daily living, p. 645-673.

Journaux nationaux

[NJ1] Jolles-Haeberli, B., Bassin, J.-P., Pichonnaz, C., Duc, C., **Fleury, A.**, Aminian, K., Djahangiri, A., Farron, A., (2009). « Un instrument qui mesure les mouvements de l' épaule avec précision – ein instrument das die schulterbewegungen präzise misst ». In : *Physioactive* 5, p. 9-16.

[NJ2] Duc, C., Farron, A., Coley, B., **Fleury, A.**, Aminian, K., Pichonnaz, C., Bassin, J. P., Jolles, B. M., (1er juin 2009). « A new way of quantifying dominant upper-limb mobility in healthy and painful shoulders ». In : *Swiss Medical Weekly* 139 (Suppl 173), S29.

[NJ3] Noury, N., Villemazet, C., **Fleury, A.**, Rumeau, P., Rialle, V., Ghorayeb, A., Baghai, R., (2008). « Systèmes d'information pour le télémonitorage de la santé à domicile : implémentation du système ailisa s' appuyant sur une messagerie électronique ». In : *TSI. Technique et science informatiques* 27.9-10, p. 1287-1314.

Conférences Internationales avec Actes et Comité de Lecture

[IC1] **Fleury, A.**, Mourcou, Q., Franco, C., Vuillerme, N., (oct. 2018c). « Implementation and evaluation of a smartphone app to assess attentional demand for balance control ». In : *Engineering in Medicine and Biology Society, 2018. EMBS 2018. 40th Annual International Conference of the IEEE*, p. 5598-5601. DOI : 10.1109/EMBC.2018.8513540.

[IC2] Franco, C., **Fleury, A.**, Diot, B., Vuillerme, N., (oct. 2018). « Applying entropy to human center of foot pressure data to assess attention investment in balance control ». In : *Engineering in Medicine and Biology Society, 2018. EMBS 2018. 40th Annual International Conference of the IEEE*, p. 5586-5589. DOI : 10.1109/EMBC.2018.8513533.

[IC3] Gransart, C., Deniau, V., Simon, É. P., **Fleury, A.**, Lecoeuche, S., Millot, P., Dayoub, I., Berbineau, M., Masson, É., (16 avr. 2018). « Cyber security of the railway wireless system : detection, decision and human-in-the-loop ». In : *TRA (Transport Research Arena) 2018.*

[IC4] Papadakis, P., Lohr, C., Lujak, M., Karami, A. B., Kanellos, I., Lozenguez, G., **Fleury, A.,** (31 jan. 2018). « System design for coordinated multi-robot assistance deployment in smart spaces. » In : *The Fourth IEEE Int. Workshop on Collaboration of Humans, Agents, Robots, Machines and Sensors (CHARMS 2018) in 2nd IEEE International Conference on Robotic Computing (IRC 2018).* DOI : 10.1109/IRC.2018.00068.

[IC5] Lujak, M., Bouraqadi, N., Doniec, A., Fabresse, L., **Fleury, A.**, Karami, A., Lozenguez, G., (14 déc. 2017). « Towards robots-assisted ambient intelligence. » In : *5th International Conference on Agreement Technologies 2017 (AT 2017).* DOI : 10.1007/978-3-030-01713-2_34.

[IC6] Talha, S. A. W., **Fleury, A.**, Ambellouis, S., (18 déc. 2017). « Human action recognition from body-part directional velocity using hidden markov models ». In : *16th IEEE International Conference On Machine Learning And Applications.* DOI : 10.1109/ICMLA.2017.00-14.

[IC7] Bilal Abidine, M., Fergani, B., **Fleury, A.,** (sept. 2017). « Integrating prior knowledge in weighted SVM for human activity recognition in smart home ». In : *IoT for Enhanced Quality of Life and Smart Living, Proceedings of the 15th International Conference on Smart Homes and Health Telematics, ICOST 2017.* Springer International Publishing, 6p.

[IC8] Mourcou, Q., **Fleury, A.**, Diot, B., Vuillerme, N., (août 2016). « iProprio : a smartphone-based system to measure and improve proprioceptive function ». In : *Engineering in Medicine and Biology Society, 2016. EMBS 2016. 38th Annual International Conference of the IEEE*, p. 2622-2625. DOI : 10.1109/EMBC.2016.7591268.

[IC9] Karami, A. B., **Fleury, A.,** (juin 2016). « Using feedback in adaptive and user-dependent one-step decision making ». In : *IJCAI-16, 25th International Joint Conference on Artificial Intelligence, Workshop Interactive Machine Learning*, p. ID 5.

[IC10] Hammouche, M., Ghorbel, E., **Fleury, A.**, Ambellouis, S., (fév. 2016). « Toward a real time view-invariant 3d action recognition ». In : *11th International Conference on Computer Vision Theory and Applications (VISAPP'16)*. DOI : 10.5220/0005843607450754.

[IC11] Lu, Y., **Fleury, A.**, Boonaert, J., Lecoeuche, S., Ambellouis, S., (2015). « Online person identification and new person discovery using appearance features ». In : *Evolving and Adaptive Intelligent Systems (EAIS), 2015 IEEE International Conference on*. IEEE, p. 1-8. DOI : 10.1109/EAIS.2015.7368794.

[IC12] Mourcou, Q., **Fleury, A.**, Franco, C., Diot, B., Vuillerme, N., (2015). « Smartphone-based system for sensorimotor control assessment, monitoring, improving and training at home ». In : *Inclusive Smart Cities and e-Health, Proceedings of the 13th International Conference on Smart Homes and Health Telematics, ICOST 2015*. Springer International Publishing, p. 141-151. DOI : 10.1007/978-3-319-19312-0_12.

[IC13] Toubakh, H., Sayed-Mouchaweh, M., **Fleury, A.**, Boonaert, J., (2015). « Hybrid dynamic data mining scheme for drift-like fault diagnosis in multicellular converters ». In : *Technological Advances in Electrical, Electronics and Computer Engineering (TAEECE), 2015 Third International Conference on*. IEEE, p. 56-61. DOI : 10.1109/TAEECE.2015.7113600.

[IC14] Yala, N., Fergani, B., **Fleury, A.,** (2015a). « Feature extraction and incremental learning to improve activity recognition on streaming data ». In : *IEEE International Conference on Evolving and Adaptive Intelligent Systems (EAIS)*. IEEE, p. 1-8.

[IC15] — (2015b). « Feature extraction for human activity recognition on streaming data ». In : *Innovations in Intelligent SysTems and Applications (INISTA), 2015 International Symposium on*. IEEE, p. 1-6. DOI : 10.1109/INISTA.2015.7276759.

[IC16] Franco, C., Gumery, P.-Y., **Fleury, A.**, Vuillerme, N., (2014a). « A new approach to wavelet entropy : application to postural signals ». In : *Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European*. IEEE, p. 2320-2324.

[IC17] Franco, C., Gumery, P.-Y., **Fleury, A.**, Vuillerme, N., (2014b). « Analyzing postural signals from a synchrosqueezing transform-based approach of wavelet entropy ». In : *4th International Congress on Complex Systems in Sports & Healthy Ageing*.

[IC18] **Fleury, A.**, Mourcou, Q., Franco, C., Diot, B., Demongeot, J., Vuillerme, N., (2013). « Evaluation of a smartphone-based audio-biofeedback system for improving balance in older adults-a pilot study ». In : *Engineering in Medicine and Biology Society, 2013. EMBS 2013. 35th Annual International Conference of the IEEE*. IEEE, p. 1198-1201. DOI : 10.1109/EMBC.2013.6609721.

[IC19] Franco, C., Diot, B., **Fleury, A.**, Demongeot, J., Vuillerme, N., (2013). « Ambient assistive healthcare and wellness management – is "the wisdom of the body" transposable to one's home? » In : *Inclusive Society : Health and Wellbeing in the Community, and Care at Home, proceedings of the 11th International Conference on Smart Homes and Health Telematics, ICOST 2013*. Springer Berlin Heidelberg, p. 143-150. DOI : 10.1007/978-3-642-39470-6_18.

[IC20] Mourcou, Q., **Fleury, A.**, Dupuy, P., Diot, B., Franco, C., Vuillerme, N., (2013). « Wegoto : a smartphone-based approach to assess and improve accessibility for wheelchair users ». In : *Engineering in Medicine and Biology Society, 2013. EMBS 2013. 35th Annual International Conference of the IEEE*. IEEE, p. 1194-1197. DOI : 10.1109/EMBC.2013.6609720.

[IC21] Chahuara, P., **Fleury, A.**, Portet, F., Vacher, M., (2012). « Using markov logic network for on-line activity recognition from non-visual home automation sensors ». In : *Ambient intelligence, proceedings of the Third International Joint Conference, AmI 2012*. Springer Berlin Heidelberg, p. 177-192. DOI : 10.1007/978-3-642-34898-3_12.

[IC22] Franco, C., Guméry, P.-Y., Vuillerme, N., **Fleury, A.**, Fontecave-Jallon, J., (2012). « Synchrosqueezing to investigate cardio-respiratory interactions within simulated volumetric signals ». In : *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*. IEEE, p. 939-943.

[IC23] Lu, Y., **Fleury, A.**, Booneart, J., Lecoeuche, S., (2011). « On-line human recognition from video surveillance using incremental SVM on texture and color features ». In : *Proceedings of the Second International Conference on Adaptive and Intelligent Systems, ICAIS 2011*. Springer Berlin Heidelberg, p. 26-39. DOI : 10.1016/j.neucom.2012.08.071.

[IC24] **Fleury, A.**, Noury, N., Vacher, M., (2010). « Introducing knowledge in the process of supervised classification of activities of daily living in health smart homes ». In : *e-Health Networking Applications and Services (Healthcom), 2010 12th IEEE International Conference on*. IEEE, p. 322-329. DOI : 10.1109/HEALTH.2010.5556549.

[IC25] **Fleury, A.**, Vacher, M., Portet, F., Chahuara, P., Noury, N., (2010). « A multimodal corpus recorded in a health smart home ». In : *LREC 2010, The International Conference on Language Resources and Evaluation*, p. 99-105.

[IC26] Vacher, M., Portet, F., **Fleury, A.**, Noury, N., (2010). « Challenges in the processing of audio channels for ambient assisted living ». In : *e-Health Networking Applications and Services (Healthcom), 2010 12th IEEE International Conference on*. IEEE, p. 330-337. DOI : 10.1109/HEALTH.2010.5556546.

[IC27] **Fleury, A.**, Noury, N., Vacher, M., (2009a). « A wavelet-based pattern recognition algorithm to classify postural transitions in humans ». In : *Signal Processing Conference, 2009 17th European*. IEEE, p. 2047-2051.

[IC28] — (2009d). « Supervised classification of activities of daily living in health smart homes using SVM ». In : *Engineering in Medicine and Biology Society, 2009. EMBS 2009. 31th Annual International Conference of the IEEE*, p. 6099-6102. DOI : 10.1109/IEMBS.2009.5334931.

[IC29] Portet, F., **Fleury, A.**, Vacher, M., Noury, N., (2009). « Determining useful sensors for automatic recognition of activities of daily living in health smart home ». In : *Intelligent Data International Workshop on Analysis in Medicine and Pharmacology (IDAMAP 2009)*, p. 63-64.

[IC30] Vacher, M., Guirand, N., Serignat, J.-F., **Fleury, A.**, Noury, N., (2009). « Speech recognition in a smart home : some experiments for telemonitoring ». In : *Speech Technology and Human-Computer Dialogue, 2009. SpeD'09. Proceedings of the 5-th Conference on*. IEEE, p. 1-10. DOI : 10.1109/SPED.2009.5156185.

[IC31] **Fleury, A.**, Noury, N., Vacher, M., Glasson, H., Seri, J.-F., (2008). « Sound and speech detection and classification in a health smart home ». In : *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*. IEEE, p. 4644-4647. DOI : 10.1109/IEMBS.2008.4650248.

[IC32] **Fleury, A.**, Vacher, M., Glasson, H., Serignat, J.-F., Noury, N., (2008). « Data fusion in health smart home : preliminary individual evaluation of two families of sensors ». In : *ISG'08*, p. 135.

[IC33] Noury, N., Hadidi, T., Laila, M., **Fleury, A.**, Villemazet, C., Rialle, V., Franco, A., (2008). « Level of activity, night and day alternation, and well being measured in a smart hospital suite ». In : *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*. T. 8, p. 20-24. DOI : 10.1109/IEMBS.2008.4649917.

[IC34] Vacher, M., **Fleury, A.**, Serignat, J.-F., Noury, N., Glasson, H., (2008). « Preliminary evaluation of speech/sound recognition for telemedicine application in a real environment ». In : *Interspeech'08*, p. 496-499.

[IC35] Vuillerme, N., Pinsault, N., Chenu, O., **Fleury, A.**, Payan, Y., Demongeot, J., (2008b). « The effects of a plantar pressure-based, tongue-placed tactile biofeedback system on the regulation of the centre of foot pressure displacements during upright quiet standing : a fractional brownian motion analysis ». In : *Complex, Intelligent and Software Intensive Systems, 2008. CISIS 2008. International Conference on*. IEEE, p. 162-168. DOI : 10.1109/CISIS.2008.32.

[IC36] **Fleury, A.**, Noury, N., Vuillerme, N., (2007a). « A fast algorithm to track changes of direction of a person using magnetometers ». In : *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*. IEEE, p. 2311-2314. DOI : 10.1109/IEMBS.2007.4352788.

[IC37] Noury, N., **Fleury, A.**, Rumeau, P., Bourke, A., Laighin, G., Rialle, V., 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*. IEEE, p. 1663-1666. DOI : 10.1109/IEMBS.2007.4352627.

[IC38] Vuillerme, N., Chenu, O., Pinsault, N., Moreau-Gaudry, A., **Fleury, A.**, Demongeot, J., Payan, Y., (2007). « Pressure sensor-based tongue-placed electrotactile biofeedback for balance improvement-biomedical application to prevent pressure sores formation and falls ». In : *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*. IEEE, p. 6113-6116. DOI : 10.1109/IEMBS.2007.4353744.

[IC39] Noury, N., Villemazet, C., **Fleury, A.**, Barralon, P., Rumeau, P., Vuillerme, N., Baghai, R., (2006). « Ambient multi-perceptive system with electronic mails for a residential health monitoring system ». In : *Engineering in Medicine and Biology Society, 2006. EMBS 2006. 28th Annual International Conference of the IEEE*, p. 3612-3615. DOI : 10.1109/IEMBS.2006.260242.

[IC40] Vuillerme, N., Chenu, O., **Fleury, A.**, Demongeot, J., Payan, Y., (2006). « Optimizing the use of an artificial tongue-placed tactile biofeedback for improving ankle joint position sense in humans ». In : *Engineering in Medicine and Biology Society, 2006. EMBS 2006. 28th Annual International Conference of the IEEE*. IEEE, p. 6029-6032. DOI : 10.1109/IEMBS.2006.260826.

Conférences Nationales avec Actes et Comité de Lecture

[NC1] Vermeulen, M., **Fleury, A.**, Fronton, K., Laval, J., (2015). « Les alpes : approches agiles pour l'enseignement supérieur ». In : *Colloque Questions de Pédagogie pour l'Enseignement Supérieur (QPES 2015)*.

[NC2] Franco, C., Guméry, P.-Y., **Fleury, A.**, Vuillerme, N., (2014). « Nouvelle approche de l'entropie spectrale à partir d'une transformée par synchrosqueezing ». In : *46ème journée de la Société Francaise de Statistiques*.

[NC3] Mourcou, Q., **Fleury, A.**, Franco, C., Vuillerme, N., (juin 2014). « Wegoto : une application smartphone pour évaluer l'accessibilité de la voirie et des espaces publics aux personnes à mobilité réduite – etude de preuve de concept ». In : *Conférence Handicap 2014*.

[NC4] Lu, Y., **Fleury, A.**, Boonaert, J., Lecoeuche, S., (2013). « Identification/réidentification de personnes sur des vidéos grâce à des algorithmes de classification incrémentaux. » In : *GRETSI 2013*, p. ID439.

[NC5] Chahuara, P., **Fleury, A.**, Vacher, M., Portet, F., (2012). « Méthodes SVM et MLN pour la reconnaissance automatique d'activités humaines dans les habitats perceptifs : tests et perspectives ». In : *RFIA 2012 (Reconnaissance des Formes et Intelligence Artificielle)*, p. 978-2.

[NC6] Franco, C., **Fleury, A.**, Villemazet, C., Demongeot, J., Gumery, P.-Y., Vuillerme, N., (oct. 2010). « Vigi-ma : vigilance maladie d'alzheimer ». In : *Conférence CIFGG 2010 - 9ème Congrès International Francophone de Gériatrie et de Gérontologie*.

[NC7] Duc, C., Farron, A., Coley, B., **Fleury, A.**, Aminian, K., Pichonnaz, C., Bassin, J.-P., Jolles, B., (2009). « A new way of quantifying dominant upper-limb mobility in healthy and painful shoulders ». In : *Annual Meeting of the Swiss Society of Orthopedy and Traumatology (SSO)*.

[NC8] **Fleury, A.**, Noury, N., Vacher, M., (2009b). « Application des SVM à la classification des activités de la vie quotidienne d'une personne à partir des capteurs d'un habitat intelligent pour la santé ». In : *SFC 2009*, pp-33.

[NC9] — (2009e). « Traitement des signaux cinématiques pour la détection et la classification des transferts posturaux : le système actim6d ». In : *GRETSI 2009*, p. ID457.

[NC10] Vacher, M., **Fleury, A.**, Portet, F., Serignat, J.-F., Noury, N., (2009). « Reconnaissance des sons et de la parole dans un habitat intelligent pour la santé : expérimentations en situation non contrôlée ». In : *GRETSI 2009*, p. ID456.

[NC11] Chenu, O., Vuillerme, N., Moreau-Gaudry, A., **Fleury, A.**, Demongeot, J., Payan, Y., (2007). « Suppléance perceptive par électro-stimulation linguale embarquée : perspectives pour la prévention des escarres chez le blessé médullaire ». In : *14ème Congrès de l'Association Posture-Équilibre*.

[NC12] **Fleury, A.**, Noury, N., Vuillerme, N., (2007b). « Traitement de signaux issus de magnétomètres embarqués-application à la détection des changements de direction d'une personne ». In : *GRETSI 2007*, pp-161.

Participations à des journées techniques

[JT1] **Fleury, A.**, Vuillerme, N., (2016). « Le smartphone comme outil de rééducation et d'entraînement au domicile ». In : *Salon Cité-Santé, Nancy, 1-2 avril*.

[JT2] **Fleury, A.**, Karami, A. B., Boonaert, J., Lecoeuche, S., (mar. 2016a). « Le smartphone comme outil d'entrainement et de rééducation au domicile ». In : *Atelier Santé, Institut Mines Télécom*.

[JT3] — (mar. 2016b). « Les habitats intelligents au service de la santé et du suivi à domicile ». In : *Atelier Santé, Institut Mines Télécom*.

[JT4] **Fleury, A.**, Boonaert, J., Rajaoarisoa, L., (sept. 2015). « Smart homes for health and well-being ». In : *Ambient Assisted Living (AAL) Forum 2015, Gent*.

[JT5] Bouis, J., Franco, C., Mourcou, Q., **Fleury, A.**, Diot, B., Vuillerme, N., (2014). « Peut-on améliorer son équilibre unipédique grâce à l'utilisation du biofeedback auditif de l'application smartphone ibalance ? » In : *Journées de la Recherche Médicale, Grenoble*.

[JT6] Delevoye-Turrell, Y., Carlier, M., Carlton, O., **Fleury, A.**, Lecoeuche, S., (mar. 2013). « Using artificial neural networks to assess the validity of a neuropsychological test battery ». In : *Symposium on Cognitive disorders and remediation in schizophrenia and other mental diseases*.

[JT7] Franco, C., Sugny, S., Vaillant, J., **Fleury, A.**, Guméry, P.-Y., Vuillerme, N., (2012). « Double tâche et contrôle de la posture du sujet âgé : apport de l'entropie par échelles ». In : *Les journées de la Recherche Médicale de Grenoble, Grenoble*, p. 44.

[JT8] Franco, C., **Fleury, A.**, Gumery, P.-Y., Diot, B., Demongeot, J., Vuillerme, N., (juil. 2012). « iBalance : un smartphone 'intelligent' pour améliorer le contrôle de l'équilibre – etude de preuve-de-concept, » in : *Université d'été e-Santé, GDR Stic Santé Thème G (e-santé), Castres, Juillet*.

[JT9] **Fleury, A.**, Boonaert, J., Lecoeuche, S., (mar. 2012a). « Analyse d'activités de la vie quotidienne avec des capteurs inertiels embarqués ». In : *Colloque Stic-Santé, Institut Mines-Télécom, Paris, Mars*.

[JT10] — (mar. 2012b). « Les habitats intelligents au service de la santé et du suivi à domicile ». In : *Colloque Stic-Santé, Institut Mines Télécom, Paris, Mars*.

Séminaires Invités

[IS1] **Fleury, A.** (2016a). « Acquisition de données dans les environnements intelligents : de la reconnaissance d'activités à la création de profils utilisateurs pour l'automatisation. » In : *Journées du laboratoire LISIC de l'Université des Sciences et Techniques Houari Boumédienne, Alger, Algérie*.

[IS2] — (2016b). « Entrainement et rééducation de la proprioception avec un outil centré sur le smartphone ». In : *Journée Handicap, Chaire e-Biomed, Université de Technologies de Compiègne*.

[IS3] — (2012). « Non-incremental and incremental SVM and their application ». In : *Séminaire du LIASD, Université Paris 8*.

Article de vulgarisation

[V1] **Fleury, A.**, Rajaoarisoa, L., (mai 2016). « Quels apports possibles, pour l'utilisateur, des technologies utilisées dans l'habitat intelligent ». In : *Minerai* 3, p. 23-29.

Mémoires

[Th1] **Fleury, A.** (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 télémédecine ». Thèse de doct. Université Joseph-Fourier-Grenoble I.

[Th2] — (2005). « Exploitation des Signaux de Capteurs Magnétométriques Embarqués. Application à la Détection Automatique des Activités de Vie Quotidienne d'une Personne. » Mém. de mast. Master EEATS, Université Joseph Fourier, Grenoble 1.

Activités d'édition

[Ed1] **Fleury, A.**, Roussos, G., éd. (déc. 2019). *Special Issue : Smartphone-Based Sensors for Posture, Movement Analysis and Human Activity Recognition*. Sensors, MDPI.

[Ed2] Sayed-Mouchaweh, M., Lughofer, E., **Fleury, A.,** éd. (mar. 2017). *Special Issue : Adaptive Intelligent Systems for Learning, Control and Optimization*. Evolving Systems, Springer.

[Ed3] *Proceedings of the IEEE International Conference on Evolving and Adaptive Intelligent Systems (EAIS)* (2015).

Bases de données publiques

[D1] Hammouche, M., Ambellouis, S., **Fleury, A.,** (2015). *URIA Kinect Dataset*. Online : http ://ia.ur.mines-douai.fr/en/datasets/.

[D2] Lu, Y., Boonaert, J., Lecoeuche, S., **Fleury, A.,** (2014). *URIA Multicamera Dataset*. Online : http ://ia.ur.mines-douai.fr/en/datasets/.

[D3] **Fleury, A.**, Noury, N., Vacher, M., (2009c). *Grenoble Health Smart Home Dataset*. Online : https ://sites.google.com/site/thonyfleury/health-smart-home-his-datasets.

Résumé

Les données sont de nos jours présentes partout, qu'elles soient volontairement ou involontairement cédées par un utilisateur. L'augmentation des puissances de calcul permet de traiter ces données plus en profondeur et plus efficacement, ce qui a poussé différentes applications à utiliser fortement le traitement de données pour offrir un service plus important aux usagers. Parmi les utilisations de ces données, les technologies pour la santé et le bien-être sont une thématique qui se développe de plus en plus ces dernières années et qui est un champ de recherche actif. Cette habilitation décrit dix années de recherche sur l'analyse de comportement et d'activité d'une personne, à partir de différents capteurs, afin de proposer des applications allant de l'aide à la rééducation ou l'évaluation de la santé d'une personne, à l'adaptation des automatismes dans un habitat en fonction de la volonté de l'habitant, en passant par la reconnaissance d'activités par smartphone ou par caméra vidéo. Organisés en trois parties, ces travaux se concentrent d'abord sur l'analyse des automatismes de la personne et la mesure de certains de ses mouvements. Avec ceci, des solutions permettant l'évaluation de la personne ou permettant de faire un retour sensoriel à celle-ci (afin de lui permettre de se corriger) se dessinent. La seconde partie se concentre sur la reconnaissance de la personne et de ses actions, avec des applications notamment sur la reconnaissance précoce d'actions. Dans cette partie sont également investiguées des techniques de classification incrémentales et en ligne afin d'améliorer les taux de reconnaissance et les contraintes d'apprentissages. Enfin, la dernière partie présente la reconnaissance d'activités dans un habitat intelligent, avec comme application finale la possibilité de créer des profils utilisateurs dans ces habitats.

Abstract

Data are everywhere. Voluntarily or not, they are given by the user to different companies or institutions. Improvement in computing power allows now to use more efficiently these huge quantities of data. It stimulated the creation of different applications and one of them is related to healthcare technologies and well-being. This habilitation describes the work that have been done these last years on human behavior analysis and activity recognition, from different sensors, to propose applications such as rehabilitation, health and well-being evaluation, profiling of person to adapt the automation of the environment, and also activity estimation from smartphone and video sensors. Organized in three different parts, this habilitation focuses first on the analysis of automation of the human body and the measurement of some movements. With this, solutions allowing the evaluation of the person and offering real-time feedbacks (to correct a posture or a movement) are drawn. The second part focuses on the recognition of the person and of his actions, with application for instance to early recognition of actions. In this part are investigated incremental and on-line classification techniques to improve the recognition rates but also to reduce the constraints in learning stage. Finally, the last part presents the recognition of activities in smart homes, with a final application to create user profiles in automated habitations.