UNIVERSITE BRETAGNE LOIRE



THÈSE / ENS RENNES
Université Bretagne Loire
pour obtenir le titre de
DOCTEUR DE L'ÉCOLE NORMALE SUPÉRIEURE DE RENNES
Mention : Traitement du Signal et Télécommunications
École doctorale MathSTIC

présentée par
Hala Abdul Rahman
Préparée à l'unité mixte de recherche INSERM U-1099
Laboratoire Traitement du Signal et de l'Image (LTSI)
et au Laboratoire Mouvement, Sport, Santé (M2S)

Multi-Sensor Based Activity Recognition: Development and Validation in Real-life Context

Thèse soutenue le 20 décembre 2017
devant le jury composé de :

Daniel MASSICOTTE
Professeur, Université du Québec à Trois-Rivières / *rapporteur*
Alan DONNELLY
Maître de conférences, Université de Limerick / *rapporteur*

Anne HUMEAU-HEURTIER
Professeur, Université d'Angers / *examinatrice*
Mohamad KHALIL
Professeur, Université Libanaise / *examinateur*

Guy CARRAULT
Professeur, Université de Rennes 1 / *directeur de thèse*
Jacques Prioux
Professeur, ENS Rennes / *co-directeur de thèse*

ACKNOWLEDGMENTS

First and foremost, I would like to extend a very special gratitude to my advisors, Prof. Guy Carrault and Prof. Jacques Prioux for their constant support and motivation. They have been wonderful mentors over my journey as a graduate student and it has been a great privilege and honor working with them. My heartfelt thanks to Dr. Di Ge and Dr. Alexis LeFaucheur for their dedicated help and advice at various stages of my research. Their continuous support and mission for providing high-quality work are appreciable. I would also like to thank the team of Sherpam project, whose members have made significant contributions to the development of this project. I gratefully acknowledge the Labex CominLabs role as a funding source that made my PhD possible. I would also take this opportunity to express my sincere gratitude and appreciation to the volunteers that participated in this study for offering their time and efforts, and without whom this work would not have been possible. Finally, I would like to thank the thesis committee for their time and insightful feedback.

TABLE OF CONTENTS

LIST OF FIGURES ..... **V**
LIST OF TABLES ..... **VII**
LIST OF ACRONYMS ..... **IX**
RÉSUMÉ ..... **XI**
CHAPTER 1 GENERAL INTRODUCTION **1**

1.1 PROJECT SHERPAM 2
1.2 DISSERTATION CONTRIBUTIONS 3
1.3 DISSERTATION STRUCTURE 4

CHAPTER 2 BACKGROUND **7**

2.1 IMPORTANCE OF MONITORING PHYSICAL ACTIVITY. 8
2.2 DIFFERENT APPROACHES FOR ACTIVITY RECOGNITION. 10
2.2.1 Visual monitoring sensors *10*
2.2.2 Wearable sensors *10*
2.2.2.1 Microelectromechanical systems (MEMS) and sensors development 11
2.2.2.2 Motion sensors 11
2.2.2.3 Vital Sensors 13
2.2.2.4 Location sensors 15
2.2.2.5 Environmental sensors 15
2.3 PROCESS OF ACTIVITY RECOGNITION 15
2.3.1 Data acquisition and pre-processing *16*
2.3.2 Data segmentation and windowing *17*
2.3.3 Features Extraction *18*
2.3.4 Features Selection *19*
2.3.5 Classification *20*
2.3.5.1 Building the classifier 20
2.3.5.2 Training the classifier 23
2.3.5.3 Testing the classifier 24
2.4 CHALLENGES IN THE RESEARCH FIELD OF ACTIVITY RECOGNITION 24
2.4.1 Decision of the activity classes *24*
2.4.2 Selection of the sensors *25*
2.4.3 User flexibility *26*
2.5 ESTIMATION OF ENERGY EXPENDITURE 26
2.5.1 Importance of monitoring EE *26*
2.5.2 Methods and tools for measuring EE *29*
2.5.2.1 Direct Measurements 29
2.5.2.2 Indirect Measurements 30

2.5.2.3 Mathematical modeling and prediction equations 30
2.5.3 Overview of existing EE studies 33
2.6 CONCLUSION 34

CHAPTER 3 EXPERIMENTAL PROTOCOLS 37

3.1 OBJECTIVES OF THE TRIPLE-SESSIONS PRAGMATIC PROTOCOL 38
3.2 LITERATURE REVIEW ON EXISTING DATA COLLECTION PROTOCOLS 39
3.2.1 Standardized controlled protocol 39
3.2.2 Semi-standardized protocol 39
3.2.3 Unstandardized free-living protocol 40
3.3 BASIC DATA COLLECTION PROTOCOL 44
3.3.1 Materials 44
3.3.2 Experimental procedure 44
3.3.3 Participants 44
3.4 PRAGMATIC DATA COLLECTION PROTOCOL 45
3.4.1 Materials 45
3.4.1.1 Shimmer3 IMU and ECG modules 48
3.4.1.2 Zephyr BioHarness3 Chest Strap 49
3.4.1.3 ActiGraph GT3x+ 50
3.4.1.4 Qstarz BT-Q1000eX 50
3.4.1.5 SenseWear BodyMedia Armband 51
3.4.1.6 ActivPAL 51
3.4.1.7 Autographer Camera 52
3.4.1.8 Cosmed K4b2 52
3.4.2 Experimental procedure 53
3.4.2.1 Standardized session (S1) 53
3.4.2.2 Semi-standardized session (S2) 54
3.4.2.3 Free-living session (S3) 55
3.4.3 Participants 56
3.5 CONCLUSION 56

CHAPTER 4 METHODOLOGY AND PRELIMINARLY RESULTS 59

4.1 INTEREST OF THE SPECTRAL DENSITY DISTANCE MEASURES 60
4.1.1 Limitation of classical approaches 60
4.1.2 New modeling approach 61
4.2 FEATURES EXTRACTION METHODS 62
4.2.1 Set of spectral distances features 62
4.2.2 Distances features illustration 64
4.2.3 Set of state-of-art time- and frequency- domain features 65
4.3 HEART RATE ANALYSIS 66
4.4 EVALUATION 66
4.4.1 Comparison between Time-Frequency and Spectral Distance Methods 67

4.4.2 Comparison between sensors' placements *67*
4.4.3 Influence of HR data *69*
4.5 CONCLUSION 70

CHAPTER 5 FREE-LIVING ACTIVITY CLASSIFICATION 73

5.1 RACHA PLATFORM: REAL-TIME APPLICATION FOR THE CLASSIFICATION OF HUMAN ACTIVITIES 74
5.1.1 Objectives of RACHA Platform *74*
5.1.2 Data Visualization Interface *75*
5.1.3 Activity Classification Interface *76*
5.2 DETAILS ON THE PRAGMATIC DATASET 79
5.2.1 Labelling the activities groups *79*
5.2.2 S3-photo analysis and annotation protocol *79*
5.2.3 Activities composition of the sessions *80*
5.2.4 Extracted Features *83*
5.3 RESULTS 84
5.3.1 Comparison between sessions *85*
5.3.2 Comparison between window lengths *86*
5.3.3 Comparison between sensors positions *87*
5.3.4 Comparison between sensors' variables *91*
5.3.5 Approach for Real-life Application *92*
5.4 DISCUSSION 93
5.5 **VE** EXTRACTION FROM ZEPHYR-BREATHING WAVEFORM 98
5.6 CONCLUSION 101

CHAPTER 6 GENERAL CONCLUSION AND FUTURE PROSPECTS 103

6.1 CONTRIBUTIONS AND FINDINGS 104
6.2 FUTURE PROSPECTS 105
6.2.1 Towards robust PAR system in free-living context *105*
*6.2.2 Towards better estimation of EE: **VE** analysis perspective* *109*

APPENDIX-A DETAILS ON PRIOR STUDIES 111
REFERENCES 121
SCIENTIFIC VALORIZATION 127

LIST OF FIGURES

Figure 1.1 - The chain of collecting, transmitting and processing data in SHERPAM project 2
Figure 2.1 - 3D representation of the number of studies based on application and sensor types 12
Figure 2.2 - The principles, fundamental elements and functioning of an accelerometer 12
Figure 2.3 - Typical design of an activity recognition chain 17
Figure 2.4 - Performance of different classifiers as a function of varying window sizes. 18
Figure 2.5 - Range of features extracted from time-series data 19
Figure 2.6 - Illustration of the k-nearest neighbor principle 22
Figure 2.7 - Hierarchical example of a decision tree 22
Figure 2.8 - Schematic representation of the iterations of a k-fold Cross Validation. 23
Figure 2.9 - Hierarchical Framework showing the assessment of PA and the prediction of EE 28
Figure 2.10 - The distribution of the EE methods based on their ease of assessment vs precision 29
Figure 2.11 - Partition of the energy expenditure components 31
Figure 2.12 - An example of the relationships between HR and EE 32
Figure 2.13 - Illustration of the electromagnetic coils placement on the subject 33
Figure 3.1 - The Shimmer nodes and the ECG electrodes positions in the basic protocol 45
Figure 3.2 - The performed activities in the basic protocol 45
Figure 3.3 - An illustration of the equipment and their positions on the body 46
Figure 3.4 - A participant wearing the sensors and gas exchange system. 47
Figure 3.5 - Description of the triple-session protocol for data collection 53
Figure 3.6 - Examples of the performed activities during Session S1 54
Figure 3.7 - Example of photos captured by the Autographer Camera 56
Figure 4.1 - Time-domain signals from activities windows and the corresponding spectrums. 64
Figure 4.2 - Color scaled 2D image resulting from the distances calculation 65
Figure 4.3 - Overall classification accuracies per unit position for the three classifiers. 69
Figure 4.4 - Overall classification accuracies before and after including the HR information. 70
Figure 5.1 - Data Visualization interface 75
Figure 5.2 - Example of a HR feature extracted from one subject performing session S1 77
Figure 5.3 - Activity Classification Interface showing an example of classification results 78
Figure 5.4 - A view of the Autographer camera interface 80
Figure 5.5 - Acceleration signals from 3D accelerometers of the Zephyr chest strap 82
Figure 5.6 - Overall performance with different window lengths for S1 test-set 87
Figure 5.7 - Overall performance with different window lengths for S3 test-set 87

Figure 5.8 – Illustration of the KNN classifier performances for S1 test-set when coupling two and three ActiGraph modules together 90
Figure 5.9 - Illustration of the KNN classifier performances for S3 test-set when coupling two and three ActiGraph modules together. 90
Figure 5.10 - Average accuracy of acceleration (Accel), (Accel + HR) and (Accel + BR) taken from Zephyr module 92
Figure 5.11 - Overall classification accuracy when S1, S3 and S1 with S3 sets are used for training . 93
Figure 5.12 - Ventilation (VE) extraction process from Zephyr Chest strap. 101
Figure 6.1 - Shimmer-pressure and Shimmer-X-axis acceleration signals variations in going up/down stairs and walking activities 107
Figure 6.2 - Example of misclassification samples among activities 108

LIST OF TABLES

Table 2.1 - Summary of the state-of-the-art in EE estimation studies 35
Table 3.1 - Summary of the strengths and limitations of the existing free-living protocols. 42
Table 3.2 - Summary of the strengths and limitations of the existing free-living protocols-continued 43
Table 3.3 - Physical Characteristics of the participants in the basic protocol. 45
Table 3.4 - The physical characteristics of the considered systems in the pragmatic protocol 47
Table 3.5 - The technical specifications of the considered systems in this study 48
Table 3.6 - Physical characteristics of the participants in the pragmatic protocol 56
Table 4.1 - The order of the information gain of sorted features 61
Table 4.2 - The overall mean accuracy for the SF and SpD methods 67
Table 4.3 - Confusion matrices for NB, KNN and C4.5 classifiers for SF and SpD method. 68
Table 4.4 - Overall Performances with un-normalized HR and normalized HR information. 70
Table 5.1 - The activity labels and their corresponding numbers as considered in RACHA platform. 76
Table 5.2 - The 7 categories created from the 17 activities..... 79
Table 5.3 - The distribution of minutes recorded of categories in each session..... 81
Table 5.4 - Overall F-score measures of the KNN classifier using SF and SpD* methods 86
Table 5.5 - Overall F-score measures of the DT classifier using SF and SpD* methods 86
Table 5.6 - F-score values per category of activities for standardized S1 test-set..... 88
Table 5.7 - F-score values per category of activities for S3 test-set..... 88
Table 5.8 - Aggregated confusion matrix, sensitivity and precision for S3 test-set..... 89
Table 5.9 - Overall F-score values of the accelerometer, gyroscope and magnetometer from Shimmer module for S1 and S3 test-sets 91
Table 5.10 - Overall F-score values of the accelerometer, HR and BR from Zephyr chest strap module for S1 and S3 test-sets 91
Table 5.11 - Description of the breathing protocol 99

LIST OF ACRONYMS

AC	Activity Counts
Acc	Accelerometer
ADL	Activities of Daily Living
ANN	Artificial Neural Network
AR	Autoregressive
Bpm	Breaths/Beats per minute
BR	Breathing Rate
CSTS	Cross Sectional Time Series
DLW	Doubly Labeled Water
DT C4.5	Decision Tree
HR	Heart Rate
EE	Energy Expenditure
FN	False Negatives
FP	False Positives
GPS	Global Positioning System
Gyr	Gyroscope
GUI	Graphical User Interface
HMM	Hidden Markov Model
HF	Heart Failure
IAMD	International Activity Monitor Database
ICAD	International Children's Accelerometry Database
IQ	Intelligent Quotient
KNN	K-Nearest Neighbors
Mag	Magnetometer
MARS	Multivariate Adaptive Regression Splines
MCS	Multiple Classifier System
ME	Manufacturer Equation
MEMS	Microelectromechanical Systems
MET	Metabolic Equivalent
NB	Naïve Bayes
NIDDM	Non-Insulin-Dependent Diabetes Mellitus
PA	Physical Activity

PAD	Peripheral Artery Disease
PAEE	Physical Activity Energy Expenditure
PAR	Physical Activity Recognition
PD	Parkinson Disease
PDA	Personal Digital Assistant
RC	Room Calorimeter
RD	Raw Data
RE	Regression Equation
RIP	Respiratory Inductive Plethysmograph
RMR	Resting Metabolic Rate
SD	Standard Deviation
SF	Statistical Features
SpD	Spectral Distances
SpO_2	Blood oxygen saturation levels
TEE	Total Daily Energy Expenditure
TEF	Thermic Effect of Food
TP	True Positives
TN	True Negatives
$\dot{V}_E$	Ventilation
$\dot{V}CO_2$	Carbon dioxide production
$\dot{V}O_2$	Oxygen consumption
WEKA	Waikato Environment for Knowledge Analysis
WHO	World Health Organization

RÉSUMÉ

Le lien entre activité physique (AP) et santé a été constamment documenté. Les bénéfices de la pratique de l'AP sur la qualité de vie ne cessent d'être démontrés. C'est dans ce contexte que l'évaluation de l'AP connaît une place de plus en plus centrale dans le secteur médical, notamment dans la prévention des maladies et le diagnostic des principales pathologiques chroniques. L'avancement des travaux de recherche sur la reconnaissance automatique et objective de l'AP est principalement dû à la révolution technologique des dispositifs de mesure et d'enregistrement permettant de quantifier les mouvements. De la même façon la quantification de la dépense énergétique (DE) connaît un regain d'intérêt grandissant. Même si les techniques directes (eau doublement marquée, chambre calorimétrique, etc.) de mesure de DE correspondante aux AP restent les techniques de référence, l'utilisation de capteurs pour estimer indirectement DE montre une précision suffisante. L'axe d'étude majeur suivi par les chercheurs est d'établir une classification des AP, pour ensuite estimer la DE associée à chacune des AP. Il a été montré que la mesure de DE est plus fiable en la précédant par une identification des types d'AP plutôt qu'en l'estimant directement. Les capteurs inertiels, particulièrement les accéléromètres, les capteurs physiologiques (ex. cardiofréquencemètres) et les récepteurs GPS sont les outils portables les plus investigués dans ce domaine. De nombreuses études attestent de leur fiabilité et de leur précision s'agissant de la reconnaissance et de la quantification d'une AP. A partir des données de sortie de ces capteurs, un modèle de reconnaissance des AP est développé à l'aide d'outils numériques. Sa performance est ensuite mesurée au travers de sa capacité à classifier correctement les activités en question.

Généralement, la collecte des données de capteurs pour la plupart des études est réalisée d'une manière standardisée en laboratoire, sur une courte durée, sur des types d'activités spécifiques et à des intensités précisément calibrées. En réalisant les activités en question, il est demandé aux sujets équipés de capteurs de suivre les instructions de l'examinateur. Ces études souffrent d'un manque de validation externe des outils et méthodes développées en particulier dans un environnement naturel de vie libre ou au moins semi-libre. Or la finalité de ces systèmes est d'être intégrée dans des scénarios de « vraie vie ». Cela exige que l'apprentissage des modèles de reconnaissance soit mené dans le même contexte que celui de l'application finale. C'est un problème critique puisque le comportement naturel des sujets diffère certainement entre un environnement contrôlé et un environnement naturel (i.e., celui de la vie courante). La nécessité de collecter les données des capteurs dans un environnement naturel de vie libre en dehors du laboratoire est donc cruciale.

Pour toutes ces raisons, l'objectif général de notre étude a été de concevoir un système de reconnaissance capable d'identifier une variété d'activités à partir des données enregistrées dans un scénario de vie libre.

Dans cette voie d'étude, les différentes questions auxquelles nous avons souhaité répondre au cours de ce travail sont : i) Comment la nature des données d'apprentissage influe sur la réponse des systèmes de reconnaissance ? ii) La classification est-elle plus efficace en combinant plusieurs types de capteurs ? iii) Quelle est la contribution de chaque capteur et de chaque variable de mesure dans la reconnaissance des activités ?

Ce manuscrit de thèse s'inscrit également dans une deuxième voie d'étude : le développement d'une méthode d'extraction de paramètres adaptés à la classification des activités. Cette méthode, présentée au chapitre 4, est basée sur la mesure des distances spectrales entre les signaux et les modèles Autorégressifs (AR) formés pour chaque activité. Bien que dans la littérature, il existe de nombreux algorithmes capables d'identifier les AP avec une bonne précision, peu d'entre eux prennent en compte l'efficacité et la pertinence des paramètres extraits. Contrairement aux études qui extraient un grand ensemble de paramètres et qui ensuite, après sélection, réduisent sa dimensionnalité, nous proposons dans cette étude d'interpréter un ensemble réduit de paramètres qui nous paraissent plus adaptés au problème de classification des AP. En fait, l'extraction et l'incorporation d'un grand nombre de paramètres dans le processus de classification peuvent entrainer (i) la redondance ou la non-pertinence de certains paramètres qui n'apportent pas d'information significative à la performance du système et (ii) une complexité et un temps supplémentaire demandé pour le calcul des paramètres.

Dans le cadre de cette thèse, deux bases de données issues de deux différents protocoles ont été mises en place (P1) et (P2) (chapitre 3). La première (P1) comporte 5 activités sédentaires et de locomotion (repos assis/debout, marche, course, vélo et voiture) exécutées par 8 sujets, chacune sur une durée d'au moins 3 minutes. Les sujets étaient équipés de 3 unités de mesure inertielles (IMU) Shimmer3 placées sur le poignet, la hanche et la cheville et d'une unité Shimmer-ECG placée sur la poitrine. Cette base de données a servi dans un premier temps pour le développement du modèle de classification basé sur les distances spectrales. Des méthodes conventionnelles disponibles dans la littérature spécifique et basées sur l'extraction des paramètres temporels et fréquentiels ont servi de méthodes de référence auxquelles était comparée notre méthode. Le modèle de distances spectrales montre une précision comprise entre 87.8% et 99.0% de bonne classification lorsque les signaux 3D-accélérométriques des 4 unités Shimmer3 sont considérées. Dans un deuxième temps, nous avons étudié la contribution de la fréquence cardiaque (FC) issue du signal Shimmer-ECG, à la reconnaissance des activités. Nos résultats montrent une amélioration de la performance lorsque qu'on couple FC aux signaux des différentes positions. Ces résultats nous ont incités à évaluer les performances de notre système de reconnaissance sur une plus grande variété d'activités, collectées dans un contexte plus ambitieux : celui de la vie courante.

Le protocole expérimental (P2) est scindé en 3 phases de natures différentes : une phase standardisée (S1), une phase semi-standardisée (S2) et une phase de vraie-vie (S3). L'intérêt de cette base de données est d'apporter des situations d'AP et de comportements sédentaires variés, en nombre suffisant, pour permettre la validation et la construction d'algorithmes efficaces sur des applications de « vraie vie ». Vingt sujets ont participé à cette étude. Ils étaient équipés de différents types de capteurs : une ceinture

thoracique de Zephyr, des unités accélérométriques Shimmer et ActiGraph, et des récepteurs GPS de Qstarz. Des signaux physiques (accélération, vitesse) et physiologiques (FC, fréquence respiratoire) ont été récupérés à la sortie de ces capteurs lors de la réalisation de 17 AP et d'activités sédentaires regroupées en 7 catégories : repos, activités ménagères, activité sportive, marche, course, vélo, escaliers. Les phases S1 et S2 ont été effectuées sous la supervision d'un observateur. Lors de la phase S3, les sujets étaient dans une situation de vie normale. Ils réalisaient leurs activités quotidiennes sur une journée (7 à 8 heures) tout en portant les différents capteurs. Un appareil photo à déclenchement automatique était également porté autour du cou. Il permettait d'annoter les activités lors de la journée. De notre point de vue, cette base de données est vraisemblablement unique et constitue une des originalités du travail. Cette base nous a permis (i) de valider et comparer la précision de ces différents capteurs de commerce et leurs variables de mesure dans la détection des AP et des activités sédentaires, (ii) de tester la robustesse des modèles développés ultérieurement (modèle basé sur les distances spectrales et modèle de référence) mais à partir des activités de la phase 3 réalisées en conditions de vie courante et (iii) d'étudier la distribution optimale des capteurs sur le corps afin d'identifier les 7 catégories d'activités avec une précision suffisante.

En accord avec notre hypothèse, nos résultats montrent une importante dégradation de la performance des deux systèmes de reconnaissance entre l'identification des activités contrôlées de la phase 1 et l'identification des activités de « vraie vie » de la phase 3. Cette dégradation est constatée quel que soit le dispositif en test et quels que soient les classifieurs utilisés (le K-plus proche voisin (KNN) et l'arbre de décision (DT)). Malgré le fait que le taux de bonne classification du système Zephyr ait atteint en moyenne 95.44% sur les données contrôlées, il diminue dramatiquement à 50.68% lorsque le système est testé sur les données de la phase 3. Pour remédier à cette chute de reconnaissance, nous avons proposé de reconstruire les modèles de classification en ajoutant des données de la phase 3 dans l'apprentissage. Cette solution permet d'améliorer les résultats mais le taux de classification des données en phase 3 reste modéré (62.15% en moyenne). D'autre part, la comparaison des performances résultant des différents capteurs et des paramètres suggère que l'accéléromètre du Zephyr porté sur la poitrine est plus performant que les autres capteurs lors de l'identification des 7 catégories d'activités. La comparaison des performances, en terme de position sur le corps, montre que les informations fournies par les capteurs des membres inférieurs sont plus utiles pour la discrimination des catégories ambulatoires étudiés (course, marche, vélo). En complément de ces résultats, ce manuscrit en propose d'autres en relation notamment avec l'étude des modèles avec différentes combinaisons de capteurs et différentes combinaisons de paramètres physiques et physiologiques.

Enfin, il est important de noter que l'ensemble des solutions développées ou testées dans ce travail a été embarqué dans une plateforme logicielle. Cette plateforme RACHA (Real-time Application for the Classification of Human Activities) facilite la lecture de la base de données collectées et sa représentation. Elle permet une exploitation directe et approfondie des signaux de tous les capteurs utilisés. Cette interface facilite la construction et l'évaluation des modèles de classification des trois

phases expérimentales. Elle constitue aussi une base fondamentale pour poursuivre le traitement, pour appliquer des modèles plus complexes et innovants, pour améliorer la reconnaissance et soutenir ainsi les recherches actuelles dans ce domaine.

Chapter 1
GENERAL INTRODUCTION

Monitoring of human activities has gained important interest in a broad range of real-world applications ranging from health and disease management (Warburton et al., 2006) to military and security domains (Thomas Wyss et al., 2010), biomechanics (Sabatini, 2006), ergonomics (Parkka et al., 2006), virtual reality and even sports and entertainment sectors (Ermes et al., 2008; Kunze et al., 2006; Shotton et al., 2013). Nowadays, physical activity recognition (PAR) is a key element in developing video game consoles. For instance, the generation of Wii consoles can now interactively control the game via their body gestures. PAR systems were successfully integrated in sport products as well, providing feedback on the physical fitness and the performance of athletes, such as the Nike+ running shoes and the Fitbit activity tracker. Apart from leisure domains, in combat environment, these systems are helpful in ensuring the safety of soldiers during their military missions by detecting their activities, behaviors, locations and health conditions. On top of that, health-related researchers have crucially relied on the monitoring of human activities to support and enhance traditional medical follow-up (Jovanov et al., 2005; Sung et al., 2005; Tao et al., 2007; Tessendorf et al., 2011). Fortunately, the measurement and the surveillance of blood pressure, blood lipid levels, ventilation ($\dot{V}_E$), etc. which were only possible in the clinical laboratory a few years ago are now possible outdoors with the help of sensing technologies (Dumond et al., 2017; Gastinger et al., 2014).

1.1 Project SHERPAM

This dissertation was initiated as part of the French project: SHERPAM (CominLabs, 2014) that stands for 'Sensors for Health Recording and Physical Activity Monitoring'. The objective of the project is to conceive, implement, and experimentally validate an ambulatory sensing system able to exploit in a continuous flow the biophysical data collected on mobile subjects. In the context of providing feedback on people behaviors, the system tends to monitor vital signals and physical activity (PA) with the related energy expenditure (EE) of both healthy and disabled populations during their daily life. To do so, it is important to carefully design the skeleton of the sensing system by selecting relevant physical and physiological sensors and parameters that don't limit the mobility of the subjects. The subjects should be able to go about their daily activities in a normal manner, while their health status is being monitored. In this project, the system is controlled by a smartphone that collects the sensor data and continuously transmits and uploads them to the local server (see Figure 1.1). To ensure an appropriate transmission of the collected data, both 3G and Wi-Fi networks are alternatively used to provide quasi-continuous connectivity of mobile phones. The data is then downloaded and explored via the SHERPAM web application. Two levels of data analysis are considered; the first being a preliminary analysis of data performed locally (on the smartphone) using lightweight algorithms, while the second, a level consisting of heavier processing, performs a more thorough analysis (on the server) via programming tools such as MATLAB (Mathworks, USA) software.



Figure 1.1 - The chain of collecting, transmitting and processing data in SHERPAM project.

Scenario of SHERPAM – Let's consider the example of an overweight 50-year-old man suffering from moderate hypertensive and hypercholesterolemia. His health and fitness consultants advised him to regularly exercise in order to lose weight and promote a healthier lifestyle. They provided him with a training program consisting of specific types of exercises with a recommended number of daily and

weekly hours of practice (e.g. 1 hour of walking with 3 sets of 10 mins of easy running and 1 min of resting between sets repeated every day for a week). The subject made use of the SHERPAM platform that allows him to track his daily activity profile and monitor his vital signs. The role of the SHERPAM platform is to accurately identify throughout his daily life, the executed activities, their duration and the EE related to each activity. If the subject was involved in a vigorous activity, the platform detects an increased effort and examines his heart rate (HR) signal. Once he enters a risky area regarding his cardiac activity, the platform recommends him to reduce his effort since excessive exercise may affect his health status.

1.2 Dissertation Contributions

In light of state-of-the-art, the main contributions of this dissertation are as follows:

Aim 1: Evaluation of the existing activity recognition algorithms in free-living context.

One of the critical challenges when monitoring PA is data acquisition, which has to be done under unconfined realistic conditions rather than laboratory controlled states. In this context, several studies have been carried out reporting high recognition accuracy. However, existing models are all built upon controlled activities only without being analyzed under real-life conditions (Bao et al., 2004). Few works evaluated the performance of the recognition system on data collected in natural out-of-lab conditions but they only relied on limited data sets and on specific activities (Parkka et al., 2006; Skotte et al., 2012). This is a critical issue since, in a restricted laboratory context, the normal human behaviors would be affected. People certainly act in a different manner depending on the environment in which they are performing the activities. To address the aforementioned limitations, our work highlights a novel data collection protocol that covers various measurement sessions. This protocol serves for testing the responses of classic recognition system used for discriminating various real-life activities. It relies on the use of a camera's photo records that can ensure the activities annotation from the naturalistic environment of the subjects.

Aim 2: Presenting a novel method for activity classification using the spectral density distances from acceleration data.

In order to achieve activity recognition, the raw output of the sensors is processed and transformed into pertinent information: the PA type, as well as its intensity, frequency and duration. In general, typical procedures of a PAR system (i) start by extracting basic statistical features from the signals in the time and frequency domain, (ii) reduce the feature sets dimensions by choosing the most relevant features to discriminate PA, and (iii) finally recognize the PA pattern using a classification tool. A detailed description of this scheme can be found in Chapter 2. Similar to other data mining systems, signal-processing researchers investigated the techniques to apply in PAR systems covering particularly features selection and features dimension reduction. While there exist many algorithms to accurately

classify activities based on acceleration data in recent literature (Altun et al., 2010; Cleland et al., 2013; Preece et al., 2009; Vähä-Ypyä et al., 2014); few are dedicated to the efficiency in terms of the relevance of the selected features. In fact, fetching and incorporating a large number of features into the classification process lead to the following issues: (1) the redundancy or irrelevance of some features that may not add significant information to the classification performance; (2) the complexity and the additional computational time required for calculating the features and training the model. Contrary to the state-of-the-art methods implying procedures to extract feature vectors (mean, variance, skewness, kurtosis...) for all acceleration signals and then reduce the dimensions, this dissertation suggests the study of an adapted feature set based on the Autoregressive (AR) spectral density model to measure distances of different activity types. Then, in order to assess the capability of the developed model in discriminating activities, a comparison of its accuracy with that of the state-of-the-art methods is presented. To do so, basic time-and frequency-domain features were extracted and analyzed.

Aim 3: Improving the recognition performance with Multi-Sensor and Vital-Sensor Modalities.
A variety of commercially available activity monitors with embedded accelerometers are used nowadays in the activity recognition domain. Along with the accelerometer, researchers were interested in testing other physical and physiological parameters and assessing their abilities in recognizing activities or enhancing the recognition performance (Lara et al., 2012b). For instance, it was stated that vital data, if added to the recognition system, could help to identify the intensity of the activity and the global positioning system (GPS) location data could help to identify the context of the activity in question. However, there is still no consensus on the best voted device or the best combination of sensors in this field and further studies must be conducted to draw strong conclusions on the performance of each sensor or parameter compared to other sensors and parameters. In addition, assessing the best placement of sensors on the body is critical. In prior studies, no work evokes which of the sensor locations provides the best information for recognizing a given set of activities even though it has been proved that by considering more sensors the recognition is improved (Bao et al., 2004; Cleland et al., 2013). In this dissertation, we are thus interested in studying, testing and comparing the validity, the reliability and the conformity of different types of commercially available devices that serve the field of activity recognition. To the best of our knowledge, none of the previous related works conducted a study of that kind.

1.3 Dissertation structure

A synthesis of the contents of this dissertation is presented here:
Chapter 2 - overviews the interest of the activity recognition field and covers the different types of sensors used in the literature. It also details the process of activity recognition and the associated

challenges. A part of this chapter is dedicated for reviewing the state-of-the-art related to energy expenditure; the measurement tools, the estimation approaches in the existing studies.

Chapter 3 - introduces an advanced data collection protocol that aims at evaluating the response of the recognition systems in a context of real-life activities. A description of the measurement sessions of the protocol as well as the used sensors is detailed.

Chapter 4 - presents a novel approach for activity classification developed based on spectral distance measures. The approach is then validated and compared to traditional approaches. This chapter also features the inclusion of HR data to the classification system and evaluates its efficiency in the recognition performance.

Chapter 5 – provides the main results issued from the manipulation of the real-life-dataset, typically the activity recognition performance of each sensor, each method and each measurement session in the data collection protocol presented in chapter 3. This chapter also introduces our developed interface RACHA 'Real-life Application for the Classification of Human Activities'. The design of the interface, the functionalities and the integrated methods are illustrated. A brief study introducing the use of ventilation as index of EE completed this chapter.

Chapter 6 - serves as a general conclusion regarding the work done during this thesis. It highlights the objectives achieved, summarizes the main findings and provides future prospects for the developed system such as the challenging approach of combining EE estimation and $\dot{V}_E$ modeling.

Chapter 2
BACKGROUND

The process of automatically identifying human activity has been of relevance since interest grew in shifting from subjective assessment techniques to sensor-based systems. Researchers have focused their efforts on providing accurate and reliable information on human's activities using an automated measurement methodology. This chapter serves as an introduction to some of the research done in the field of activity assessment. Its objective is to provide a clear understanding of the measurement of a PA, both qualitatively (type of activity) and quantitatively (intensity of activity and expended energy). First, we present the importance of PA monitoring typically in medical and health domains. Then we provide a review of the different types of sensors that are currently used in PAR literature, while mainly focusing on the wearable sensing approaches. Following that, we discuss the activity recognition process, highlighting the main components of any PAR system. We also address some of the key challenges facing the recognition process. The most relevant and latest published studies that were mainly interested in real-life context application of activity recognition were also surveyed and presented in Appendix A. In the last part of this chapter, we highlight different health applications in which EE estimation is essential, motivating us to explore this objective in future works. We explain the approaches for EE measurement and EE estimation, and survey some of the common EE studies.

2.1 **Importance of monitoring physical activity.**

Since decades, the field of PAR is extensively researched and forms a topic of high interest specifically in medical and health applications. The link between the common diseases and the levels and types of activities has proved the need and benefit of automatically recognizing PA.
The efficiency of monitoring PA can be grouped in three major aspects:

- Rehabilitation, health and disease management:

Automated PAR systems are used nowadays for rehabilitation, to derive treatment programs, and to assess effectiveness of treatments, providing potential support to traditional medical diagnosis. (Sung et al., 2005) via their health care platform LiveNet, tend to monitor the user activity and context together with a range of physiological signals to create a reliable health profile of patients suffering from hypothermia for example. Furthermore, to improve clinical diagnosis of Parkinson's patients, an automated Parkinson symptom detection system helps identifying the movement states of Parkinson's patients and continuously assessing the intensity of the symptoms and the complications (Dunnewold et al., 1997; Moore et al., 2008). Automated PAR systems have been also used in the detection of abnormal activities in mental pathologies such as dementia, in the course of depression treatment and in the detection of epileptic seizures (Sung et al., 2005; Yin et al., 2008). It is worth mentioning as well the application of ambulatory activity recognition in the field of assessing the walking capacity in patients with lower extremity peripheral artery disease (PAD) (Le Faucheur et al., 2008; Müllenheim et al., 2016). Therefore, recognizing the patient behavior becomes a central element of care to provide feedback about the patient's health status and thereby avert undesirable events. Furthermore, the treatment of patients with diabetes, obesity, hypertension, or cardiovascular diseases requires to follow a well-defined exercise program while assessing the efficiency of each required PA on their health status (Jia, 2009; Martin et al., 2006; Verberk et al., 2007). In stroke as well, the identification of the upper extremity movement can rely on a body-worn activity recognition system which could help determining the appropriate medical treatment (Uswatte et al., 2000).

- Management of sedentary and inactivity awareness:

In recent decades, both developed and developing communities witness an important shift towards sedentary lifestyles (Owen et al., 2010). The modern technology design of houses, work places, schools, and public spaces are nowadays limiting the population PA and encouraging prolonged laziness. At the same time, the intensity of PA is inversely and linearly associated with mortality (Warburton et al., 2006). The World Health Organization (WHO) ranked the sedentary lifestyle as the tenth leading cause of mortality in the world. According to the INSERM Report in 2008 - "Physical Activity: Contexts and health effects", all-cause mortality can be reduced from 2% to 58% depending on the intensity and type

of activities. The sedentary population also suffers from several chronic diseases, such as cardiovascular disease, type 2 diabetes and certain types of cancer and osteoporosis (WHO 2010). Consequently, researchers have deployed the PA recommendations to solve the current health epidemic resulting from physical inactivity (Dugdill et al., 2009; Puska et al., 2004). The study of Leitzmann et al., (Leitzmann et al., 2007) that considered more than 250 000 people between 50 and 71 years showed that moderate activities for at least 3 hours/week, or vigorous activities for at least 20 min 3 times/week reduce 30% the risk of death. In order to meet these recommendations of exercising, a better monitoring of daily and weekly activities of the individual is crucial. PAR systems are used to study the influence of exercise routines aimed at improving human activity profile (Van Sluijs et al., 2007). For instance, different types of exercise stimulate different types of physiologic activity and different health effects: Long-duration activities improve the endurance for example, playing ball games improves balance and postural musculature and neurocoordination, swimming has great impact on upper body musculature, exercising at the gym improve muscle strength and aerobic activities have important effect on the cardiovascular system (e.g., improved stroke volume, decreased peripheral resistance) (Pärkkä, 2011; Powell et al., 2011). In order to establish the appropriate activity profile recommended for a particular subject, an advanced monitoring of subjects activities is demanded, to record his daily life activities, his types of sport and exercise and his transportation trips (Krumm et al., 2006). Another important goal is the early assistance to motivate individuals to adopt a healthy lifestyle and to adhere to a PA target (Pärkkä, 2011). By providing the subject sufficient and reliable information about his care process, the PAR system motivates him to do self-care and thus be aware and follow the changes in his health status (Lorig et al., 2001).

- Prevention of physical disabilities in aging population:

A study conducted in the US in 2000 showed that 1 person in 5 of people aged 65 and over suffers from a physical disability and 1 person in 10 is unable to take care of his personal needs (Waldrop et al., 2003). Another study showed that continuous care provider for this portion of population costed the US economy $123 billion dollar each year. This number is estimated to rise to $346 billion dollars in 2040 (Corbin-Jallow et al., 1999). Clearly, any preventive plan to reduce these costs and promote a better health-related quality of life is of major benefit. Activity profiling systems can play an important role by detecting the presence of functional and mobility limitations as well as determining the context of a situation so that an appropriate assistance can be provided. In other words, the central controller responsible of the data processing may either send an alert call to the nearest emergency center and inform them about the health status of the patient being monitored and may help detecting early warning signs of a disease or a health threat (Teng et al., 2008).

2.2 Different approaches for activity recognition.

Studies of activity recognition can be separated into two classes: visual monitoring that considers external non-wearable sensors and on-body monitoring that uses wearable sensors. Traditionally, visual monitoring studies that employ vision-based systems with cameras and video recordings had been at the forefront of this work (Aggarwal et al., 2011; Moeslund et al., 2001).

2.2.1 Visual monitoring sensors

In this approach, the devices are installed in pre-defined points of interest at home or laboratory environments to monitor user's activity profile. This consists of a visual motion capture system including video cameras in addition to sensors embedded on target items which the user is supposed to interact with (sofa, table, washing machine, etc.) (Aggarwal et al., 2011). In general, the continuous recognition of human activities is performed in this case by detecting and analyzing the types and durations of ongoing activities in a recorded video (a sequence of image frames). This visual monitoring technique is efficient in accurately detecting complex activities and is practical since it discharges the user from any controller device on his body. However, it suffers from several issues that caused the shift toward the use of wearable sensors. First, it can only be applied in constrained settings and environments. Additionally, it is unable to follow the user's movement once he/she is out of the perimeters captured by the camera. Another issue is the complexity of this technique in calibrating each camera separately along with the long cost of recording, storing and processing images and video records throughout days and weeks. To overcome the limitations of visual monitoring, some researchers turned toward the fusion of both visual and on-body monitoring systems to benefit from both modalities and enhance the recognition performance (Tao et al., 2007; Viéville et al., 1990). On the other hand, the field of activity recognition is currently witnessing the use of small and wearable cameras that are automatically triggered to capture videos or images from the user's environment (such as GoPro Hero3+ for video records and SenseCam camera). Some studies consider the cameras as recognition tools, as it is the case in our study, while others utilize them for validation. For the former group, cameras are used along with wearable motion sensors to detect performed activities (Doherty et al., 2013). For the latter group, cameras replace the human observer that normally accompanies the subject during the experiments and can serve for annotating the performed activities via the recorded pictures (Carlson et al., 2014; Kerr et al., 2013). The main reason for using such technology is to ensure free-living scenarios. In this dissertation, we conducted a similar study and proceeded to explain in depth the related benefits of such annotation type.

2.2.2 Wearable sensors

The study of (Khusainov et al., 2013) overviewed 488 research works across different application areas (gait assessment, fall detection, classification of ambulation activities, etc.) and identified 433 studies

(more than 88%) that adopted the use of wearable sensors and the rest 55 studies used fixed sensors. This review also deployed a graphical distribution of the reported studies with respect to the sensor types and applications (Figure 2.1). This distribution demonstrates that researchers are interested in various sensor types depending on the intended application. For example, as shown in Figure 2.1, the accelerometer is the most used sensor in several applications such as ambulation activities, leisure/communication activities and fall detection. However it is not used for the prediction of diabetic foot iteration. Instead, sole pressure, humidity and body temperature sensors were preferred in such application.

2.2.2.1 Microelectromechanical systems (MEMS) and sensors development

Recently, with the prompt development of MEMS, the size, weight, and cost of sensors have decreased significantly and their computational power has been improved. With the advance of storage technology, the memory capacity of the sensors has grown. And with the development of wireless technology, the continuous data transmission in body area networks has been possible (Perez et al., 2010). These improvements have made the wearable electronics devices reliable tools for long term health recording (Mathie et al., 2004). Numerous off-the-shelf sensors with novel specifications are currently available on the market, such as tracking and fitness devices, acceleration and GPS modules, HR monitors, smart watches and wrist bands activity trackers, etc. (Lo et al., 2014). Researchers are now exploring these sensors with their main purpose: extracting useful information for recognizing the subject's activities and running context-aware applications (Perez et al., 2010), allowing people to interact with these devices and track and analysis their PA during their day.

2.2.2.2 Motion sensors

2.2.2.2.1 Accelerometer-based activity monitors

The physical principle of an electronical accelerometer is that of the spring mass damper system. Technically speaking, an accelerometer has 2 fundamental parts: a housing and a mass, that, while tethered to the housing via the springer can still move. Once the accelerometer is displaced, the mass responds by forcing the spring to either stretch or compress proportionally to the applied acceleration. Based on Hook's Law, the motion of this central section is measured by how much the spring stretches proportional to the force applied (Bao, 2000). The output produced by the accelerometer is a voltage charge measured using differential capacitance and is then translated into a PA and physiological metrics (Figure 2.2).

Different types of accelerometers exist in the market: piezoelectric, piezoresistive or capacitive transducer, they can be uniaxial (single axe), biaxial (dual-axes) or triaxial (3D axes) depending on the number of axes sensible to the acceleration. Accelerometers have also different amplitude ranges and different sampling rates (Chen et al., 2005).

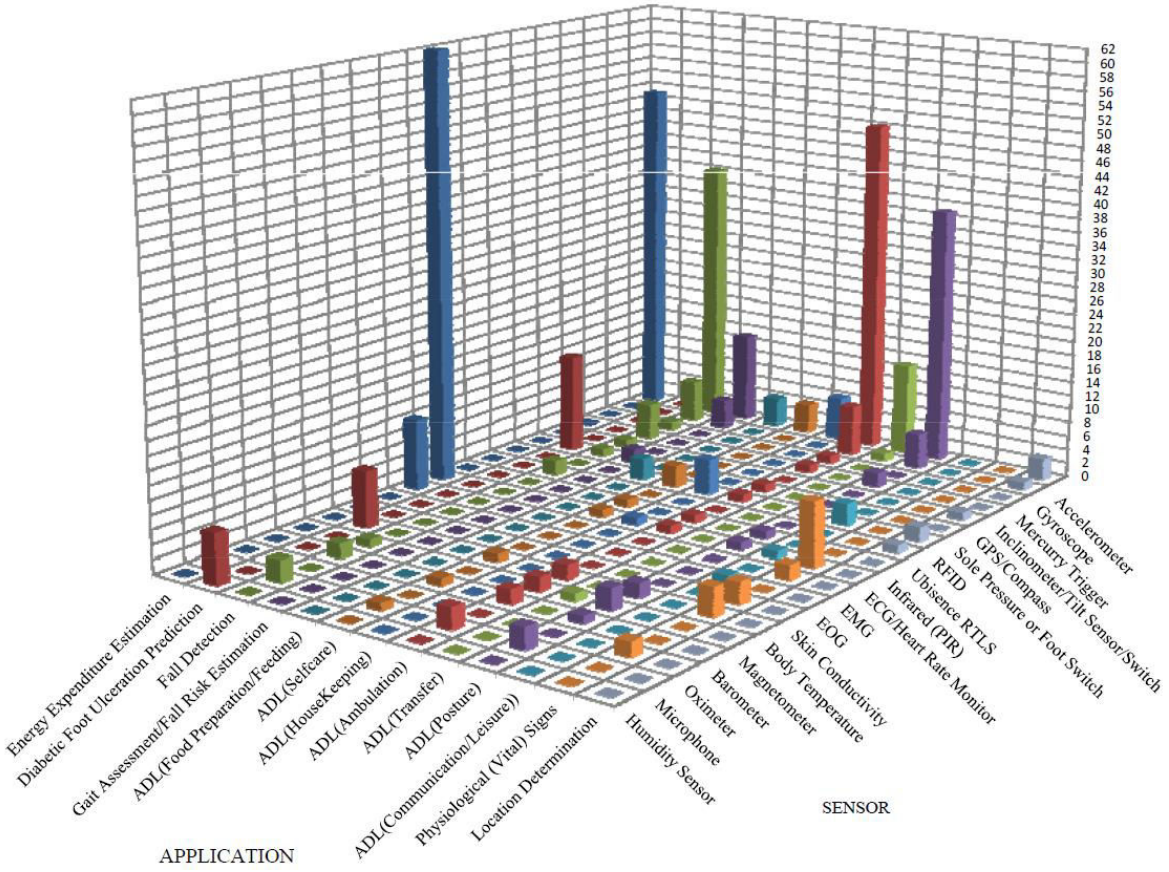


Figure 2.1 - 3D representation of the number of studies based on application and sensor types.
Source : (Khusainov et al., 2013).

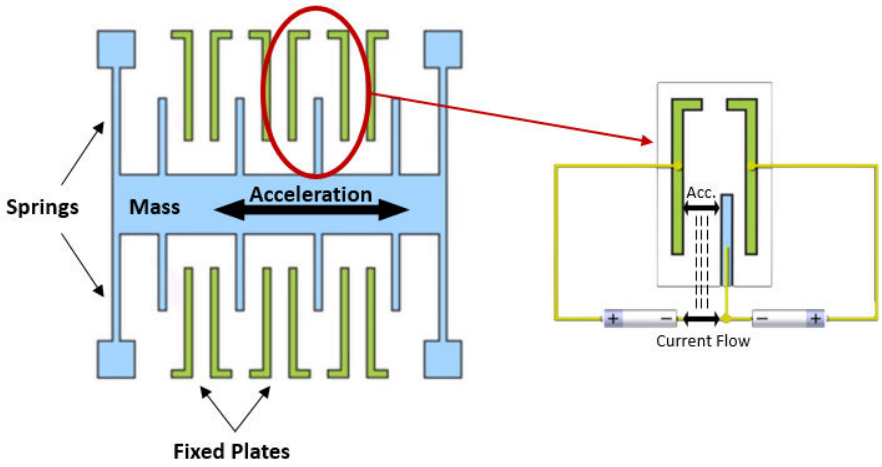


Figure 2.2 - The principles, fundamental elements and functioning of an accelerometer.
This illustration is adapted from an instructional video on YouTube: "How a Smartphone knows up from down (accelerometer)".

The impact of these technical specifications on the PAR performance is well studied. For example, the study of (Maurer et al., 2006) demonstrated that the sampling frequency should be at least 20 Hz to maintain better accuracy that outperforms devices with lower sampling rates. Besides, it was demonstrated that the acceleration range at body extremities can reach ±12g and at the torso and hip can

reach ±6g (Bouten et al., 1997). However, the literature suggests that the acceleration amplitude of ambulation activities ranges solely between [2;-2] g. Promising results can thus be obtained with sufficiently a ±2g range of acceleration data (DeVaul et al., 2008). Being light-weighted, inexpensive and with low-power consumption, the accelerometers are the most commonly used sensors for the objective assessment of PA.

2.2.2.2.2 Pedometers

A pedometer has the internal mechanism of an accelerometer. Its principle is simply based on counting accumulated steps in response to a determined body acceleration level. In research, it has been proved that, contrary to questionnaires, pedometers can objectively assess PA. Due to their small size, low cost, light weight and noninvasiveness, pedometers are widely used within the field and have been endorsed by national programs. Indeed, pedometers helped promoting better health status for subjects by encouraging them to reach a required number of steps per day and thus become more active. However, despite their accuracy in monitoring ambulatory activities, pedometers have a lot of weaknesses in monitoring other types of activities involving upper body movement. Furthermore, pedometers are unsuitable to assess gait types (patterns), walking speeds and behavior as demonstrated by (Le Masurier et al., 2003) and (Storti et al., 2008). In latest years, manufacturers have embedded the step counters with the inertial and GPS sensors in cellphones, which may improve the accuracy of pedometer applications in the field of activity recognition (Bonomi et al., 2012).

2.2.2.3 Vital Sensors

2.2.2.3.1 Hear Rate monitors

The second most common signal studied in the field of activity recognition, after the acceleration, is the HR. This physiological parameter plays an important role in assessing the disease and health status of the monitored subject. The commercially available HR monitors consist of chest strap transmitter and a receiver watch. Their storage capacity allows a recording on several days with a sufficient sampling frequency. Literature presented two diverging points of view regarding the role of HR in this field. While some have reported a number of limitations in the use of HR signal in the PA monitoring, others have proved that higher overall accuracy is achieved when adding the HR data to the recognition system. (Chen et al., 2005; Lara et al., 2012b) stated that factors other than exercise can affect the HR measure. For example, HR can be influenced by an emotional stress and consequently its variation does not reflect the effort exerted during a particular exercise. They also declared that the measure of HR signals requires more sensors in addition to the accelerometer and thereby increases the system cost and the level of obtrusiveness. Furthermore, (Tapia et al., 2007a) pointed out that the latency of the HR onset with respect to the transition of PA are the cause of systematic classification errors for vigorous activities of short durations such as walking upstairs. However, (Maurer et al., 2006) found that vital signs would

definitely yield more reliable classification results since they provide more clear patterns to distinguish between activities such as ascending and walking.

2.2.2.3.2 Electrocardiogram sensors

Advanced electrocardiogram (ECG) monitoring systems are currently emerging. Wearable ECG monitors are widely used for the ambulatory monitoring of the cardiac activity and for the capture of rhythm disturbances for patients with cardiac diseases. The diagnosis of cardiac disease relies on the study of the rate and the regularity of the heart beats, information extracted from the ECG signal (Parkka et al., 2006). Recent research has also focused on the development of wireless body area network (WBAN) platforms with integrated ECG monitoring. For instance, (Lo et al., 2005) developed a WBAN system incorporating 3-Lead ECG, 2-Lead ECG strip and a blood oxygen saturation levels (SpO2) sensor. Another WBAN platform dedicated for ECG data streaming and monitoring has been deployed (Wang et al., 2010). The CardioNet system of (Ross, 2004) records ECG data subsequently and transmits it in real-time to a central server in order to be analyzed by a care provider. The role of these WBAN systems is also to signal a warning if a negative event is detected as in the study of (Standing et al., 2001). The MIThril platform captures various physiological signals: the ECG data, the galvanic skin response and the skin temperature (Pentland, 2004). Similarly, the project 'Context-aware cardiac long term monitoring' (CALM) aims at providing a daily-life surveillance of ECG and blood-pressure for assisted treatment of patients with cardiac problems (Jatobá et al., 2008). Different systems are now commercially available providing healthcare for cardiac patients. A well-known device is the Actiheart (CamNtech Ltd, Cambridge, UK) system, a chest worn monitoring device. The ECG signal processor of the Actiheart assesses data collected from two standard ECG electrodes attached to the chest. The Shimmer Company offers as well ECG recordings via the ECG module and its 6-leads electrodes.

2.2.2.3.3 Body temperature

Body temperature is also one of the most studied vital signals in the field of PA monitoring. As reported by (Mukhopadhyay, 2015), the measurement of body temperature is used as health indicator (e.g., fever), and help studying the illness conditions in stroke, heart attacks and shock. The temperature of the subject can be measured using a simple thermistor embedded in a chip. Several studies suggest that a combination of body temperature signal together with other physiological and acceleration signal is of great interest, since it allows a more accurate assessment of different types of moderate and vigorous activities (Anliker et al., 2004; Gjoreski et al., 2015; Parkka et al., 2006; Winkley et al., 2012). (Cvetkovic et al., 2015) showed however that the body temperature doesn't add any valuable information to their recognition model since it is highly correlated with the HR. The temperature sensor has also been integrated in commercially available monitors. For example, the Zephyr BioHarness BT and the BodyMedia sensor which are used by (Gjoreski et al., 2015).

2.2.2.4 Location sensors

In recent days, the emergence of GPS in people's smartphones and Personal Digital Assistant (PDA) devices has made it more possible than ever to track their location and mobility mode. Researchers have been interested in recognizing the context of people's activity based on GPS data. They focused on detecting the place of the subject, on classifying his outdoor movement and on mapping his transportation routine (Zheng et al., 2008). The study of (Liao et al., 2007) infer the subject behavior on the basis of GPS data that provides more context information. They first extract the locations of the subject, predict his movement among these locations and then identify his activities on each location. Similarly, the study conducted by (Huss et al., 2014) was able to differentiate transport modes as either walk, bicycle, car, bus or train from speed metrics alone. The advantage of GPS devices is however limited to the outdoor applications, as they don't work properly indoor. Consequently, this device is usually used along with the accelerometers as in (Ellis et al., 2014; Nguyen et al., 2013; Parkka et al., 2006; Reddy et al., 2010) to better predict PA performed indoors and outdoors by combining data from these two devices. (Reddy et al., 2010) used successfully a mobile phone with a built-in GPS receiver and an accelerometer to detect whether the subject is stationary, walking, running, biking, or in motorized transport.

2.2.2.5 Environmental sensors

Although inertial and vital sensors were commonly used in the activity recognition systems, other environmental parameters such as microphones, lux measurements, ambient temperature and humidity sensors can also add useful contextual information to the system (Ermes et al., 2008; Foerster et al., 1999; Tapia et al., 2006). For instance, a low intensity of light and audio level can be good indicator that the subject is sleeping. In recent days, various miniature sensors are now integrated in one small device, offering the possibility to monitor human activities on the basis of different types of signals with no complexity cost. As an example, the BodyMedia armband device from SenseWear is a multi-sensor device and is used in our experiments. A detailed description of this device and the integrated sensors can be found in Chapter 3 (section 3.4.1.5).

2.3 Process of activity recognition

The process of recognizing human activities consists of the sequence of problem statement, signal processing, pattern recognition and machine learning methods that can automatically classify outputs (activity labels) based on inputs (sensor data) (Lara et al., 2013). It is a multi-stage process that the majority of the studies follow (Freedson et al., 2000; Lara et al., 2012b; Parkka et al., 2006; Tapia et al., 2007a). It can be briefly described like so (Figure 2.3): Firstly, once a target set of activities to identify has been set, the appropriate sensors are selected and their positions on the body are determined. Next

the streams of sensor data are collected while a sufficient number of participants are following a well-defined protocol of activities. The raw data recorded is organized and if needed pre-filtered out of artifacts at the signal-preprocessing stage. This stage also includes the segmentation of the signal into a number of small time windows of interest that contain an activity. For each window, a set of features (e.g. time- and frequency-domain features) that characterize the signal is computed. In the features selection stage, the aim is to identify the best representative subset of features to be used as inputs for the classifier algorithm. In general, optimal features are those that minimize the inter-class correlations and the intra-class variability. Afterward, the selected features are used as predictor variables to design, structure and train the classifier. A performance evaluation stage is then used to test and re-design properly the trained model to further improve its accuracy, if necessary. In the classification stage, the final version of the system that is able to assign correct class labels to new inputs is finally considered. This section overviews the development of the activity recognition system following the mentioned process step by step.

2.3.1 **Data acquisition and pre-processing**

The development of an activity recognition system requires at first collecting the data properly and preparing the acquired signal for features extraction. While some applications require the coupling of several sensors (e.g., accelerometers, gyroscopes, HR monitors), others might adopt only one sensor that can capture multiple parameters (e.g., 3D accelerations on X, Y and Z axes). A vector of multivariate time series with different sampling rates is thus obtained and a synchronization stage across the sensor outputs might be needed. Moreover, artefacts might contaminate the signal for various reasons, such as a loosely attached sensor or external artefacts, noises and vibrations. In this signal filtering stage, the amount of artifacts must be minimized. For example, (Wang et al., 2011) tested several filters: median (MF), Butterworth low-pass (BLF), and Kalman filters (KF) to de-noise 3D acceleration signals recorded during regular walking gait. They found out that the KF showed the largest Signal to Noise Ratio (SNR) and correlation coefficients (R) values, followed by the MF, and finally BLF filter. Similarly, an example of noise reduction for ECG signal is presented in the study of (Celka et al., 2005). The Wiener filters (WF), wavelet de-composition (WD), and principal component analysis (PCA) were investigated; it was shown that the choice of the filters depends on the nature of the signal such as the stationarity, the statistics of the information, the origin of the signal and its noise.

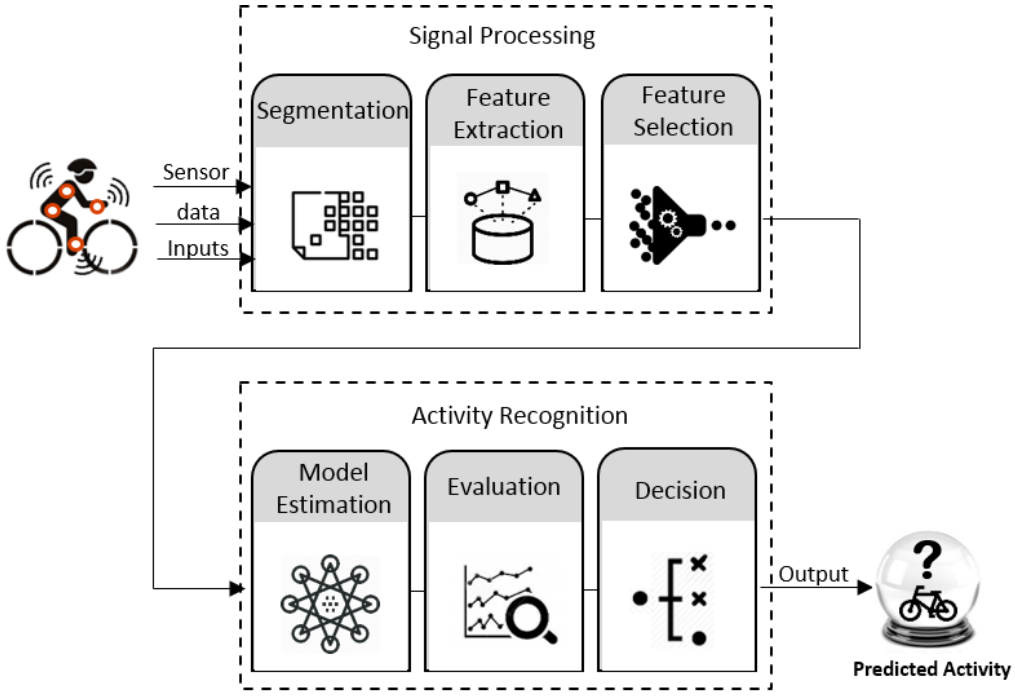


Figure 2.3 - Typical design of an activity recognition chain.

2.3.2 Data segmentation and windowing

In order to effectively extract the features from sensor data, a stage of windows segmentation is firstly required. It consists of partitioning the measured time series into time windows, each including a potential activity. Subsequently, the algorithm of activity recognition is applied separately to each window. Two key factors that can influence the classification performance must be carefully selected: the window length and the windowing technique. For instance, the longer the window length the longer the delay of the recognition system and lower the temporal resolution of the PAR system. Conversely, the shorter the window length the less the contained information within the window that describes the activity. Ranges of different window lengths have been tested in the literature: 1s (Ermes et al., 2008), 2s (Skotte et al., 2012), 4s (Maurer et al., 2006), 5s (Altun et al., 2010), 10s (Parkka et al., 2006), 20s (Sasaki et al., 2016) and up to 60s (Rosenberg et al., 2016). It was shown that a 30-s trade-off window length must be adopted for physiological signals (Tapia et al., 2007a) whereas 5s window length is optimal for acceleration signals (Altun et al., 2010). Furthermore, two windowing techniques were also tested in literature. They can be either overlapping (Bao et al., 2004; Lara et al., 2012b) or non-overlapping (Altun et al., 2010). The choice between the two techniques depends for instance on whether the intended application must handle transitions or not. (Khusainov et al., 2013) tested the response of a number of classifiers for different window sizes with no-overlap, as presented in Figure 2.4. From this figure, it can be inferred that, for all tested methods, accuracies first arise with the increase of window length before going downwards, showing a potential trade-off.

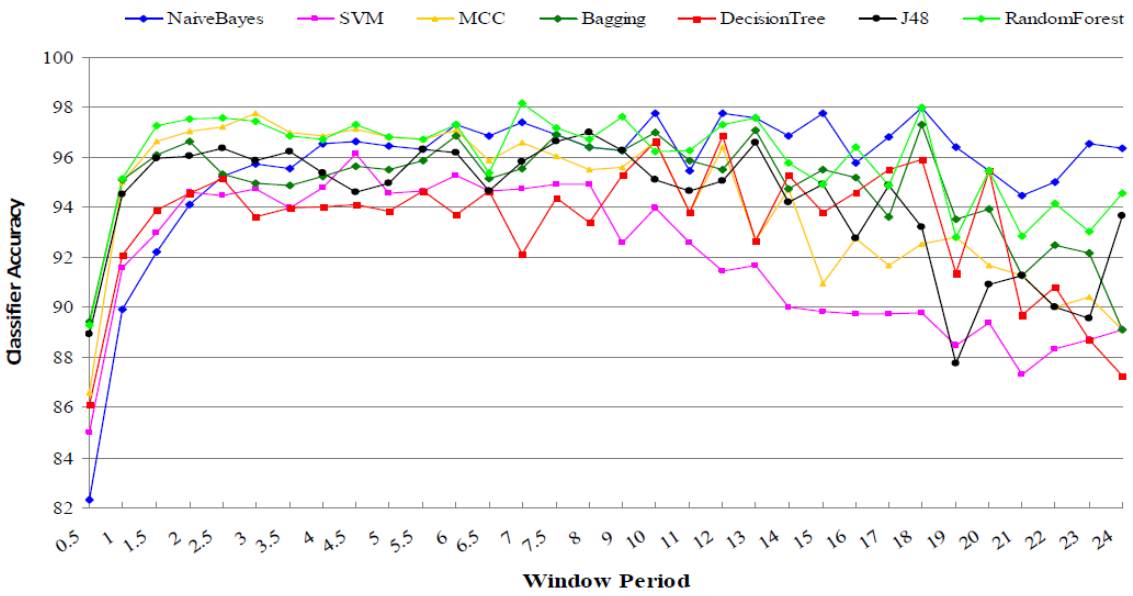


Figure 2.4 - Performance (measured as accuracy) of different classifiers as a function of varying window sizes. SVM: Support Vector Machine, MCC: Multi Class Classifier, J48: C4.5 (as implemented in WEKA). Source:(Khusainov et al., 2013).

2.3.3 Features Extraction

Sensor data, most specifically acceleration signals, cannot be directly used in their raw form to extract the pattern they represent since they are characterized by their oscillation and fluctuation. Applying feature extraction techniques aims at characterizing time series data before their classification into activity patterns. The features filter the relevant information contained in each time window and are then used to feed the classification algorithm. The types of features can be divided into two main categories (Figure 2.5): Time domain features, such as mean, max, min, standard deviation, variance, skewness, kurtosis, correlation coefficients and frequency domain features such as total energy, peak frequencies of the Fourier Transform (FFT) and power in different bandwidths. Other approaches based on the Wavelet Transform (WT) and the Discrete Cosine Transform (DCT) were also proposed in the PAR literature, as reviewed by (Lara et al., 2013).

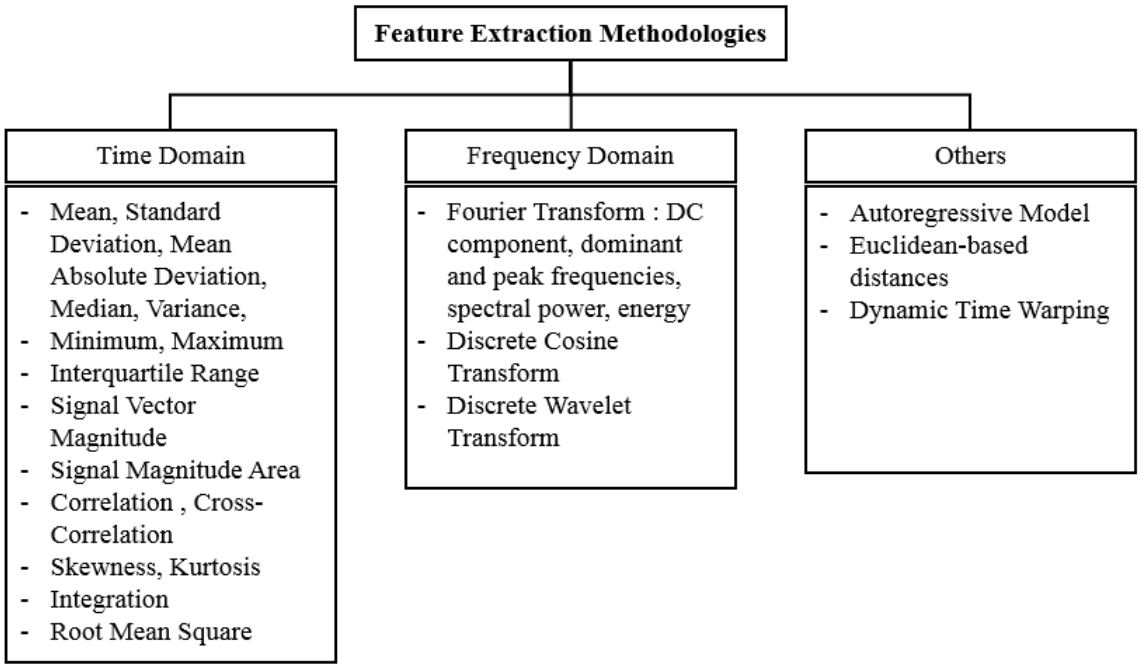


Figure 2.5 - Range of features extracted from time-series data. Adapted from (Figo et al., 2010; Lara et al., 2013).

2.3.4 **Features Selection**

When directly used in classification, a large feature set with redundant or irrelevant information can increase the computational load, slow down the classification process and negatively affect the recognition accuracy. Thus, implementing techniques for selecting the most appropriate features is an important factor towards effective classification. (Fang et al., 2012) demonstrated that the performance of the classification algorithms strongly depend on the set of features. The objective of the feature selection is to minimize the number of features without a significant loss of classification accuracy. The best discriminative feature set should have high correlations between repetitions of the same activities and across different subjects but should also show small correlations between different activities. Different methods have been presented for choosing the optimal feature set for classification. Some methods are based on selecting a subset from the existing features and are referred to as feature selection methods. Others combine the original features and create a new set of features and are referred to as dimensionality reduction methods. Examples of feature selection methods that are widely used in general machine learning problems are the Minimum Description Length (MDL) and the Sequential Forward Search (SFS). (Zongker et al., 1996) evaluated the performance of 13 different feature selection methods and found that the Sequential Forward Floating Selection (SFFS) was the most powerful method for the classification of hand-printed characters. In PAR studies, the Correlation-based Feature Selection (CFS) and the Minimum Redundancy and Maximum Relevance (MRMR) techniques were commonly used in literature (Jatobá et al., 2008; Maurer et al., 2006; Pande et al., 2014). In the study of (Altun et al., 2010), after selecting a large set of 1170 features, the principal component analysis (PCA)

technique was used to reduce the number of features to 30. Similarly, (van Hees et al., 2013) used PCA to reduce the number of the extracted features while (Leutheuser et al., 2013) have chosen to integrate the whole set of 152 features directly at the input of the classifiers. Interestingly, (Kate et al., 2016) compared the performance of different classifiers in two cases: when all the computed features are taken together and when only selected features are used. This test resulted in higher accuracy when reduced number of features is considered. For this reason, it is necessary to reduce the dimension of the features set, especially when it is blindly selected with no prior knowledge on the usefulness of each feature separately. This problem statement is further discussed in Chapter 4, in which a novel contribution to directly select appropriate measures while avoiding the long process of extracting and selecting features is presented.

2.3.5 **Classification**

The classification in any PAR system consists of assigning an activity class label (such as resting, walking, running, etc.) to each instance in the set of features. To do so, a classifier is trained using a set of learning data and is then evaluated using a set of unobserved testing data. The principle of most classifiers is to draw decision borders that are able to separate the classes in the feature space. The learning approaches can be divided in two categories, either supervised or unsupervised. For supervised learning, the set of training data must be labelled before training the classifier whereas for unsupervised learning, the classifier handles the unlabeled data and automatically identifies a number of clusters, each of which matches certain activity. While the majority of PAR systems work in a supervised manner, few classify activities using unsupervised techniques. Some other works used a semi-supervised techniques by labeling part of the data, since the annotation might be practically difficult and time consuming (Ali et al., 2008; Huynh et al., 2006; Trabelsi et al., 2013).

2.3.5.1 Building the classifier

The classification algorithms that are used in PAR systems varies from simple threshold-based methods to more complicated methods such as Naïve Bayes (NB) and Artificial Neural Networks (ANN). These advanced classification algorithms and their computational statistics are nowadays implemented in newly developed software featuring machine learning tools and dedicated for recognizing patterns based on input features. For instance, the Waikato Environment for Knowledge Analysis (WEKA) (Hall et al., 2009) is a well-known machine learning platform. It offers the possibility to build and evaluate a variety of common learning algorithms implemented in the software. In order to choose the best classifier for a given PAR application, a compromise must be achieved between the computational complexity of the classifier and its recognition performance. According to the study of (Lippmann, 1989), the decision tree classifier is the most efficient and the K-Nearest Neighbor classification algorithm is the most demanding in terms of computational memory requirements for the classification.

Most of the literature studies tested and compared separately single classification algorithms before building their recognition system. In some studies, more than one classifier were used together to infer predictions (Kuncheva, 2004). It was shown that this alternative, the multiple classifier systems (MCS), provides better accuracy in classifying class labels despite the additional complexity and computation cost required for learning these systems. The different classification techniques that were considered in the most recent and relevant studies can be found in Appendix A. Three classifiers that were manipulated in our work: K-nearest neighbors (K-NN), Naïve Bayes (NB) and Decision Tree (DT), are described here, while a description of other machine learning techniques used in the field of PAR can be found elsewhere (Pärkkä, 2011; Preece et al., 2009).

2.3.5.1.1 K-Nearest Neighbors

The K-nearest neighbor classifier is based on instance-learning algorithm, where classification decision is made by identifying the activity class of the majority nearest neighbors to the instance. This classifier requires the determination of two parameters, the value k that represents the number of nearest neighbors taken into consideration and the distance metric, such as the most commonly used Euclidean-distance. A normalized feature space which dimension corresponds to the number of features is first constructed from all-training samples. Unknown instance is also represented in the feature-space in order to compute its distance towards each training sample. We created, in Figure 2.6, a simple illustration of the KNN technique in a 2D feature space. KNN classifier is a simple algorithm yet computational expensive and requires a lot of memory, since no model is created but all training samples are used. Nevertheless, KNN classifier was widely used in literature and its performance in recognizing PA had been demonstrated in several studies (Altun et al., 2010; Bao et al., 2004; Leutheuser et al., 2013; Shoaib et al., 2013; Trabelsi et al., 2013). With this classifier, (Zhang et al., 2006) were able to differentiate between falls and common activities with more than 95% accuracy. Similarly, the KNN classifier in the study of (Bussmann et al., 2001) successfully classified a wide range of postures and movements with an accuracy up to 93%.

2.3.5.1.2 Decision Tree

The decision tree resemble to a hierarchical scheme. It consists of nodes, branches and leaves. At each node, a decision is made following threshold-based rule applied on the features set. The principle of DT can be thought of as a number of consecutive questions on which the answers will lead to the class label to be assigned. An example of a DT is illustrated in Figure 2.7. In the training phase, the set of rules are determined, simultaneously generating the complete elements of the DT. Its structure can be manually created or can be based on automated algorithms (Parkka et al., 2006). This type of classifier are easy to understand and are widely used in activity recognition applications. DT and random forest (RF) classifier, a combination of multiple decision trees, are the preferred classifiers in the PAR field.

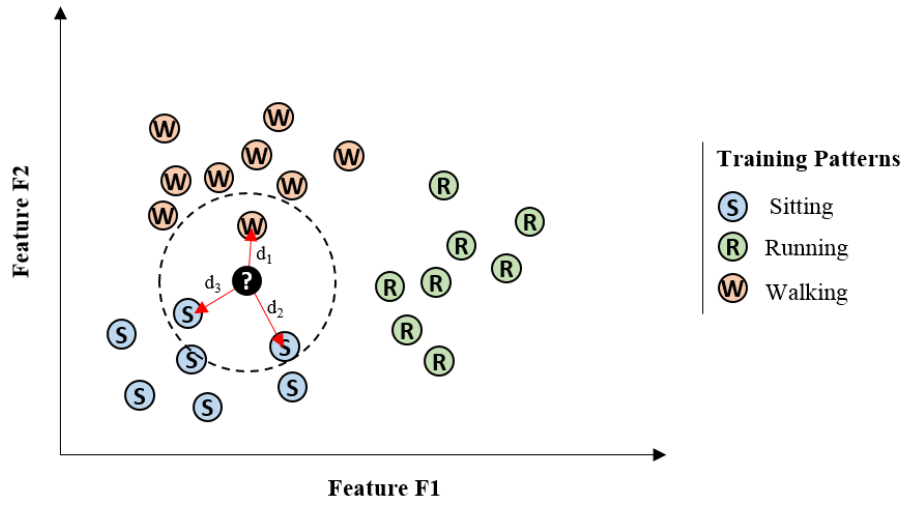


Figure 2.6 - Illustration of the k-nearest neighbor principle simplified in a 2D features space. Training patterns from three classes: sitting, running and walking are plotted and an example of an unknown instance which is assigned with a sitting label is presented.

In different studies, DT and RF achieved highest recognition accuracy in detecting PA when compared to other classifiers (Ellis et al., 2014; Kerr et al., 2016; Maurer et al., 2006; Parkka et al., 2006). For example, (Bao et al., 2004) performed the classification of 20 everyday activities using four classifiers: Conventional decision table, KNN, DT and NB and compared their performance. They showed that DT was the best voted classifier with an average accuracy of 84%.

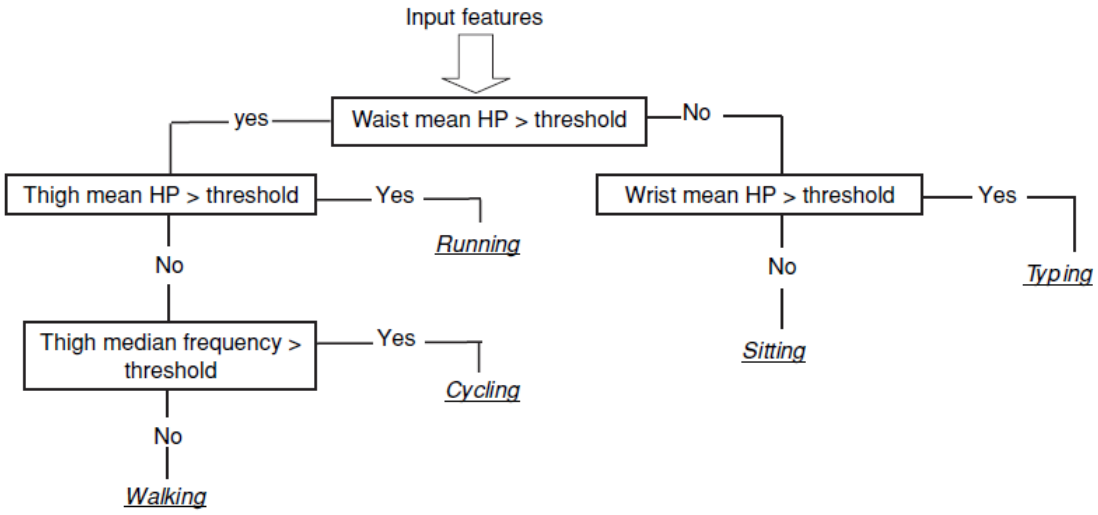


Figure 2.7 - Hierarchical example of a decision tree based on simple threshold rules from four input features: waist, thigh and wrist means, along with thigh median frequency. All features having been (High-Pass) HP filtered. Source: (Figo et al., 2010; Preece et al., 2009).

2.3.5.1.3 Naïve Bayes

The Naïve Bayesian approach is based on the probabilistic Bayes' theorem. It estimates the conditional probabilities for unknown sample using likelihoods from patterns in each activity class. This type of technique can only do classification and not regression. To simplify the computation in a NB classifier, the features are assumed to be conditionally independent of each other given the output. Yet in truth, physical and physiological signals are highly correlated in detecting PA variations. Although the assumption of independency doesn't hold in PA problems, the NB classifier is well-studied in PAR field due to its robustness and fast computations compared to other classifiers (Cleland et al., 2013; Kate et al., 2016; Ravi et al., 2005). Different results were reported after using the NB classifier in PAR application. While some studies showed quite good accuracy matching with other classifiers (Maurer et al., 2006; Ravi et al., 2005), others found reduced performance (Bao et al., 2004).

2.3.5.2 Training the classifier

In order to effectively train the model, two main approaches can be adapted: the percentage split and the cross-validation. Both can be applied in an inter-subject way or intra-subject way. The former approach divides randomly the entire dataset into two sets: training and testing. For example, half of the data are used for training the model and the other half is used for the test. The k-fold cross-validation is more recommended since the dataset is divided into k folds of equal size and a train-test cycle is performed at each iteration k. In the first iteration, the fold 1 constitute the testing set and the remaining [2 to k] folds constitute the training set. The cycle is repeated k times in a manner that each fold at its time is used for testing. This procedure is described in Figure 2.8.

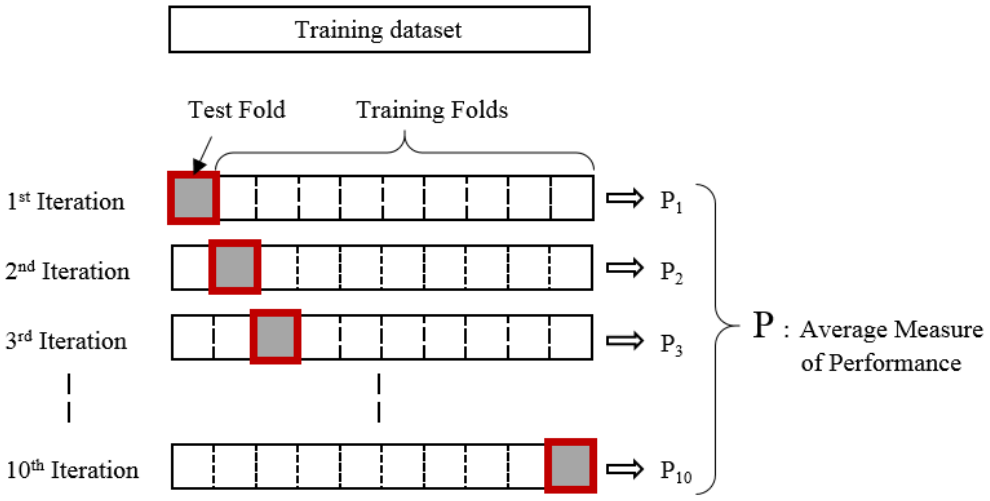


Figure 2.8 - Schematic representation of the iterations of a k-fold Cross Validation with k =10.

When an inter-subject evaluation is considered, the model is trained using data from all available subjects except one excluded subject that is used for testing. While an intra-subject evaluation is based

on the data from only one subject. In this case, a part of the samples of this specific subject is used for training, and the remaining samples of the same subject are used for testing.

2.3.5.3 Testing the classifier

In order to evaluate the quality of the classifier, several evaluation metrics can be used. The accuracy is the main standard metric that reflects the overall classification performance for all classes. Other metrics are also frequently used alongside the accuracy, such as the Receiver Operating Characteristic (ROC) curves, the confusion matrix and its related measures; True Positives (TP) and True Negatives (TN), False Negatives (FN) and False Positives (FP), the precision, the sensitivity (or recall), the specificity and the F-score. These metrics are computed as follow: TP: The number of positive samples classified as positive. TN: The number of negative samples classified as negative. FN: The number of positive samples classified as negative. FP: The number of negative samples classified as positive.
Accuracy = (TP+TN)/(TP+TN+FP+FN). Sensitivity (Recall) = TP/(TP+FN). Specificity = TN/(TN+FP). Precision = TP/(TP+FP). F-score = 2 x (Precision x Recall)/(Precision + Recall).

2.4 Challenges in the research field of activity recognition

To conceive and implement smart systems that are able to continuously monitor daily human activities is not trivial. Several challenges face the design, the development, the implementation and the manipulation of a PAR system. In contrast to other machine learning problems such as computer vision and speech recognition, PAR cannot be formulated in a clear problem definition since its structure relies on many issues as it will be cited next. A change in one of the factors turns a given PAR problem into a completely different problem statement.

2.4.1 Decision of the activity classes

In previously published studies, there is no common definition of the activities under investigation making it impossible to compare and validate one PAR system in regard with another. Each system depends on the chosen set of activities to be recognized. For example, a set of ambulatory activities (walking, running and cycling) could be identified using only accelerometer-based systems, whereas the identification of talking activity might need an additional microphone sensor. The PAR research community is trying to adopt a clear taxonomy of activities that can serve as a good reference for organizing the PAR studies. Therefore, researchers are invited to define the set of activities that fits their application on the basis of well-known resources such as the compendium of Ainsworth[1] (Ainsworth et

[1] The Adult Compendium of Physical Activities is a coding scheme that links a five-digit code to each specific PA by rate of EE. It was developed to enhance the comparability of results across studies.

al., 1993) and the Activities of Daily Living (ADL) index tool (Katz et al., 1970). In 2014, the International Activity Monitor Database (IAMD) project was deployed after reporting the consensus of 18 PA experts for the need a global accelerometry data repository (Wijndaele et al., 2015). This project aims at standardizing the data collection, processing and analytic procedures, developing an online infrastructure and a methodological comparison work among international studies. A similar data pooling project occurred in 2008: the International Children's Accelerometry Database (ICAD). This database holds information on approximately 26 000 children from 20 studies worldwide (Sherar et al., 2011). Its objective was 'to allow new analyses, to generate a clearer understanding of predictors of activity, activity–disease associations, and the types and levels of activity that should be promoted to maximize health benefits'.

The range of activities that are currently under study can be grouped into eight categories based on their nature: posture (sitting, lying and standing), ambulation (walking, running, going up/down stairs), transportation (riding a car/bus, cycling, driving), housekeeping (sweeping, cleaning), leisure (reading, watching TV, working on PC), self-care (eating, brushing teeth), communication (using the phone, making phone calls), sports (playing basketball) and into four categories based on their intensities: sedentary, light, moderate to vigorous, vigorous.

2.4.2 **Selection of the sensors**

The choice of the number, the type and the location of the sensors that subjects will tolerate is essential. Sensors characteristics may highly vary due to environmental factors, temperature sensitivity, calibration issues, time deviations or shifts in sensors positions. For example, the barometers that are particularly sensitive to the location and the altitude and the magnetometers that are sensitive to ferromagnetic influences, both require recalibration before each time used. Moreover, the sensing equipment embedded in mobile phones must be carefully employed since they can be influenced by the position of the phone on the body. A main criterion that must be taken into consideration while designing a PAR system is its obtrusiveness. The monitoring system should be comfortable, light, easy to wear and noninvasive in order to be effective in practice. Between a system that requires carrying a heavy rucksack, and a system that is worn as a sport watch, the later system will be certainly accredits by the user.

In terms of energy efficiency, an important feature that current applications are successfully integrating is that sensors may be dynamically added or removed based on each application requirements during monitoring. This technique is referred to as dynamic and adaptive sensor selection (Shoaib et al., 2015). In such applications, intelligently switching the sensors on and off saves their battery life and thus provides a long-lasting recognition system. For example, (Ryder et al., 2009) designed a recognition system that relies on the GPS along with the accelerometer only when the subject is outdoor otherwise the GPS is turned off since it is not useful indoors.

2.4.3 User flexibility

The ideal PAR system is the one that is monolithic, meaning that is able to successfully identify the samples of activities performed by an unseen user. This system is based on person-independent training making it robust against variability among users instead of considering individual training. It is challenging to delicately find the most discriminative feature set among classes but also more steady across different users. In many cases, a personalized system is used instead since each user performs the activities on his own manner. In this case, the person-dependent classifier is re-trained by integrating the data of the new user too. This option is less practical, yet it can with no doubt enhance the performance of the system (Tapia et al., 2007a). For example, the periodicity of cycling activity depends on the age, the fitness level and the weight of the subject. An elderly person would surely cycle differently that a young person. A classifier trained for other subjects regardless their characteristics may not perform well on a specific subject with newly introduced characteristics. Further considerations in designing a person-dependent or person independent systems are still needed (Lara et al., 2013).

2.5 Estimation of energy expenditure

It is of great interest to study the intensity, the frequency and the duration rather than just the type of the activity (Figure 2.9). In fact, it is not sufficient to only determine the sequence of the performed activities during the day. Yet, quantifying each activity and its related level of energy demands is also required to afford a real comprehension of the etiology of among others, obesity, diabetes and cardiovascular diseases (Leonard et al., 2002; Schoeller et al., 1990). EE estimation would allow to substantiate the physiological and biological impact of a given PA pattern, regardless its frequency, duration, intensity and type. From an epidemiological point of view, it's particularly interesting to standardize the PA measure, allowing a unified study of the relationship between PA and health. "There is ongoing uncertainty about the various dimensions which are biologically relevant and important for health but one key dimension is overall physical activity energy expenditure which is naturally the most important consideration for weight loss or maintenance." (Thompson et al., 2015). This strong relationship between the health outcomes and the EE clarifies the reason why investigators for more than a decade are dedicating their efforts to precisely measure EE. A quick overview of some of the health domains interested in EE estimation is presented next.

2.5.1 Importance of monitoring EE

- Mortality

A study conducted by (Myers et al., 2004) that tested the ability of different clinical and exercise test variables in predicting the mortality for men population, showed that the exercise capacity and the EE from weekly activity are the only significant predictors of all-cause mortality. It reported that a weekly

increase in activity by 1000 Kcal, similar to an increase of 1 metabolic equivalent task (MET) in physical fitness, leads to a 20% survival benefit. Similarly, another study (Hu et al., 2004) examined the health-related inactivity in middle-aged women population. It conferred that less than 1 hour/week of exercise conveys to an increase of 52% in overall mortality and a doubling of mortality from cardiovascular disease. These findings, supported by several other investigations witness the dose-response relation between the expenditure of energy and the health benefits. The highest is the amount of expended energy (so the level of activity), the lowest is the risk of premature mortality.

- Diabetes Diseases

Furthermore, monitoring the EE helps assessing a number or risk factors for hyperlipidemia, hyper-tension, cardiovascular disease and type 2 diabetes (Warburton et al., 2006). In particular, EE during leisure time had shown to be protective against the development of Non-Insulin-Dependent Diabetes Mellitus (NIDDM) as reported by (Helmrich et al., 1991; Lynch et al., 1996). The study of (Helmrich et al., 1991) showed that a decrease of 6% in the incidence of type 2 diabetes is achieved due to 500 kcal/week of increase in the EE. Accordingly, the study of (Lynch et al., 1996) pointed out that activities, regardless their duration, must be of an intensity of at least 5.5 MET in order to effectively reduce the risk of NIDDM.

- Nutrition Assessment

Malnutrition in patients could destroy the lean muscle mass, increase the risk of infection, affect the immunity level, cause organs dysfunction and even conduct to mortality (Druyan et al., 2012). Monitoring the physiologic and metabolic responses of these patients is critical for providing optimal nutrition support and clinical care and preventing complications related to under- or overfeeding. To do so, an accurate assessment of EE is essential to control the caloric needs of the patients and facilely design the required nutrition programs during illness (Psota et al., 2013).

This short documentation regarding the health effects of the monitoring of EE involves a heavy burden on EE researchers to deliver appropriate measures of energy that serve most health and fitness organizations.

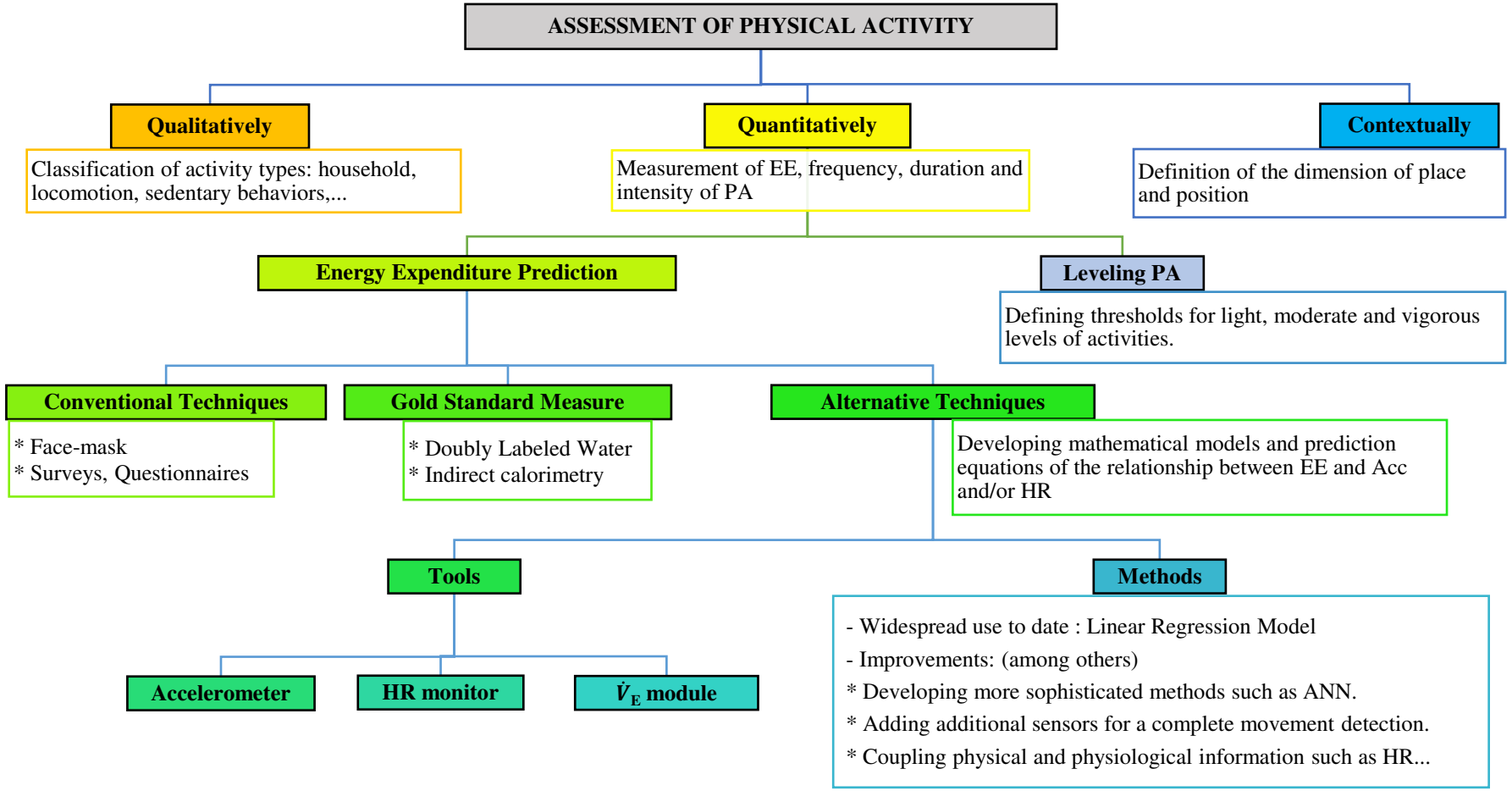


Figure 2.9 - Hierarchical Framework showing the assessment of physical activity and the prediction of energy expenditure.

2.5.2 **Methods and tools for measuring EE**

This section provides an overview of the different available techniques that have been used for measuring the EE including: doubly labeled water, room respiration calorimeter, portable pulmonary gas exchange systems, HR monitors, motion sensors and self-reports. Figure 2.10 shows the distribution of these techniques with respect to their ease of assessment and their degree of precision as presented by (Khusainov et al., 2013).

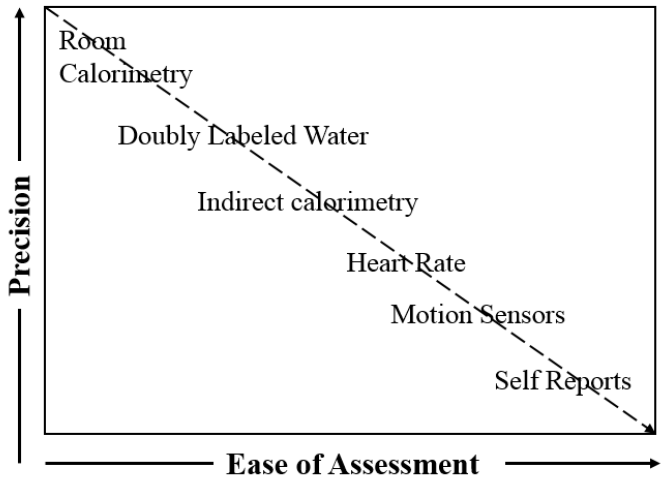


Figure 2.10 - The distribution of the EE methods based on their ease of assessment vs precision. Source: (Khusainov et al., 2013), (modified).

2.5.2.1 Direct Measurements

The accurate measurement of EE relies on either the doubly labeled water (DLW) technique or the room respiration calorimeter technique (Melanson Jr et al., 1996; Valanou et al., 2006). The gold standard is the DLW that measures the EE with small error rates (less than 10%) as reported by (Laporte et al., 1985). This technique is based on labeling specific isotopes of hydrogen and oxygen atoms in the water drunk by the subject. The analysis of the proportion of no-metabolized isotopically labeled water contained in the urine reflects the overall energy expended by the subject during 1 to 2 weeks of measurement duration. It is a practical, non-cumbersome approach since the subject is free from any device fastened on his body. However, the isotopes are expensive and their implementation requires special expertise (Murphy, 2009). Furthermore, the DLW technique can only measure the overall long-term EE. The principle of a room respiration calorimeter is the measure of the heat produced by the subject that is locked inside the calorimeter room. Since it is not possible to measure directly the consumption of oxygen ($\dot{V}O_2$) and the production of carbon dioxide ($\dot{V}CO_2$), a current of air is pumped through the calorimeter room and by examining its volume and composition it is possible to quantify the amount and type of the respiratory gas produced by the subject (Atwater et al., 1905). Although the

method is very accurate, it is costly, complicated and only usable in laboratory conditions and thus impractical in free-living context (Melanson Jr et al., 1996; Pärkkä, 2011).

2.5.2.2 Indirect Measurements

Indirect methods for the measurement of EE include the measure of $\dot{V}O_2$ and $\dot{V}_E$ which correlate with heat production (Schoeller et al., 1990). As examples of the most common tools that can do this job, the COSMED K4b2 portable pulmonary gas exchange analyzer (COSMED Srl, Rome, Italy) and the Oxycon Mobile (Yorba Linda, CA). These systems require the subject to wear a breathing face mask covering the nose and mouth of the subject, a mouthpiece and the equipment for the gas exchange analysis. A gas calibration must be performed prior to each test. Then during the tests, $\dot{V}O_2$ is measured breath-by-breath; the air flow passes through the mask towards the gas analyzers where the concentrations of oxygen and carbon dioxide in the inspired and expired air are determined (Valanou et al., 2006). It was shown that these monitors are reliable and deliver valid measures of the respiratory gas exchange and thus the EE (McLaughlin et al., 2001; Rosdahl et al., 2010). Although, these portable respiratory devices can be used outside the laboratory, it is still not practical for everyday measurements since they require wearing a breathing mask. Indeed, they were used as ground truth for evaluating the accuracy of the EE prediction methods from accelerometry data for example (Crouter et al., 2006a; Dugas et al., 2005; Kozey et al., 2010).

2.5.2.3 Mathematical modeling and prediction equations

Traditionally, the estimation of EE relied on direct observations or subjective self-reports techniques such as activity diaries, questionnaires or 24-h recalls. However, these techniques provided poor, biased and inconsistent measures of EE (Sallis et al., 2000). The latest technologies through movement sensors, HR monitors or $\dot{V}_E$ measurement modules and the advanced analytic tools (e.g., computational methods, artificial intelligence) are now being applied to overcome the limitations of subjective techniques and obtain better estimations.

The total daily energy expenditure (TEE) consists of 3 main components, as presented in Figure 2.11: the resting metabolic rate (RMR), the thermic effect of food (TEF) and the physical activity EE (PAEE) (Ravussin et al., 1992). The RMR represents the energy expended while at rest. It is the minimal amount of energy required by the organs to live; to pump blood and to sustain the metabolic activities of cells and tissues. It represents 60 to 75% of TEE, and is relatively constant over days. The TEF reflects the EE needed by the body for the digestion of the foods after eating a meal. It is estimated to be equal to 10% of TEE. Of all the components of TEE, the one that varies the most and influences the measure of EE is PAEE. It can be calculated as follows: PAEE = TEE – RMR – TEF (Lagerros et al., 2007).

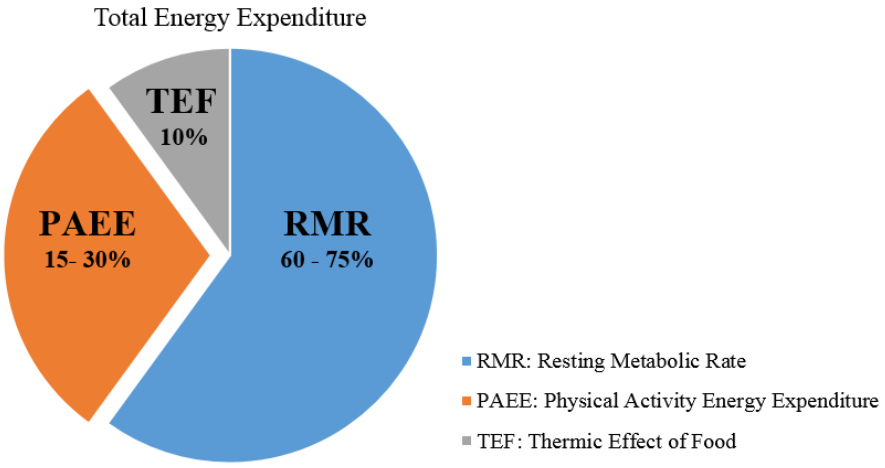


Figure 2.11 - Partition of the energy expenditure components.

2.5.2.3.1 EE estimation from acceleration

Acceleration can be assumed as an image of EE during an activity. Actually, when the limbs and torso are moving, the muscular forces respond and reflect the expended energy. Accelerometers, acting as motion counters, have thus been used to estimate the activity caloric expenditure. The raw output from accelerometers is calibrated to meaningful indicator of EE. The total caloric expenditure can thereby be predicted from the RMR, calculated from the height, weight, age and gender of the subject, summed to the activity caloric expenditure obtained from the acceleration signal. A number of prediction equations ranging from simple linear regression to more sophisticated modeling techniques have been tested to accurately establish the correlation between acceleration and EE (Table 3.1). The evaluation of these techniques is done by comparing the estimated values of EE to the corresponding measures of the DLW and indirect calorimetry serving as references (Melanson Jr et al., 1996; Valanou et al., 2006). The use of accelerometers in the EE estimation has been reliable and witnessed the publication of numerous articles. However, there is still a gap in finding a prediction method that stands accurate across all intensities (Lyden et al., 2011). Several studies have suggested the combination of acceleration and HR measures in order to overcome this challenge (Kozey et al., 2010; Pande et al., 2014; Schneller et al., 2015; Zakeri et al., 2013). Furthermore, it was interestingly proved that by identifying the PA type and then estimating the EE on the basis of the recognized activity provides more valid estimations than conventional approaches (Trost et al., 2012). The reason is that the relationship between the activity counts and EE is more significant within a specific type of activity rather than between different types of activities (Bonomi et al., 2009; Kozey et al., 2010).

2.5.2.3.2 EE estimation from heart rate

The physiologic variable of HR presents a quasi-linear association to $\dot{V}O_2$ and can be thus used as indicator of EE (Bassett 2000). The optimum technique that is widely adopted in literature to estimate EE on the basis of the HR is called the HR-Flex. This method is represented by a curve relation between EE and HR (Figure 2.12). The HR Flex point is the average between the minimum HR during exercise and the maximum HR during rest. This point splits the curve into two parts. The values of HR above the HR Flex point have a linear relationship with the EE. Below this point, the relationship is more variant, and thus the value of resting HR is considered (Leonard, 2003)

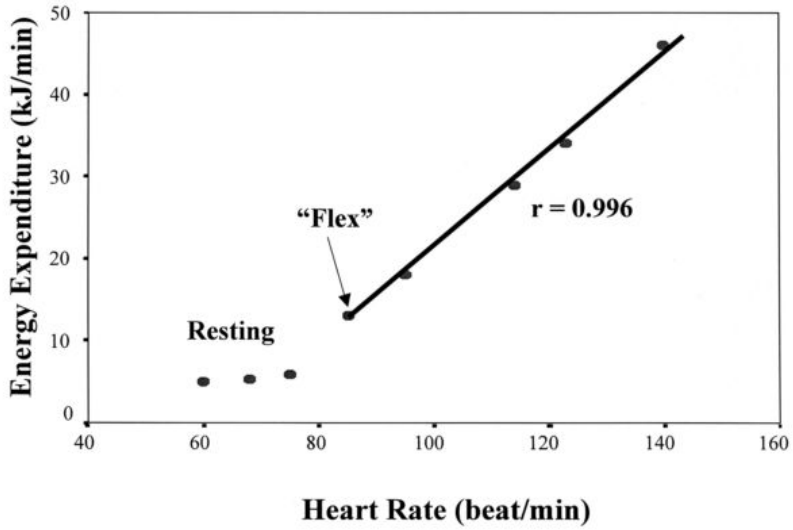


Figure 2.12 - An example of the relationships between HR and EE, illustrating the flex-HR threshold point that separates between the linear and nonlinear regressions. Source: (Leonard, 2003).

2.5.2.3.3 EE estimation from ventilation

Since 1980, the strong relationship between pulmonary ventilation, oxygen consumption and EE had been demonstrated (Bernard et al., 1979; Boutellier et al., 1985; Durnin et al., 1955). Yet recent developments in sensors technology that are able to replace face masks, have led to a renewed interest in measuring and using $\dot{V}_E$ as an index of EE. Light-weighted, easy-to-use and non-invasive, the electromagnetic coils system is the promising module to do this job. The principle of these coils is based on measuring the anteroposterior displacement of the rib cage and abdomen and the axial displacements of the chest wall and the spine (Dumond et al., 2017; Gastinger et al., 2010a) (Figure 2.13). Other types of sensors using Respiratory Inductive Plethysmograph (RIP) technology were also employed to monitor $\dot{V}_E$ (Clarenbach et al., 2005; McCool et al., 1986; Witt et al., 2006). They are based on two-belt sensors assembly fitted on the ribcage and the abdomen of the subject, in such way the belt dimensional change is correlated to the respiration movement (Konno et al., 1967).

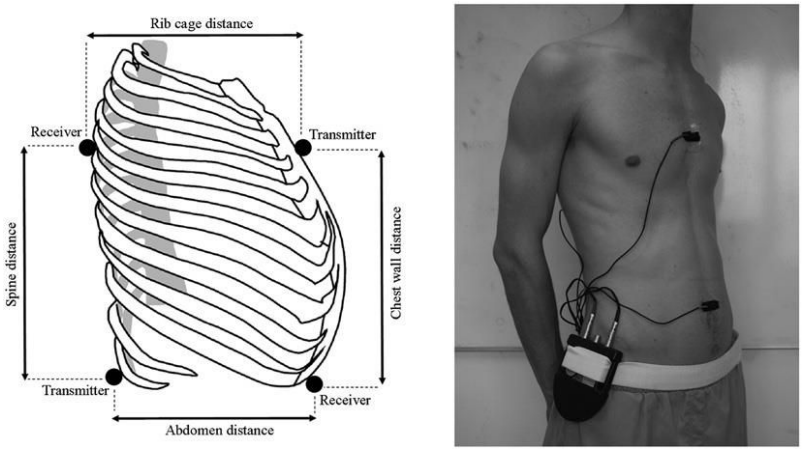


Figure 2.13 - Illustration of the electromagnetic coils placement on the subject as designed and employed by (Gastinger et al., 2010a).

However, a calibration maneuver that preceded the distances variations measurement is necessary. A wide range of researches have been conducted in this field to validate such technique in the aim of delivering surrogate measure of EE. A survey of the different approaches that measure $\dot{V}_E$ in order to estimate EE and the recent investigations in this domain can be found in (Gastinger et al., 2014). Attempts are recently being made to integrate this technology into instrumented garments and thereby shift the EE monitoring towards free-living environment outside the laboratory. Interestingly, a prior study (Gastinger et al., 2010b) compared between $\dot{V}_E$ and HR to determine which variable is better correlated to $\dot{V}O_2$. Their results showed that the relation between $\dot{V}_E$ and $\dot{V}O_2$ was significantly higher than between HR and $\dot{V}O_2$ during steady state activities and different intensities of exercise. Commenting on these findings, they argued that 'the use of HR presents a number of limitations for delivering accurate measure of EE, since external factors (such as emotional stress, high ambient temperature, high degrees of humidity, dehydration, body posture, or disease) might affect the HR variations without affecting $\dot{V}O_2$ and thus EE'. As such, we sought to contribute to this field by studying the breathing waveform output of a chest strap device that was used in our experiments (Chapter 5 (section 5.5)). Due to the observed limitations of this approach, we then proposed, as a perspective, a modeling $\dot{V}_E$ mechanism method, which will be clarified in the conclusion.

2.5.3 **Overview of existing EE studies**

Table 2.1 summarizes some of the most recent and most cited studies within the field with an overview of the used experimental tools, the validation system, the developed estimation method and the main results concluded from each study. As it can be observed, current literature has investigated different EE prediction methods (RE, ANN, etc.) that are based on acceleration and HR signals. These algorithms are validated against EE values measured by different ground truth techniques: DLW, RC or face mask devices. In this table, we have shown that the ActiGraph device is the most commonly used accelerometer in the field of EE estimation. A detailed description of this device can be found in Chapter

3 (section 3.4.1.3). We would like here to shed the light on the Activity Counts (AC) data type that is used to find the relation between acceleration and EE. AC is a well-known term in EE domain, which is used to indicate that the data are filtered using manufacturers' proprietary algorithms. Explaining what counts are, ActiGraph support team points out: '*Counts* are a result of summing post-filtered accelerometer values into epoch "chunks."' They added: 'The filtering process by which counts are produced is proprietary to ActiGraph' (ActiGraph, 2016). Similarly, many other activity monitors deliver their outcomes in AC. As a result, for a same input measuring acceleration signal, each type of activity monitor will generate different count values depending on their closed proprietary algorithms (Vähä-Ypyä et al., 2014). A vast majority of the studies that used the ActiGraph device, established cut-points algorithms for assessing the intensity of PA and developed count-based prediction techniques for EE estimation (Crouter et al., 2006a; Freedson et al., 1998; Lyden et al., 2014). Yet, such studies remain narrow in focus dealing only with count data and the developed equations won't be valid once a different accelerometer is employed.

2.6 Conclusion

This chapter covered a wide range of works done in the fields of activity recognition and EE estimation. We surveyed the different wearable sensing systems that were extensively investigated in both fields. However, despite their ability to identify a range of activities, we have shown that the full potential of the PAR systems still needs further investigation in design, development and validation issues. Generally, data acquisition is done in a controlled environment and rarely do studies investigate the response of their systems in a naturalistic environment. We hypothesis that once evaluated in a naturalistic environment, the recognition performance of the existing systems will decrease dramatically. We believe that the accuracy of the PAR system employed in real-world applications would decline drastically when trained using dataset of controlled activities. The data acquisition protocol is thus critical and a comprehensive study is needed to ensure flexibility of the recognition systems to out-of-the-laboratory applications. Furthermore, the feature extraction process in most of the existing studies consists of computing a big set of features with no sufficient investigation of their contribution. For the purpose of this work, these issues were carefully evaluated in the next chapters in which we were motivated to (i) extend the activity recognition field to a more realistic scenario and (ii) propose a well-adapted type of features to enhance the recognition process.

Table 2.1 - Summary of the state-of-the-art in EE estimation studies.

Reference /Year	Experimental Tools	Validation Tools	Participants (N°/Age)	Data (AR/AC)	Methods	Activities	Findings
(Schneller et al., 2015)	- Actiheart - ActiGraph - ActivPAL	Cosmed K4b2	14/ 20 - 40	AC[1]	RE[2] per activity	Lie, Stand, walk, run, stairs	Accuracy of PAEE estimation improved after an activity type recognition.
(Pande et al., 2014)	- Smartphone Acc - Polar HR monitor	Cosmed K4b2	7/ 6 - 10	RD[1]	Bagging ensemble technique	Lie, sit, walk (slow, fast)	Superior accuracy and correlation for EE during low- and higher-energy activities.
(Zakeri et al., 2013)	- ActiGraph GT3X - Actiheart	- RC[1] - DLW[1]	69/ 3 - 5	AC	CSTS[1] and MARS[1] models	static, low, moderate, active, very active play	Both CSTS and MARS models are strong prediction models for EE.
(Kozey et al., 2010)	- ActiGraph GT1M, - Polar HR monitor	Oxycon mobile	277/ < 40	AC	Linear RE	Household, sporting, locomotion	Compendium should be updated with additional data and more activities.
(Staudenmayer et al., 2009)	- ActiGraph M7164	Cosmed K4b2	48/ 35 (11.4)	AC	ANN[1] model	Low, locomotion, vigorous, household	Successful implementation of an ANN to estimate PAEE.
(Rothney et al., 2007)	- Biaxial Acc - ActiGraph - IDEEA[1]	RC	102/ 18 - 70	RD	ANN model	Stretch, pace, walk, jog, type, bike, play cards.	With ANN models, minute-by-minute and total EE estimations can be improved.
(Crouter et al., 2006b)	- ActiGraph M7164	Cosmed K4b2	48/ 35 (11.4)	AC	2-model RE	Low, locomotion, vigorous, household	The model measures walk/run and lifestyle activities EE better than single-RE.
(Crouter et al., 2006a)	- ActiGraph - Actical AMP31	Cosmed K4b2	48/ 35 (11.4)	AC	17 published RE + 1 ME[1]	Low, locomotion, vigorous, household	The equations are valid for the tested activities, but not on a wider range.
(Puyau et al., 2004)	- ActiWatch - Actical	RC	32/ 7 - 18	AC	Power Function	computer, clean, walk, jog, run, aerobics	A power function best described the relationships between counts and EE.
(Zhang et al., 2004)	- IDEEA	- Mask - RC	37/ 33 (13)	AC	Ainsworth compendium	sit, stand, lie, run, breakfast, lunch, treadmill	IDEEA provides a suitable method for estimating EE with high accuracy.

[2] RC: Room Calorimeter, DLW: Doubly Labeled Water, AC: Activity Counts, RD: Raw Data, CSTS: Cross Sectional Time Series, MARS: Multivariate Adaptive Regression Splines, RE: Regression Equation, ME: Manufacturer Equation, ANN: Artificial Neural Network, IDEEA: Intelligent Device for Energy Expenditure and Activity.

Chapter 3
EXPERIMENTAL PROTOCOLS

Data acquisition is one of the critical procedures in PAR systems. The existing activity recognition models marked high level of performance in detecting various activities of daily living. However, their validation was restricted to a controlled laboratory environment instead of being evaluated under real-life conditions (Altun et al., 2010; Leutheuser et al., 2013; Pober et al., 2006). People perform activities in a different manner depending on the environment that they are in (Parsons, 1974). In a restricted laboratory context, normal human behaviors would be affected. Therefore, data acquisition has to be done under unconfined realistic conditions rather than laboratory controlled states. Unfortunately, few studies have been interested in collecting data in a naturalistic environment but they only relied on limited datasets and on specific types of activities (Parkka et al., 2006; Skotte et al., 2012). In this chapter, the idea is to shift from bounded to pragmatic protocols thereby generating a realistic dataset that ensures further validation of the recognition systems in this context.

Initially, our work was established using a dataset collected in a controlled protocol. This 'basic' protocol encompasses a limited number and types of activities performed by 8 subjects for a predefined duration under the supervision of a researcher. Because it is important to ensure a naturalistic data collection protocol, we believe that this first dataset is not sufficient for a complete validation of the recognition system. For this reason, we carried out a new expanded protocol that could offer important insight on the system response in challenging circumstances. This 'pragmatic' protocol encompasses a larger number of daily activities each performed by 20 subjects in three separated sessions (standardized, semi-standardized and free-living sessions). This chapter begins by presenting the objectives of the triple-session pragmatic protocol that we designed. Then, it overviews the most relevant protocols existing in the literature. Following that, it gives a detailed description of the materials used and the experimental procedures followed in both the basic and the pragmatic protocols to generate the datasets that we used throughout our studies.

3.1 Objectives of the triple-sessions pragmatic protocol

This novel protocol will allow:

1. To assess the accuracy of human activity recognition algorithms developed on both controlled and naturalistic data and then tested on the latter. In our previous basic protocol, we only collected controlled data in supervised conditions. In this protocol, new data were collected in unsupervised context as well. Our goal was to study the effects of the naturalistic environment on the performance of the recognition methods.
2. To evaluate the response of different sensors in recognizing the activities in the context of free-living. We pointed out, in the previous chapters, that a variety of sensors with embedded accelerometers are used nowadays in the human activity recognition domain. Along with the accelerometer, researchers were interested in testing other physical and physiological parameters and studying their abilities in recognizing activities or enhancing the recognition performance (e.g. GPS, HR...) (Lara et al., 2012b; Wu et al., 2011). However, there is still no significant evidence on the best voted device in this field. By combining several commercially available sensors, this experimental protocol aims at validating each sensor/parameter compared to other sensors/parameters.
3. To enrich the international data repository of activity monitor with a dataset suitable for a further validation of existing recognition systems. This dataset that we have generated will have the potential to disclose the limitations of previously developed classification algorithms caused by controlled labelling of the activity datasets. It will allow a clearer understanding of human behaviors and thus new analyses would be conducted.

3.2 Literature Review on existing data collection protocols

The data collection protocols considered in the literature can be divided as follows, regarding the circumstances in which the dataset was collected:

3.2.1 Standardized controlled protocol

This type of protocol is being used in the vast majority of current studies (Altun et al., 2010; De Vries et al., 2011; Farkas et al., 2011; Leutheuser et al., 2013; Mannini et al., 2010; Pober et al., 2006). The datasets are recorded in a confined laboratory settings where the participants receive instructions on the sequence of activities to perform in a fixed duration and intensity and where the experimenter define the beginning and the end of each activity in a specific order. In these circumstances, the recognition algorithms achieved high accuracies in discriminating various types of activities, as reported in the related studies (Altun et al., 2010; De Vries et al., 2011; Farkas et al., 2011; Leutheuser et al., 2013; Mannini et al., 2010; Pober et al., 2006; Staudenmayer et al., 2009). The laboratory environment is adequate for collecting data in a balanced way, meaning, the contribution from all the considered activity types is the same. Datasets issued from supervised protocol can serve as proof of concepts for the system development but further validations are still needed using less controlled dataset. Actually, some previous studies have demonstrated a degradation in the system performance when the classification algorithms, developed on controlled dataset are tested on free-living datasets (Foerster et al., 1999; Skotte et al., 2012). The reason behind this is that the PA of the participants are being performed in an unnatural manner in the controlled settings. Their normal behavior was affected by what is called: Hawthorne phenomena (Parsons, 1974). Consequently, the signals recorded in free-living conditions and those recorded in laboratory conditions are considerably different. A simple example of what we have seen in literature is when the participant in the study of (Leutheuser et al., 2013) was asked to do the vacuuming activity using the right hand since the sensor was fastened on the right wrist. The researchers of this study argued that, otherwise, no suitable information would be delivered by the right wrist sensor signal. However, such studies remain narrow in focus dealing only with controlled activities. The algorithm development must be issued from datasets in uncontrolled context where they will ultimately be applied.

3.2.2 Semi-standardized protocol

The semi-standardized protocol, as the name suggests, is a compromise between the standardized controlled protocol and the unstandardized free-living protocol. Some recognition systems work in a semi-standardized fashion when it is not feasible to implement a complete unsupervised protocol. In this case, no specific instructions are given to the participants. They are asked to perform, on a given period of time, a series of activities at self-selected intensities, duration and order, without imposing minutes of rest between the activities. However, an experimenter that labels the start and end of each activity

accompany the participants during the data acquisition (Lara et al., 2012b; Schneller et al., 2015). In order to produce more representative data, a task-based protocol can be deployed as it is the case in the study of (Grant et al., 2006). In this approach, the experimenter asks the participant to perform a desired task, for example, "solve an Intelligent Quotient (IQ) test written on a paper" instead of asking him to "sit and write". This type of protocol will subsequently generate more naturalistic activities. The semi-standardized protocols usually serve for the validation of the classification algorithm developed in the standardized protocol (Schneller et al., 2015).

3.2.3 **Unstandardized free-living protocol**

Once the participant is left to freely perform his usual routine in his own environment with no restrictions on his behavior during the day, this fashion is referred to as unstandardized free-living protocol. Although it is preferable to collect datasets in a naturalistic environment as much as possible yet researchers in this domain face many challenges regarding its feasibility in tracing the participant's activities during his day. Different approaches to annotate the collected datasets have been considered in the existing free-living protocols. In this section, we summarized the most pertinent state-of-the-art free-living protocols while a wider overview of the related studies can be found in appendix A. We pointed out in Tables 3.1 and 3.2 the characteristics that we believe are strengths and limitations related to each considered experimental protocol. 1) The studies of (Bao et al., 2004; Ermes et al., 2008) are based on the participant-labelling approach where the participant is asked to hand annotate throughout his day the type, the start and the end of the performed activities or tasks. 2) Whereas in (Foerster et al., 1999; Nguyen et al., 2013; Sasaki et al., 2016), the direct human observation approach was considered, an experimenter accompanying the participant tracks his movement in real-time and notes the performed activities on a worksheet or a PDA. 3) Other studies rely on a previously validated sensors to serve as reference tools. For instance, (Skotte et al., 2012) considered a pressure sensor settled in the hip pocket to detect and validate the sitting posture in free-living context. 4) Although these approaches have been commonly used since decades, a better and more accurate approach of validation have been recently practical and feasible. It is based on videos recordings and images' captures that allow observation of the participant's environment and surroundings and thus allow convenient identification of the type of performed activities. Along with the existing studies (Bourke et al., 2017; Carlson et al., 2014; Kerr et al., 2013; Rosenberg et al., 2016) that considered the last approach, our free-living protocol relies on the use of a camera that can automatically capture pictures from the user's perspective. We believe that this approach provides important benefits over the previously cited approaches. The video/or camera system allows the detection of transitions, brief and sudden activities that can be missed with the direct observation or the user annotation approaches. Furthermore, the observational study of (Sasaki et al., 2016) identified general and wide categories of activities (e.g. sedentary, household, and recreational) while camera approach would allow a more refined and expanded scope of categories if employed. For instance, (Kerr et al., 2013) were interested in studying the sedentary behavior, so they relied on the

SenseCam camera that successfully provided domain and contextual information about the sitting behaviors, such as sitting driving a car, sitting using computer screen, sitting eating and watching TV and sitting reading. Camera system, contrary to the other approaches, also permits for a post-observation coding with the possibility to recheck the collected dataset and improve or revise labelled data. In addition, the direct observation periods reported in the literature had short durations as in (Foerster et al., 1999) because of the complication on the observer and on the participant as well while accompanying him throughout his daily life. Yet with camera systems, it is more practical to collect data on prolonged periods with less burden to participants and researchers, as it is the case with (Carlson et al., 2014; Kerr et al., 2013).

Table 3.1 - Summary of the strengths and limitations of the state-of-the art free-living protocols.

	Study	Strengths	Limitations
User-Annotation	**(Bao et al., 2004)**	• Classification of 20 different types of activities (Walking, sitting, watching TV, Running, Stretching, etc.). • Five accelerometers on the thigh, ankle, arm, wrist and hip.	• Subjects executed an obstacle course consisting of a series of activities listed on a worksheet. • The subjects self-labeled the start and stop times of activities following the prescribed order.
	(Ermes et al., 2008)	• 4 hours of recording in uncontrolled conditions. • Supervised and unsupervised data were used in the training phase compared to the use of only supervised data. • Wide variety of activities: nine daily and sports activities.	• The user used a PDA to annotate his activities during uncontrolled protocol. • The user wore a rucksack during the experiments.
Direct Observation	**(Sasaki et al., 2016)**	• 3 observers trained to use the PDA. • Train recognition algorithms on laboratory and free-living accelerometer data and test their accuracy in classifying the activity types in free-living older adults.	• The concerned population: 35 older adults over 65 years of age. • One type of sensor is used: the ActiGraph GT3X+ • 2 - 3 hours of direct observation. • Recognize general categories of activities: standing, sedentary, locomotion, household, and recreational.
	(Nguyen et al., 2013)	• Classification of common PA types and durations and the estimation of EE. • A combination of accelerometers and a GPS (with the portable calorimeter mask).	• The subjects were asked to perform randomly and spontaneously 5 different activities indoors or outdoors over 2 hours.
	(Foerster et al., 1999)	• Retest the reliability of the controlled protocol.	• Only 50 min of observations per subject • The observer suggested some types of activities in order to obtain a wide range of postures and motions.

Table 3.2 - Summary of the strengths and limitations of the state-of-the art free-living protocols – continued.

	Study	Strengths	Limitations
Video and image captures	**(Bourke et al., 2017)**	• Validation of a daily life activities, transitions and postures. • Two types of protocol: Supervised task-based and free-living unsupervised task-based protocols. • Using high frame-rate video technology of ≥25 fps. • Selection of different wearable inertial sensors (Shimmer, SenseWear, ActiGraph, and ActivPAL). • Five raters labelled the videos.	• The concerned population: 20 older adults over 65 years of age. • Aims particularly at developing fall-risk assessment algorithms. • Exclusion of activities related to sport and confounding activities. • Interested in general postures and behaviors only. • The out-of-lab activities were recorded on less than 2 hours per subject.
	(Rosenberg et al., 2016)	• SenseCam Camera: ground truth annotation. • Recordings during waking hours on 7 days for 39 women: 3000 hours of data. • Development of new algorithms from free-living settings and validation on new sample.	• The concerned population: 39 older women (55-96 years old) • Only activities of posture and locomotion: sitting, riding in a vehicle, standing still, standing moving, walking/running. • One accelerometer (ActiGraph) on the hip. • Gold standard observational data only recorded for walking activity.
	(Carlson et al., 2014)	• Use of GPS receiver. • SenseCam Camera used as ground truth annotation • 3 research assistants were trained to code and annotate images. • Validation of GPS algorithms using annotated image. • 3–5 day monitoring period per subject. • Population age range: 18 – 70 years.	• Only considered GPS parameters: speed and distance + PALMS (Personal activity location measurement system). • Classify the trip mode (Walking/running, cycling, in vehicle) and the time spent in the trip. • Threshold based algorithms: cut-point on the 90th percentile speed value.
	(Kerr et al., 2013)	• SenseCam Camera used as ground truth annotation. • A standardized coding protocol based on existing behavioral taxonomies (e.g., SOFIT) and principles of nominal group technique. • Inter-rater reliability of image coding was established. • 3 –5 days monitoring period per subject. • Population age range: 18 – 70 years.	• Only validating the detection of sedentary behaviors. • Considering simple method of 100 (counts per minute) cpm accelerometer cutpoint. • One ActiGraph on the hip (working on counts data).
Sensors	**(Skotte et al., 2012)**	• Validation tool: pressure sensor in the hip pocket to detect sitting. • 9-hour measurement period in free-living context.	• Only validating sitting posture in free-living conditions. • The subjects filled a diary, specifying the start and end of working hours, travel time and time spent lying.

3.3 Basic data collection protocol

The dataset collected using this first protocol served for the development of the classification methodology. At first sight, we established a feature extraction method on the basis of this dataset and proved its pertinence in terms of recognition performance compared to existing methods in literature (as it will be presented in the next chapter). However, to further validate the proposed method, an extended dataset was needed. To fulfill these demands, we designed a new data acquisition protocol: the triple-session pragmatic protocol that includes enlarged number of activities and subjects and will be presented afterwards.

3.3.1 Materials

The dataset was collected using the Shimmer3 IMU and ECG modules manufactured by SHIMMER Company (Shimmer Research, Dublin, Ireland). These devices can accurately record the acceleration in three directions in the dynamic range of ± 16g at different sampling frequencies. A more detailed description of the Shimmer devices can be found next. Three IMU modules were attached to the wrist, the hip and the ankle of the participant and one ECG module was attached to his chest as shown in Figure 3.1. The modules were wrapped and secured to the body with straps. Thus, the participant could execute the vigorous activities without any limitations on his movement or any risk to harm the electronics. The IMU modules were calibrated to store acceleration data at 90 Hz and the ECG module was calibrated to record ECG signals at 512 Hz through 6 lead electrodes placed on the appropriate positions on the chest. In this study, only the HR is extracted from ECG signals to contribute in the classification algorithms.

3.3.2 Experimental procedure

The dataset includes five sedentary and ambulatory activities performed in a random order: running, walking, cycling, car riding (as a passenger) and resting (Figure 3.2). Subjects could freely choose the speed of their movement while the duration of each activity was three minutes at least. An observer accompanied the subjects during the experiments to annotate the start and the end of each activity.

3.3.3 Participants

Eight healthy volunteers (4 male and 4 female) participated in this study. They were recruited from the University of Rennes 1 and from the Ecole Normale Supérieure (ENS) - Rennes, and were aged between 18 and 30 years (Table 3.3).

Table 3.3 - Physical Characteristics of the participants in the basic protocol.

Characteristics	Mean (SD)
Age (*year*)	26.3 (4.7)
Height (*cm*)	172.1 (12)
Weight (*kg*)	65.9 (12.1)



Figure 3.1 - The Shimmer sensor nodes and the ECG electrodes positions on the body in the basic protocol.

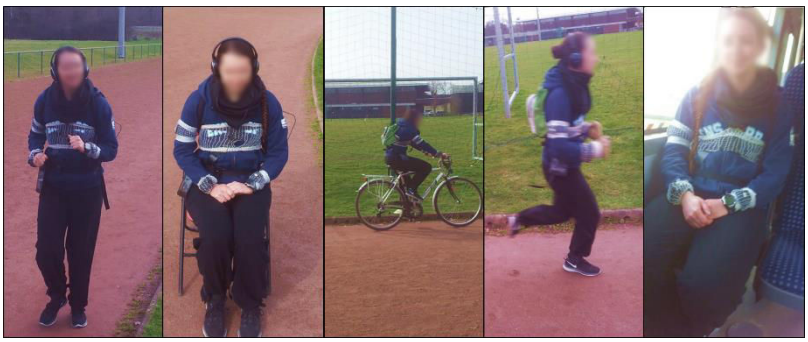


Figure 3.2 - The performed activities in the basic protocol, respectively from left to right: walking, resting, cycling, running and car riding.

3.4 Pragmatic data collection protocol

3.4.1 Materials

One of the objectives of the pragmatic protocol is to select the most appropriate platform of sensors/and parameters among the widely used and commercially available systems. The main challenge is that the sensing platform must be able to accurately recognize the variety of daily life activities in free-living context. Aside from the good precision, the design aim of the monitoring platform is to also respect a multitude of constraints: miniaturization, easy usage, confidentiality, accessibility and affordability (Khusainov et al., 2013). In the data acquisition process of the pragmatic protocol, eight off-the-shelf commercially available systems placed on different parts of the body were considered: ActiGraph

GT3X+, Shimmer 3 from Shimmer Company, SenseWear Armband from BodyMedia, Zephyr BioHarness 3, Cosmed K4b2, Qstarz BT-Q1000eX, Autographer Camera and ActivPAL from PAL technologies. Figure 3.3 and 3.4 show the different devices, their placement on the body and the sessions in which they were used during the pragmatic protocol. The Cosmed K4b2 and the Autographer camera were used as gold standard devices for validation and the other platforms were tested and compared to each other. We have chosen to particularly investigate these sensors due to their extensive usage and the pivotal role they played in the field of activity assessment. All six platforms are particularly conceived for wearable sensing applications and are broadly used in human health monitoring researches. Further, they all feature the desired specifications from miniaturization and ease of usage to long monitoring lifetime. Table 3.4 presents a detailed description of the chosen monitoring devices: their dimensions, weight, location on the body and the principal parameters recorded. Table 3.5 shows the configuration mode considered for each device, the sampling frequency with the related battery lifetime and memory capacity. In order to achieve a complete synchronization of all the devices prior each experiment, the internal clock of the devices was synchronized to the clock of the PC which was initially synchronized to UTC time as given on the website: https://www.timeanddate.com/worldclock. The last column in Table 3.5 displays how each device could've been synchronized.

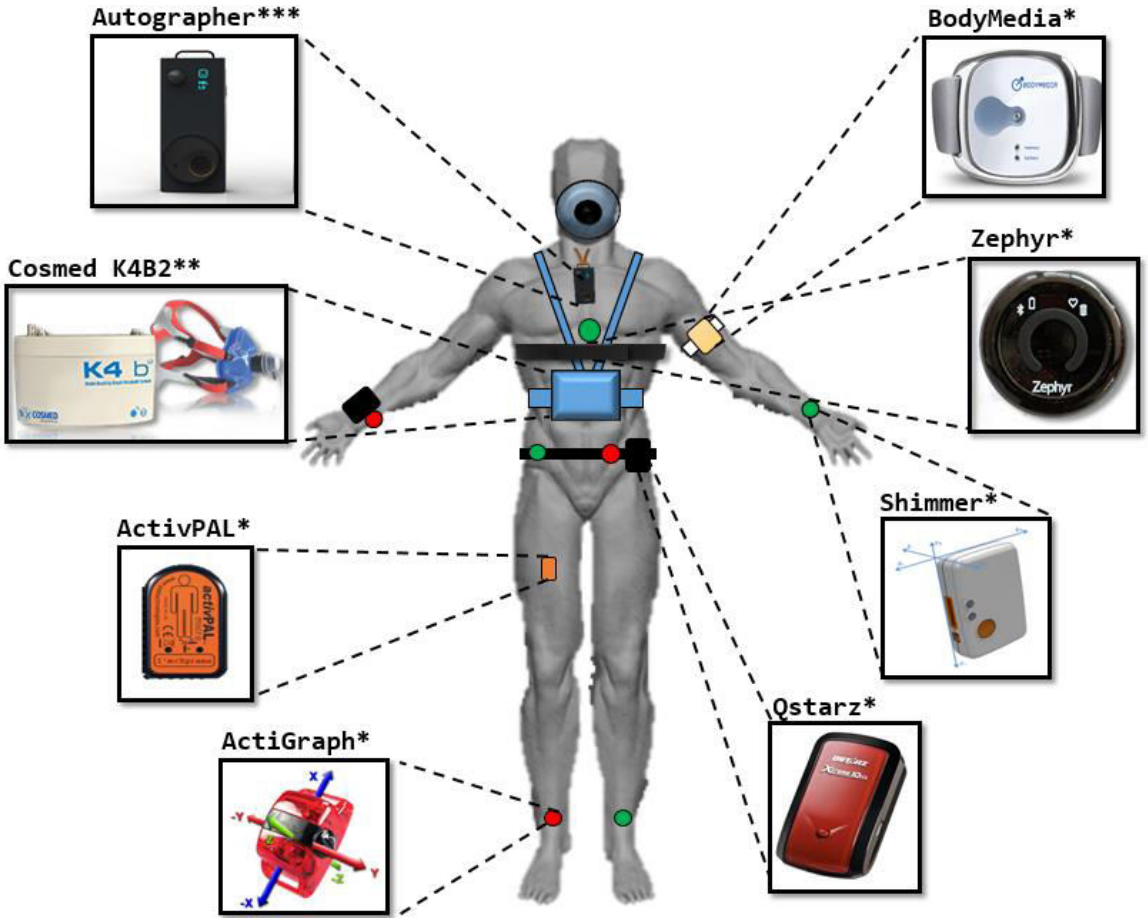


Figure 3.3 - An illustration of the equipment and their positions on the body, where the * indicates the systems used in all sessions. The ** indicates the systems used in S1 and S2 and * indicates the systems used in only S3.**



Figure 3.4 - A participant wearing the sensors and gas exchange system. Some sensors must be placed in direct contact with the skin as seen in the picture on the left while others are worn on top of the clothes as seen in the picture on the right.

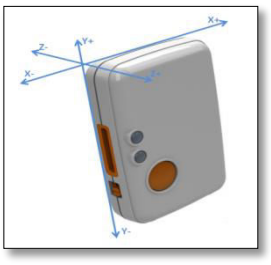
Table 3.4 - The physical characteristics and the principal parameters measured by the considered systems in the pragmatic protocol.

Module	Dimensions (Lxlxh) mm	Weight (grams)	Location on the body	Principal parameters
Zephyr BioHarness 3	28(D)x7 (h)	18	Chest strap	- 3D acceleration - Heart Rate - Breathing Rate
Shimmer 3	51x34x14	31	Ankle, hip, wrist Chest (ECG module)	- 3D acceleration - 3D gyroscope - 3D magnetometer - ECG signal
ActiGraph wGT3X+	46x33x15	19	Ankle, hip, wrist	- 3D acceleration - Lux Measurement
Qstarz BT-Q1000eX	72x47x20	65	Wrist, hip	- Distance - Speed - Location
BodyMedia SenseWear™	55x62x13	45	Arm	- Skin temperature - Acceleration - Energy expenditure
ActivPAL	53x35x7	20	Thigh	- Acceleration - Inclination - Posture
Autographer Camera	90x37,4x22,9	58	Around the neck	- Automatic photo captures

Table 3.5 - The technical specifications of the considered systems in this study.

Module	Configuration	Battery Lifetime	Memory Capacity	Synchronization
Zephyr BioHarness 3	Summary and Waveform mode. F_{ECG} = 250 Hz F_{HR} = 1 Hz F_{Acc} = 100 Hz	More than 11hours	55 hours	Manually via the software.
Shimmer 3	F_{ECG} = 250 Hz F_{IMU} = 100 Hz	23hours	Around 25 hours (ECG module)	Synchronized automatically to PC clock once configured.
ActiGraph wGT3X+	Raw mode. F = 100 Hz Wireless: OFF	11 hours	12.5 days	Synchronized automatically to PC clock once configured.
Qstarz BT-Q1000eX	F= 1 Hz Vibration detector: OFF	24 hours	13 hours 20 mins	Internal clock automatically synchronized to UTC.
BodyMedia SenseWear™	Researcher Mode. F = 1 Hz	5 to 7 days	More than 13hours	Synchronized automatically to PC clock once configured.
ActivPAL	Raw mode. F = 20 Hz.	More than 7 days	About 10 days	Synchronized automatically to PC clock once configured.
Autographer Camera	High Capture: Max: 6 photos/min. Avg: 3 photos/min. GPS Off, Bluetooth OFF and sound ON.	8 hours	27 000 photos	Manually via the software.

3.4.1.1 Shimmer3 IMU and ECG modules



The Shimmer sensor nodes are miniature unobtrusive wearable modules. In this study, we proposed to use this inertial sensing platform since it integrates not only an accelerometer but also allows other physical measurements. Each Shimmer IMU node contains a low noise 3-axis accelerometer, a wide range 3-axis accelerometer, a 3-axis gyroscope, a 3-axis magnetometer, a temperature sensor and a relative pressure sensor. The Shimmer Company has also developed wearable ExG nodes that can be configured for measuring EMG (Electromyography) and ECG (Electrocardiograph) signals. A considerable amount of literature that made use of the Shimmer sensing platform has been lately published. (Cleland et al., 2013) considered six Shimmer nodes and analyzed the optimal placement and the needed number of accelerometers to sufficiently

detect everyday activities. The Shimmer IMUs have been also adopted in gait and fall risk assessment (Doheny et al., 2013) and in Parkinson disease (PD) applications. For instance, (Jalloul et al., 2015) deployed a monitoring system consisting of a set of 6 Shimmer IMU nodes for the detection of dyskinesia in PD patients. Furthermore, the ECG data produced by the Shimmer ECG module have been validated in several studies. As an example, (Gradl et al., 2012) developed a mobile application that is able to monitor real-time ECG and automatically detect arrhythmia by analyzing the Shimmer ECG signals. The Shimmer nodes were also successfully used for estimating human EE (Gjoreski et al., 2015). Interestingly, (Leutheuser et al., 2013) produced in 2013 a benchmark dataset of Daily Life Activities (DLA) 'The DaLiAc database' collected using Shimmer sensing platform and made it publicly available on their website to be implemented in new classification algorithms. In our study, we used one ECG node placed on the chest and three IMU nodes placed on the wrist, hip and ankle. The nodes were attached using elastic Velcro straps and five electrodes connected externally to the ECG module are adhered on the chest as shown in Figure 3.4. The Shimmer platform is also accompanied by the 'Consensys' software interface for configuring, synchronizing and data managing the Shimmer nodes. Although the system allows a Bluetooth communication for real-time data transmission, we only logged the recorded data on the Shimmer SD card for offline processing.

3.4.1.2 Zephyr BioHarness3 Chest Strap



The BioHarness 3 is a newly available multi-sensor monitoring device strapped around the chest touching the skin in a direct contact. This device features three physical as well as physiological sensors housed within the chest strap. An electrode sensor captures ECG waveform and HR data. A capacitive pressure sensor provides the breathing waveform (BW) and the breathing rate (BR) data by detecting the expansion and contraction movements of the ribcage. A MEMS triaxial accelerometer delivers the acceleration data in gravitational force unit (g) along 3 orthogonal axes (vertical (x), sagittal (z) and lateral (y)). The applicability of the BioHarness chest strap in the field of activity recognition was mainly done in combination with a mobile phone. For instance, (Lara et al., 2012a) designed a mobile platform 'Vigilante' that recognizes PA in real-time after receiving data from the BioHarness sensors via Bluetooth. Furthermore, (Gjoreski et al., 2015) extracted the HR, the BR and the skin temperature from the Zephyr sensor and used these vital signals among others to develop an algorithm for estimating EE. The validity and the reliability of the Zephyr BioHarness variables were successfully assessed by (Johnstone et al., 2012a; Johnstone et al., 2012b). However, the validation tests of both studies were conducted in a controlled laboratory context. To the best of our knowledge, our current study would be the first to analyze the reliability of the Zephyr BioHarness in free-living situations outside the laboratory.

3.4.1.3 ActiGraph GT3x+



Contrary to the aforementioned devices, the ActiGraph GT3X+ device only incorporates a triaxial accelerometer and a light sensor. However, the ActiGraph accelerometer is to date the most widely studied activity monitor in the research area of activity recognition. A review of great interest conducted in March 2014 (Wijndaele et al., 2015) that pooled the accelerometry data collected internationally in adults' population has found out that more than half of the studies (51% across 76 studies) have used an ActiGraph device. In PAR field, the ActiGraph accelerometer had been successfully validated on big sample data. As an example, data were collected from a sample of 222 older women in the study of (Rosenberg et al., 2016) and from a sample of 146 adults in the study of (Kate et al., 2016). The ActiGraph was also validated in real-life context (Hu et al., 2016; Rosenberg et al., 2016; Sasaki et al., 2016) and on relatively long-term monitoring: on 7 days during waking hours in the study by (Hu et al., 2016) for example. Though, traditional analysis of the ActiGraph accelerometer data was based on what is known as 'counts measurements' (AC). The limitations in using such type of data is explained in Chapter 2 (section 2.5.3).

3.4.1.4 Qstarz BT-Q1000eX



The Qstarz BT-Q1000eX is a GPS data logger device that records altitude, longitude, height speed, distance, the number of satellites in view, the number of satellites used and satellites signal information (e.g. the signal to nose Ratio (SNR)). Similarly to other GPS devices, the Qstarz device is dedicated for outdoor use where satellite acquisition is faster and provides better positioning accuracy. The GPS is mainly used in complement with accelerometers to identify the activity's context and transportation modes in ambulatory applications. The use of GPS devices in the field of activity recognition is reviewed in Chapter 2 (section 2.2.2.4). Particularly, the Qstarz brand of GPS is well-documented and is used by a number of PAR researchers (Carlson et al., 2014; Ellis et al., 2014; Hu et al., 2016; Rosenberg et al., 2016; Wu et al., 2011). (Rosenberg et al., 2016) recorded time-stamped location data with the Qstarz in free-living daily life context. The aim of the study was to identify walking and sedentary time in older women over 6 days of wearing both accelerometers and Qstarz sensor. Furthermore, (Ellis et al., 2014) made use of the Qstarz BT1000X to identify active travel behaviors under different environment conditions (e.g. urban canyon with tall buildings that interfere with GPS signal and open space where GPS connectivity is better). The Qstarz device allowed collection of satellite information and helped detecting outdoor from indoor locations.

3.4.1.5 SenseWear BodyMedia Armband



The BodyMedia armband is a multi-sensor device designed to be worn at the left upper arm over the triceps. The integrated sensors deliver triaxial acceleration measures and a variety of physiological variables: the heat flux, the skin temperature, the near-body ambient temperature, the galvanic skin response. This multi-sensor armband was conceived as an EE assessment tool. The manufacturer's software permits to study the type and the intensity of the performed activity and to estimate the EE on the basis of the measured variable, together with the user's characteristics (gender, age, body weight, height, smoking status and handedness). The BodyMedia sensor is a well-known state-of-art commercial sensor to predict daily EE in free-living situation (Bonomi et al., 2012; Gjoreski et al., 2015). A great deal of previous research focused on evaluating the validity and reliability of the armband's EE estimation method against the doubly labeled water approach and the indirect calorimetry measurements (Fruin et al., 2004; St-Onge et al., 2007). However, the integrated-EE algorithms are proprietary algorithms that remain unpublished and it is unclear how the different variables contribute in the built-in prediction equations (Bonomi et al., 2012; Chen et al., 2005).

3.4.1.6 ActivPAL



The ActivPAL device is designed to be attached to the thigh. It integrates an accelerometer that detects the thigh movement and inclination. From this information, the ActivPAL sensor is conceived to (i) discriminate between postures and activity types, (ii) calculate the time spent in each posture or activity and (iii) estimate the related EE on the basis of default MET values and default regression equations. Although the ActivPAL can classify postures and activities into only three categories: standing, stepping and sitting/lying (Grant et al., 2006), it is considered the field ground truth device for posture measures (Rosenberg et al., 2016). Several researchers interested in identifying postures and sedentary activities compared the performance of their algorithms to the outputs of the ActivPAL, the reference device (Skotte et al., 2012). However, it was reported that the EE calculation by the ActivPAL software is not accurate enough. The integrated thresholds and equations underestimate the measure of EE in walking, running and cycling activities for example (Schneller et al., 2015).

3.4.1.7 Autographer Camera



The Autographer is a wearable light-weight digital camera that can automatically capture photographic images from the user's perspective via its 136° eye-view lens. The camera clips on to clothing or can be worn around the neck with a lanyard. It features five integrated sensors: an accelerometer, a magnetometer, a light, a motion and a temperature sensors. Once these sensors are triggered, meaning once they detect a change in movement, light or temperature, etc., an image is captured. Otherwise, an image is captured each 20 seconds if the sensors aren't activated. In high frequency mode, the number of images taken typically ranges between 3 and 6 images per minute. The images are time-stamped with a 1-sec resolution and can be downloaded and labelled via the Autographer software. As discussed previously, these types of cameras currently present real proxy for human observation in the research field of activity recognition (Kelly et al., 2013). They allow to track the movement of the wearer and identify the activities he was performing during his day due to the images captured. Although some existing studies have used the camera as an activity recognition tool along with the accelerometer (Doherty et al., 2013), the Autographer camera serves here as a validation tool only. Other brands of cameras manipulated for the same purpose exist in literature, mostly SenseCam (Carlson et al., 2014; Doherty et al., 2013; Kerr et al., 2013; Rosenberg et al., 2016) and GoPro camera (Bourke et al., 2017). (Bourke et al., 2017) have monitored the participants using high resolution video technology from the GoPro Hero3+ camera. However, the battery lifetime and the memory capacity of the GoPro video recording system limited the data collection period to just 2-hours. We believe that 2 hours period is not long enough to track the subject's behavior in daily-life context. Nevertheless, in our study, the Autographer camera took continuous images with much lower resolution (maximum 6 photos/min) than the video recordings but the data collection in this case lasted sufficiently for almost 8 hours.

3.4.1.8 Cosmed K4b2



The Cosmed K4b2 machine is a portable indirect calorimetry system that allows the measurement of breath-by-breath pulmonary gas exchange. The system also allows the direct field study of human performance and the detection of cardiopulmonary limitations during exercise (Ainslie et al., 2003). The Cosmed K4b2 has been proved to be a valid and a reliable metabolic measurement system when tested against the Douglas bag method across a wide range of cycle ergometer intensities (McLaughlin et al., 2001). The Cosmed system is generally used as a ground truth device for the EE estimation. It serves as a reference for developing and learning EE prediction equations and then for evaluating the performance of the developed EE models (Crouter et al., 2006a; Dugas et al., 2005; Gjoreski et al., 2015; Kate et al., 2016; Schneller et al., 2015; Swartz et al., 2000). The fully portable

Cosmed system consists of a facemask, a harness, a mouthpiece, a battery and a gas exchange analyzer. Prior to each test, the oxygen and carbon dioxide analyzers of the Cosmed system must be calibrated. Four steps calibration process is performed following the manufacturer's guidelines: 1. Room Air calibration relative to ambient air. 2. Reference air calibration against a specific percentage of gas mixture. 3. Turbine calibration using a 3L syringe and 4. Gas delay calibration to determine the lag time between the turbine measurement and the gas analysis. Information on the participant characteristics (age, gender, height and weight) and on ambient temperature and humidity are then entered to the Cosmed K4b2 device before initiating a test. During the test, the system outputs $\dot{V}O_2$, $\dot{V}CO_2$ and $\dot{V}_E$ values measured breath-by-breath. The Cosmed also calculates EE function of $\dot{V}O_2$ (EE = $\dot{V}O_2$ x 4.825). The calculated EE is then compared to the estimated EE issued from prediction models.

3.4.2 **Experimental procedure**

In order to achieve our objectives in first evaluating our developed features extraction method on an extended dataset and second on studying the effects of the type of datasets on the recognition performance, we designed a triple-sessions based protocol. The standardized (S1) and semi-standardized (S2) sessions were performed under supervision, while participants in the third free-living session (S3) are left to freely perform their daily-life activities with the no supervisor around. Yet an Autographer camera was worn around the neck to automatically take photos of their environment (Figure 3.5).

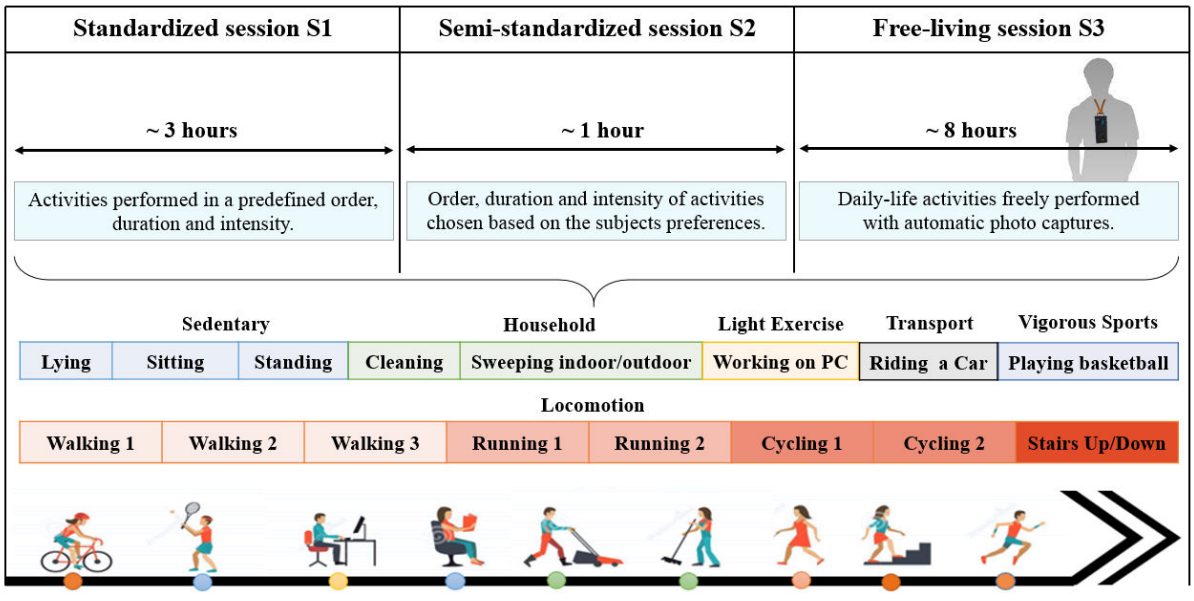


Figure 3.5 - Description of the triple-session protocol for data collection.

3.4.2.1 Standardized session (S1)

Based on the dataset collected in this session, we will be able to develop mathematical models for the recognition of daily life activities and the estimation of the associated EE. It was required that each

subject performs a series of sedentary and dynamic activities in standardized conditions: in a predefined duration, intensity and order. The subjects were equipped with a number of sensors deployed on different parts of the body in addition to the gas exchange measurement system, which were presented previously. The activities in this session are performed under the supervision of an observer and are carried out indoor or outdoor depending on their nature. Each activity lasted at least 6 min and activities were separated by transition periods. These periods consist of the recovery time needed after performing a vigorous activity. The initial values of resting $\dot{V}O_2$ must be retrieved as much as possible (± 1mL·O_2/min/kg) before starting next activity. The ranges of activities to be performed were chosen to reflect the diversity of physical and sedentary activities that can be carried out on a daily basis by any individual (Figure 3.6). The list of activities is: (i) Sedentary activities: lying, sitting, standing. (ii) Light activity: working on PC. (iii) Locomotion activities: walking at 3 different speeds (slow, normal, fast), running at 2 different speeds (slow and fast), cycling at 2 different speeds (slow and fast), and going up/down stairs at self-selected pace. (iv) Transportation: riding in a car as a passenger and the experimenter is the driver. (v) Sports: playing basketball, meaning doing repetitive acts related to basketball; passing the ball, shooting (jump shots or layups) and dribbling, and (vi) Household activities: cleaning the windows, sweeping indoor and sweeping outdoor.



Figure 3.6 - Examples of the performed activities during Session S1.

3.4.2.2 Semi-standardized session (S2)

The session S2 aims at collecting a testing dataset that serves for validating the robustness of the mathematical models developed on the dataset of S1. The activities in this case were performed in semi-standardized condition, which is considered to be closer to the conditions of daily-living than S1. The

session S2 also allows to refine the recognition model by adding to it augmented dataset. In this session that lasts around 50 min, each subject is required to carry out a sequence of activities from the same list of sedentary and physical activities as in S1. The subject now can freely choose the order, the duration and the intensity of the activities, but with two restrictions: (i) the duration of each performed activity type must last at least 20 s to reduce transition periods between the activities, which would not be taken into consideration in this study and (ii) each activity type must be performed at least once to ensure fair contribution from all the considered activities. During this session, the participants are equipped with the same sensors and gas exchange system that are mounted on the same locations on the body as in S1. An experimenter accompanies the subjects during this session as well and annotates the sequence of all the performed activities and marks their starts and ends on a worksheet. Moreover, this session was filmed and the recorded videos were afterwards analyzed to verify if the activities were correctly allocated.

3.4.2.3 Free-living session (S3)

The session S3 is the fundamental session in the acquisition protocol. It permits to evaluate the recognition models of sessions S1 and S2 on situations of daily-living. During this session, the participants were asked to carry on their normal day. Throughout the monitoring day (7 to 8 hours), they were equipped by a number of devices, as presented in Figure 3.3. Besides, the Autographer camera, which was not used in S1 and S2, will serve here to capture automatic photos from the user's environment on the monitoring period. In agreement with similar studies (Carlson et al., 2014; Doherty et al., 2013; Kelly et al., 2013; Kerr et al., 2013), ethical concerns were considered for the sake of protecting the privacy and security of the photos. Participants were instructed on how to pause the automatic photo capture whenever they wish. At the end of the session, they were also able to review the photos and delete the undesired ones. Participants were also instructed to remove the sensors if they have to shower or during any activity where there is a direct risk to damage the equipment or to get hurt because of the equipment (e.g. contact sports). An example of the photos captured during this session is presented in Figure 3.7.

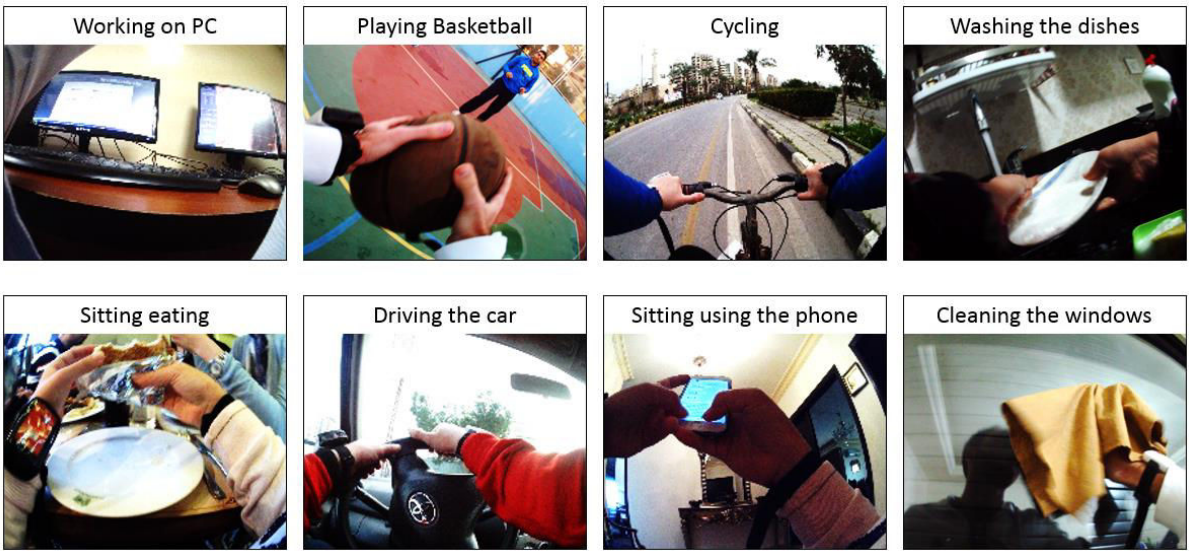


Figure 3.7 - Example of photos captured by the Autographer Camera with the corresponding annotations.

3.4.3 **Participants**

A convenience sample of 20 subjects who were between 18 and 35 years old volunteered to participate in this study which was conducted in Lebanon. The study was approved by the Ethics committee of the Doctoral School of Sciences and Technology of the Lebanese University (Ref. Number: CE-EDST-1-2016). After a detailed explanation about the protocol, the three sessions and their objectives, participants provided written informed consent. The participant is excluded from the study if he/she is: (i) subject to medical treatment that may affect his EE at rest and/or during an exercise or (ii) suffering from pulmonary or cardiac diseases or (iii) pregnant. The participants' characteristics are provided in Table 3.6.

Table 3.6 - Physical characteristics of the participants in the pragmatic protocol.

	Age (*years*)	Height (*cm*)	Weight (*kg*)	Gender (*Female/Male*)
Mean (SD)	21.5 (4.4)	168.5 (8.8)	62.5 (12.6)	10/10

3.5 **Conclusion**

In this chapter, we described the design of two protocols that were used in our studies: a basic and a pragmatic protocols. Basic protocol included five sedentary and ambulatory activities performed during a fixed duration under supervision, and the dataset was collected from 8 subjects. The pragmatic protocol, however, consisted of an expanded set of activities in order to cover various daily postures and activities from locomotion to household and sports. The dataset in this case was collected from 20

subjects who performed the protocol on three different sessions (standardized (S1), semi-standardized (S2) and free-living (S3)). An observer accompanied the subjects during the first and second sessions. Whereas in the third session, the subjects were left to continue their usual activities on the course of the day. A wearable camera that was capturing automatic photos from their environment, served as gold standard for activity annotations. The basic protocol was initially used for developing and optimizing the classification algorithms. But its small sample size did not allow for a complete validation. We were interested in testing the developed method using the dataset from the pragmatic protocol as well. This work is presented in the following chapter. An important issue that wasn't addressed in most of the existing studies in the literature is to test the accuracy of the recognition models in a context of free-living. The three sessions in the pragmatic protocol could serve this matter. By training the models using dataset from session S1 and then testing their performance in recognizing data from session S3, this can introduce interesting findings regarding the impact of free-living data on the system response. This aspect is closely investigated in Chapter 5.

Chapter 4
METHODOLOGY AND PRELIMINARLY RESULTS

Note to Readers: *Part of the study presented in this chapter was published in the international journal 'Biomedical Signal Processing and Control' and in the '2015 International Conference on Advances in Biomedical Engineering'. The work was also presented as an oral communication in the '21st annual congress of the European College of Sport Science'.*

Automatically identifying daily life activities has become a focus in research in the biomedical domain. Although in the literature, there exist many algorithms to accurately classify activities based on acceleration data; few are dedicated to the efficiency in terms of the relevance of the selected features. In fact, fetching and incorporating a large number of features in the classification process lead to the following issues: (i) the redundancy or irrelevance of some features that may not add significant information to the classification performance, and (ii) the complexity and the additional computational time required for calculating the features and training the model based on the big features set (Altun et al., 2010; Chen et al., 2008; Leutheuser et al., 2013; Parkka et al., 2006; Ravi et al., 2005). Furthermore, few researchers were motivated to study physiological, environmental and location signals in addition to the acceleration signals, in the aim of improving the recognition accuracy and identification precision (Freedson et al., 2000; Lara et al., 2012b; Tapia et al., 2007a).

In this chapter, we propose a formal methodology to improve the activity recognition and assess the impact of integrating other physiological signals, the HR as in (Lara et al., 2012b; Tapia et al., 2007b). The study presented here proposes an original feature selection approach based on the spectral distances between a given signal and an activity model. The new spectral distances method (SpD method) is evaluated and compared to existing methods in literature. This study also investigates the improvement of classification performances brought by the HR data in addition to the accelerometer data. The experimental dataset used in this study is the one collected during the basic protocol. It consists of acceleration and HR recordings from eight volunteers performing five sedentary and ambulation activities (running, walking, cycling, car riding and resting). Four wearable sensor modules, including an ECG node are employed. For more details on the experimental procedure, the measurement tools and the physical characteristics of the participants in this protocol please refer to the previous chapter. The response of the system to three widely used classifiers, the K-nearest neighbors (K-NN), the Naïve Bayes (NB) and the decision Tree (DT-C4.5) is reported along with the classification rates. A brief description of the three classifiers can be found in Chapter 2 (section 2.3.5.1).

This chapter is organized as follows: section 4.1 explains the interest of the distance measure using spectral density. Section 4.2 gives methodological details on the AR spectral model training and presents the reference feature selection method in the literature. Section 4.3 presents the inclusion of HR for activity classification. Then, section 4.4 compares the proposed method with the reference feature selection method in the literature (Altun et al., 2010) and outlines the most important contributions. Finally, section 4.5 summarizes the presented study.

4.1 Interest of the Spectral Density distance measures

4.1.1 Limitation of classical approaches

It is difficult to recognize the pattern for each PA using the raw acceleration signals due to the fact that they are by nature noisy and containing repetitive variations. In general, typical procedures of a PAR system start by extracting basic statistical features from the signals in the time and frequency domain, then reduce these feature dimensions in order to choose the most relevant features to discriminate PA, and finally recognize the PA pattern using a classification tool (Altun et al., 2010; Parkka et al., 2006). However, in these methods, features are mostly designed arbitrarily to contain first, second and third order statistics truncated to a certain dimension in the first place, some may hold meaningless or loose important information with respect to the PA recognition task. Furthermore the selection results vary from one training database to another, which in turn proves that the selected features without an appropriate model might not be representative of the PA of interest. For example, in (Altun et al., 2010) time- and frequency- domain features were computed from 1st, 2nd and 3rd order statistics, before reducing their number to a vector composed of 30 features. In order to prove our hypothesis, we

reproduced the work of (Altun et al., 2010) in our dataset and sorted the features based on their pertinence in discriminating the activities.

Table 4.1 represents the features in a descending order according to their mutual information MI calculated on the basis of the equation:

$$MI(\boldsymbol{X},\boldsymbol{Y}) = \sum_{y\epsilon Y}\sum_{x\epsilon X} p(x,y)\log\frac{p(x,y)}{p(x)p(y)} \quad (1)$$

Where $p(x,y)$ is the joint probability density function of $\boldsymbol{X}$, the PA class, and $\boldsymbol{Y}$, the feature, and $p(x)$ and $p(y)$ are the marginal probability density functions of X and Y respectively.

As it can be noticed in the ranking, the 3rd order statistics (skewness and kurtosis) firstly appear at row 140. Furthermore, the rest of the group of skewness and kurtosis features are ranked in the end of the list. Since only the first 30 features are finally selected for the activity classification, the 3rd order statistics are thus completely discarded. On the other hand, in this type of feature selection, the first five selected features refer to the signal variability and its energy, yet other feature selection methods may not classify the features in this same order. This brief study prove then how unstable is the choice of the most significant features. That's why one of the interests of our proposed method is to avoid this critical step.

Table 4.1 - The order of the information gain of sorted features. (MI: Information Gain).

Order	MI	Feature	Order	MI	Feature	Order	MI	Feature
1	1.945	varYCh	139	0.855	corr2XHi	307	0.175	skewZAn
2	1.904	peak1YHi	140	0.852	kurtYCh	308	0.167	skewZWr
3	1.885	varXWr	141	0.849	skewYWr	309	0.167	kurtXCh
4	1.875	Peak4YAn	142	0.847	Indf3YPg	310	0.142	skewXHi
5	1.874	VarYHi	143	0.846	Corr5ZHi	311	0.121	skewZHi
:	:	:	:	:	:	312	0.117	skewXCh

Symbols Definition: *X, Y* and *Z* refer to the X-, Y- and Z- accelerometer axes, *Ch, Hi, An* and *Wr* to the Chest, Hip, Ankle and Wrist positions respectively, *var* to the variance, *peak1* and *peak4* to the 1st and 4th peaks in the spectrum values respectively, Indf3 to the 3rd frequency index value, *corr2* and *corr5* to the 2nd and 5th correlation coefficients respectively, *kurt* to the kurtosis and *skew* to the skewness.

4.1.2 New modeling approach

Contrary to the state-of-the-art methods implying procedures to extract feature vectors for all acceleration signals (mean, variance, skewness, kurtosis...) and then reduce the dimensions, our study suggests an adapted feature set based on the AR model spectral density to measure distances of different activity types. As the Fourier transform of the auto-correlation function, the spectral density includes all statistical information up to the second order without truncation. The distance measure applied on the

spectral density yield low-dimensional vectors, and thus no more feature selection is needed here, resulting in a substantial gain in computation complexities for both learning and real-time applications. In this study, we developed a spectral recognition model using measurements of spectral distances and then in order to study the capability of the developed model in discriminating activities, we compared its accuracy with that of the state-of-the-art methods (Altun et al., 2010) constructed using basic statistical features (SF method). However, unlike (Altun et al., 2010), we calculated the classification accuracies using the whole set of features without reducing their dimensions, in an attempt to reach the best reference results (highest accuracies) to compare with the results of our proposed method of spectral distances-based classification. Data streams are splitted into sliding windows of 5 s on which we applied the two feature extraction methods. Since the duration of the each activity is 3 min, an altogether 288 windows (8 subjects x 3min x 60s/5s) are created as valid samples representing the PA in question. In this way, when dealing with 5 different activities, each represented by 288 instances, this can make the classification procedure devoted enough regarding the important number of instances for training and testing the classifiers (Franke et al., 2007).

4.2 Features Extraction methods

4.2.1 Set of spectral distances features

The SpD method relies on the fact that power spectral densities of a random process (activity's signal) include information of the periodicity and the cyclostationarity of the process. Hence, spectral distances can be used to measure the differences between two such processes (Basseville, 1989).

The Itakura-Saito distance dIS is the spectral distance considered in our study. It is based on the power spectral density (PSD) of a random process obtained by the DFT (Discrete Fourier Transform) of its discrete autocorrelation function r(n), n ∈ N. The distance dIS between the PSD of a random process s1 and that of the given model s2 is:

$$d_{IS}(\boldsymbol{s_1}, \boldsymbol{s_2}) \stackrel{\text{def}}{=} \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\frac{\boldsymbol{s_1}}{\boldsymbol{s_2}} - \log\frac{\boldsymbol{s_1}}{\boldsymbol{s_2}} - 1\right] dw \tag{2}$$

Using an auto-regressive model of p-th order AR(p) driven by a White Gaussian Noise of variance σ^2_s (Basseville, 1989), Eq.(2) is shown to be equivalent to:

$$d_{IS}(\boldsymbol{s_1}, \boldsymbol{s_2}) = \frac{1}{\sigma_2^2}\boldsymbol{a_2^T} R_p(\boldsymbol{s_1})\boldsymbol{a_2} - \log\frac{\sigma_1^2}{\sigma_2^2} - 1 \tag{3}$$

where Rp(s) is the square Toeplitz matrix of dimension (p+1)*(p+1) whose (k, j) element is **r** (k - j), 0 ≤ k, j ≤ p, and **a** = $(a_0, a_1, \ldots, a_p)^t$ is the column vector composed of the autoregressive model coefficients. Basically, $d_{IS}(\boldsymbol{s_1}, \boldsymbol{s_2})$ is different than$d_{IS}(\boldsymbol{s_2}, \boldsymbol{s_1})$, so we preferred to implement in our experiments a symmetrized version (dcosh) of the (dIS):

$$d_{\,cosh} \stackrel{\text{def}}{=} \frac{1}{2}\left[d_{IS}(\boldsymbol{s_1}, \boldsymbol{s_2}) + d_{IS}(\boldsymbol{s_2}, \boldsymbol{s_1})\right] \tag{4}$$

such that $d_{\,cosh}(\boldsymbol{s_1}, \boldsymbol{s_2}) = d_{\,cosh}(\boldsymbol{s_2}, \boldsymbol{s_1})$.

From Eq (3) and Eq (4), it is obvious that learning the reference model of each PA type is reduced to estimating ($\boldsymbol{a}$, σ_2^2) while to calculate the distance, the couple ($\mathbf{r}$, σ_1^2) of a raw signal is required. In order to estimate the reference model parameters for each activity, we proposed to avoid the concatenation of the raw signal from the different subjects in the dataset. Instead, we adopted a different process, explained as follows; we begin with the Yule-Walker equation for the i-th subject:

$$\boldsymbol{r}^{\boldsymbol{i}}(p) = T^{i}(p) \cdot \boldsymbol{a} \tag{5}$$

where T^i (of order p) is the upper left square sub-matrix of Rp for the realization i. Then, the computation of the model coefficient $\boldsymbol{a}$ for a particular class of activity can thus be realized by the global system equation, which can be written as:

$$\begin{pmatrix} r^1(p) \\ r^2(p) \\ \vdots \\ r^M(p) \end{pmatrix} = \begin{pmatrix} T^1(p) \\ T^2(p) \\ \vdots \\ T^M(p) \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix}$$

for which the Moore-Penrose inverse gives the optimal coefficients of $\boldsymbol{a}$ in a minimum squared error sense.

The distances $d_{IS}(\boldsymbol{s_1}, \boldsymbol{s_2})$ are finally calculated between the 5s-segments and the activity reference models so that each of the segments will be presented as a set of distances values. The resulting feature vector calculated consists of 60 (= 5 distances x 3 axes x 4 modules) features.

Example of the PSD spectrums constructed for each activity using the AR model coefficients are shown in Figure 4.1. The first peaks clearly detected for the run, walk and cycle activities, reflect the periodicity of the signals at specific frequencies and thus they show that they can usefully identify the activity under test. However, rest and car ride activities are only identified by a high energy in low frequency range.

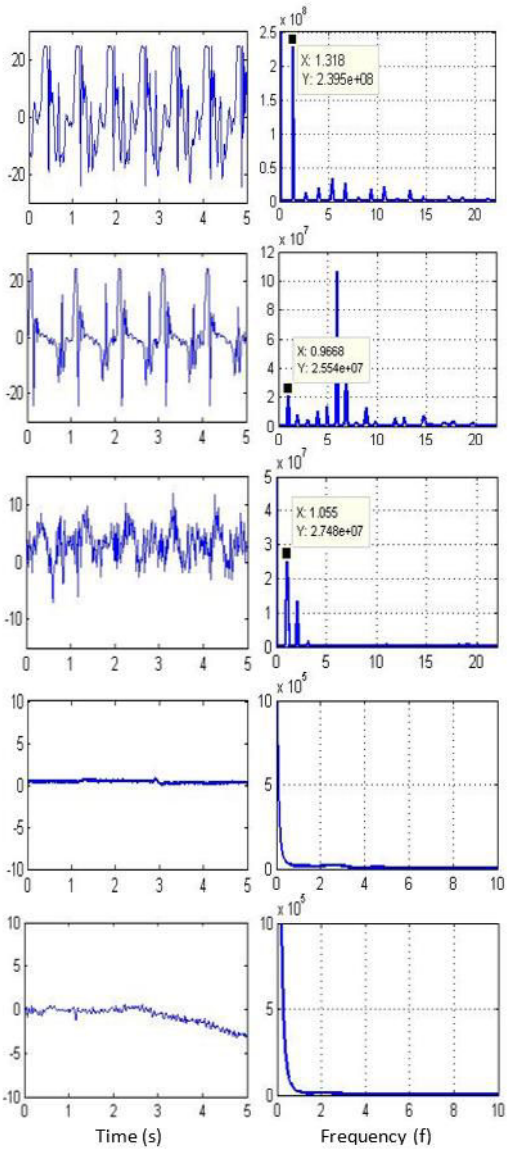


Figure 4.1 - On the left, time-domain signals (m/s^2) from activities collected in a 5-sec window from the x-axis accelerometer of a Shimmer ankle-worn device. On the right, the corresponding spectrums obtained at the output of the AR model for running, walking, cycling, resting and car riding respectively from top to bottom.

4.2.2 Distances features illustration

A representation phase prior to classification is, as we believe, a basic phase to understand the generated data; it consists of a visualization of the activities in terms of the extracted information. This can give a preliminary judgment about the model under construction. In order to visually analyze the data, the measured distances are clustered in color scaled 2D images, whence comes the nomination "distcolorimetrics" (Figure 4.2). Based on the use of power spectral distances, the distcolorimetrics are regarded as features representing each activity's distances towards other activities under study. This illustration of the parameters provides efficient visual representation as to the similarity and the variation among activities. The color of each squared block represents a distance between two respective activities, ranging from red to blue (cf. color bar definitions in Figure 4.2). Naturedly, all diagonal

elements are blue since the dcosh (s1; s1) is zero by definition. We also observed that the x-axis accelerometer of the chest worn device detected a more important difference between walking and running than that of walking and cycling. This colorimetric information allowed a quick evaluation of the difficulties for the activity classification task.

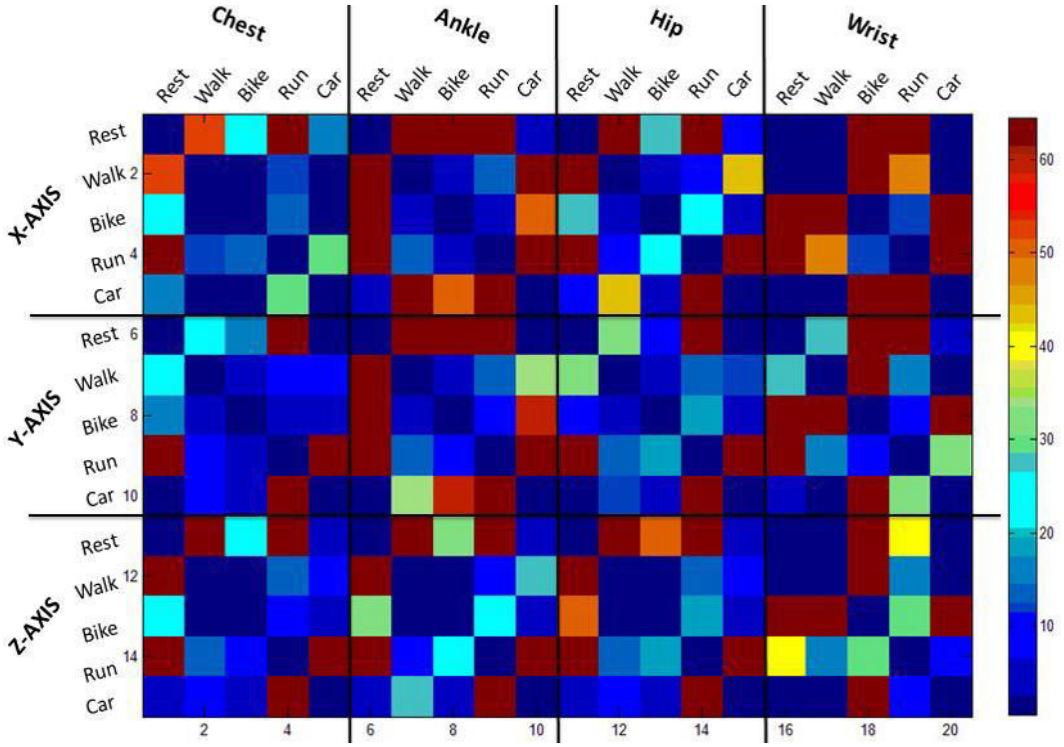


Figure 4.2 - Color scaled 2D image resulting from the distances calculation by pairs of two distinct activities.

4.2.3 Set of state-of-art time- and frequency- domain features

We define a collection of features as chosen by the authors of (Altun et al., 2010). In fact, we have chosen to compare our method to the one presented in (Altun et al., 2010) because its study lead to very high correct classification and it has an important citation rate and research impact. Furthermore, we compared our study specifically to this one because both have the same types of sensing devices (which are based on accelerometers), the same sensor placements (hip, waist, wrist and ankle) and the same type of activities (walking, running, cycling, car riding and resting). The set of features collected by (Altun et al., 2010) includes: mean, minimum, maximum, variance, skewness, kurtosis, 11 autocorrelation samples and five maximum peaks with their corresponding frequency values of the discrete Fourier transform. Hence, for the SF method, the resulting feature vector calculated consists of 324 ((6+11+5+5) x 3 axes x 4 modules) features. We didn't apply the PCA method for features selection as for (Altun et al., 2010); the total feature vector is used for the classifiers training and testing.

4.3 Heart Rate analysis

Few works has been done to measure the impact of physiological signals including HR in PAR applications. Among them, (Lara et al., 2012b) compared the classification accuracy of the two PAR systems either with or without the HR data, and found out that higher overall accuracy is achieved using the HR data. They found that vital signs would definitely yield more reliable classification results since they provide clear patterns to distinguish between activities such as ascending and walking. (Tapia et al., 2007a) also proposed a PAR system combining data from accelerometers and HR monitors. Their study showed that the addition of HR data improves accuracy of recognition by around 2%. They pointed out however that the latency of the HR onset with respect to the transition of PA are the cause of systematic classification errors for vigorous activities of short durations such as walking up stairs.

In our study, we were motivated by the enhancement of the recognition of ambulatory activities by including the HR data. From the recorded ECG signals by the Shimmer electrodes, R waves are first detected using an existing algorithm (Pan et al., 1985), to calculate the instantaneous HR in beats per minute (bpm). Subjects showed notable differences in HR level even with stationary activities. The problem of individual discrepancies is addressed in the training step by estimating the mean and standard deviation of the HR for each subject and for each activity. The difference between the current HR value and its reference value is divided by the standard deviation to form the normalized HR distances. These HR distances are then considered as physiological evidences to the classification task in addition to the physical evidences as calculated by the power spectral density distance measure.

4.4 Evaluation

In this section, we discuss and evaluate the system on the basis of three comparative studies. The first one compares our proposed method with the state-of-the-art method in (Altun et al., 2010). The second one compares the responses of the system regarding the sensors' placements. The third one measures the impact of the HR information inclusion on the system accuracy.

To test the strength of the system regarding the classification algorithms, we have chosen to apply three different types of classifiers: the Naïve Bayes approach (NB), the K-Nearest Neighbors algorithm (KNN) and the Binary decision tree (C4.5) (Rf. Chapter 2 (section 2.3.5.1)). Among the classification algorithms that have been implemented for activity recognition earlier (Khusainov et al., 2013), these classifiers were among the best voted in literature in discriminating activities. The training and testing were done by splitting data randomly: 50% of them were used for training, while the remaining 50% for testing purpose. This process is repeated 20 times, and averaged results were presented for validation.

4.4.1 Comparison between Time-Frequency and Spectral Distance Methods

Using this validation technique and for all the classification types considered, the overall mean accuracies of the spectral distance (SpD) method and the statistical features (SF) method range respectively from 87.8% to 99.0% and from 96.3% to 98.1% (Table 4.2).

Table 4.2 - The overall mean accuracy ± the standard deviation (SD) given by the classifiers: KNN, NB and C4.5 for each of the SF and SpD methods (validation technique: 50% random split).

	SF method	SpD method
KNN (%)	**98.1 ± 0.7**	**99.0 ± 0.4**
NB (%)	96.3 ± 0.6	87.8 ± 1.1
C4.5 (%)	97.2 ± 0.7	95.9 ± 1.1

The 99% ± 0.4 correct classification rate reached by the KNN classifier for the SpD system showed that the computed spectral features have successfully proved their strength in the activity classification. Consequently, these results support the hypothesis that implementing only the spectral distances measures is accurate enough to distinguish between activities; dcosh distance can capture signal variability based on an efficient estimation of the power spectral density of the signal. The information in the skewness and the kurtosis features are missing in the spectral distance dcosh since both are higher-than-second-order statistics, which further proves the robustness of the spectral approach. In fact, these two features are irrelevant and do not provide information that can successfully improve the classification performance (cf Table 4.1 and discussion in section 4.1.1).

4.4.2 Comparison between sensors' placements

Table 4.3 shows the distribution of correctly classified and misclassified instances between the activities using SF and SpD methods with the KNN, NB and C4.5 classifiers after a 10-fold cross-validation technique. From this table, it can be noticed that the activities 'running' and 'walking' are accurately detected using all different classifiers. They were correctly classified (100%) using SpD method with a KNN or C4.5 classifier, which support the idea that ambulation behaviors involve quasi-periodic movement of the body and can be thus identified using spectral distances measures. However, 'resting' and 'car riding' tend to be mixed up with each other. They contribute mainly to the performance drop of our method as compared to the reference SF method for all the three classifiers: for example, over a hundred more car riding sequences are interpreted as resting by the NB classifier against 32 misclassified sequences with SF method.

Table 4.3 - Confusion matrices for the Naïve Bayes (NB), K-nearest neighbors (KNN) and Decision Tree (C4.5) classifiers. On the left, for the SF method and on the right, for the SpD method. (Validation technique: 10-fold cross validation).

K-Nearest Neighbors (KNN)

	SF method					SpD method					
	Run	Walk	Bike	Rest	Car		Run	Walk	Bike	Rest	Car
Run	**288**	0	0	0	0	Run	**288**	0	0	0	0
Walk	0	**288**	0	0	0	Walk	0	**288**	0	0	0
Bike	0	0	**288**	0	0	Bike	0	0	**287**	0	1
Rest	0	0	0	**281**	7	Rest	0	0	0	**285**	2
Car	0	0	0	3	**285**	Car	0	0	1	6	**281**

Naïve Bayes (NB)

	Run	Walk	Bike	Rest	Car		Run	Walk	Bike	Rest	Car
Run	**288**	0	0	0	0	Run	**288**	0	0	0	0
Walk	0	**287**	1	0	0	Walk	0	**278**	10	0	0
Bike	0	0	**287**	0	1	Bike	0	0	**285**	0	3
Rest	0	0	1	**266**	21	Rest	0	0	1	**260**	27
Car	0	0	0	32	**255**	Car	0	0	1	134	**154**

Decision Tree (C4.5)

	Run	Walk	Bike	Rest	Car		Run	Walk	Bike	Rest	Car
Run	**287**	1	0	0	0	Run	**288**	0	0	0	0
Walk	0	**288**	0	0	0	Walk	0	**288**	0	0	0
Bike	0	1	**286**	0	1	Bike	0	0	**287**	0	1
Rest	0	0	1	**275**	12	Rest	0	0	0	**285**	2
Car	0	0	0	9	**279**	Car	0	0	1	6	**281**

Figure 4.3 also shows that the best classification accuracy (95.7%) is achieved with the KNN classifier using only the sensor data from the ankle (less than 5% drop). Most importantly, the KNN classifier has consistently higher performance results compared with the other two methods. The results also suggest that the information provided by the sensors from the lower limbs is more valuable for the discrimination of the investigated ambulatory types.

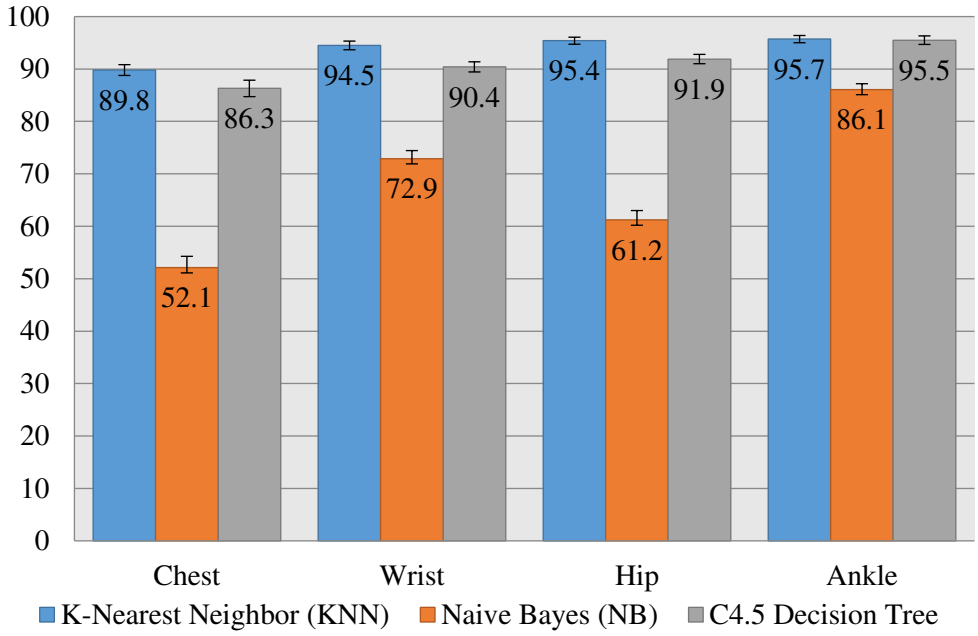


Figure 4.3 - Overall classification accuracies per unit position for the three classifiers using SpD method.

4.4.3 **Influence of HR data**

We focus in this part on analyzing the ability of HR information in the enhancement of the PAR results. Thus, we followed the same procedure in combining the acceleration and HR data. It is important to mention that subjects were executing each activity with no constraints in manner and speed, which yields to widely different HR signals from a person to another. First, the mean value of the HR was computed and used for PAR with a performance ranging from 50% to 60% only (2nd column), while after normalization the same HR data could achieve an accuracy rate of up to 87% (4rd column). Indeed, the inter-subject variability can be compensated through the normalization process using both the mean and standard deviation statistics of the subject.

Likewise, the overall accuracy of the system combining acceleration and HR information was substantially enhanced whatever the considered classifier, as shown in Table 4.4, (reaching 91.3% compared to the 87.8% with no HR for the NB classifier, 99.7% compared to 99% for the KNN classifier and 96.3% compared to 95.9% for the C4.5 classifier).

Finally, in Figure 4.4, we provide an overall comparison of the PAR results with the different sensor positions (chest, wrist, hip, and ankle) and the different classification algorithms (KNN, NB, C4.5) for the distance measures (SpD vs SpD+$\overline{HR}$). An increase of performance for all the sensor positions with all the classification algorithms is clearly noticed. Furthermore, we would like to note that adopting the HR monitor plus only a single accelerometer sensor placed on the ankle or the hip could be sufficient for ambulatory activities classification without risking excessive performance drop (less than a 1% drop) (Table 4.4, 5th column compared to Figure 4.4).

Table 4.4 - Overall Performances (accuracies ± standard deviation) with un-normalized HR and normalized $\overline{HR}$ information.

	SpD	HR	SpD + HR	$\overline{HR}$	SpD + $\overline{HR}$
KNN (%)	99.0 ± 0.4	54.0 ± 1.7	99.2 ± 0.4	87.9 ± 1.1	99.7 ± 0.3
NB (%)	87.8 ± 1.2	49.4 ± 2.0	88.0 ± 1.2	82.8 ± 1.8	91.3 ± 0.9
C4.5 (%)	95.9 ± 1.1	60.5 ± 1.8	95.9 ± 0.9	87.5 ± 1.4	96.4 ± 0.6

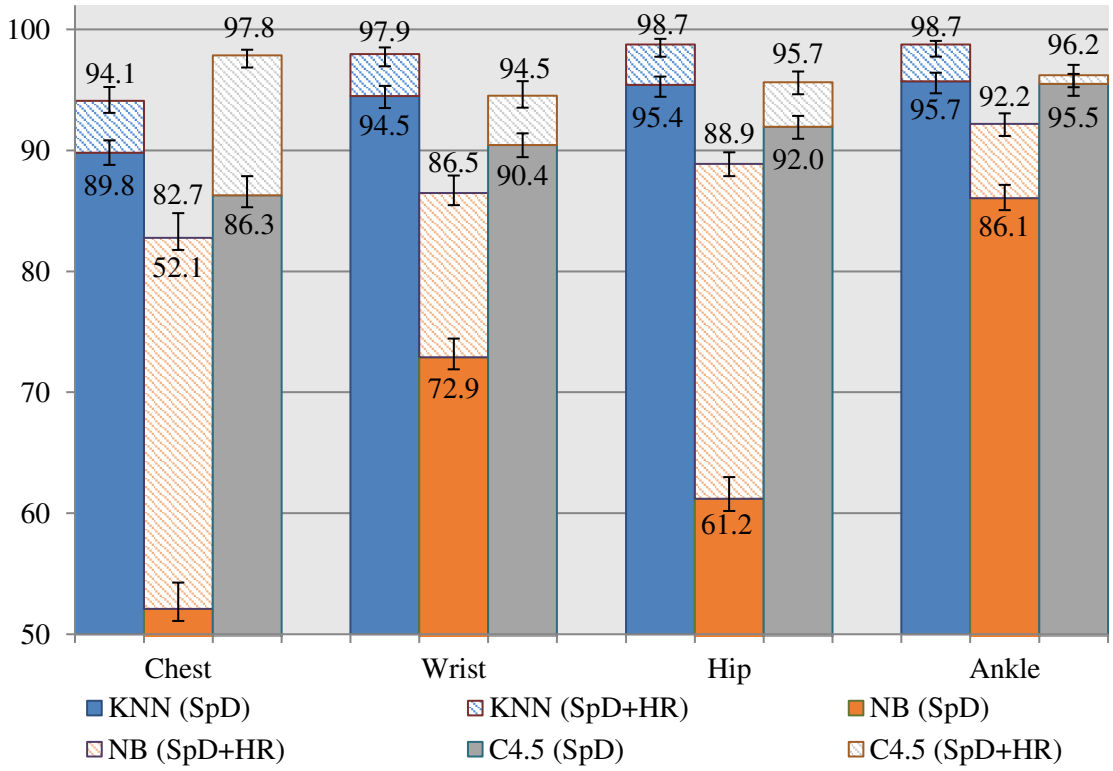


Figure 4.4 - Overall classification accuracies and their standard deviations for each sensor position before and after including the normalized HR information.

4.5 Conclusion

In this chapter, we developed a novel classification method based on spectral distances measures. The overall mean accuracy proves the pertinence of this technique to automatically distinguish between locomotion behaviors while reducing the computation and the complexity of the classification algorithms. We believe that the spectral features contain sufficient information required to classify PA. A new reduced set of spectral-based features (60 spectral distances when considering all the 4 sensor positions) hence are better voted to implement in the PAR systems than the large traditional set of empirically hand-picked time-and-frequency-domain features (324 features when considering all the 4 sensor positions). Moreover, we joined the acceleration signals with the HR signal and showed that the

classification performance can benefit from sensor data fusion of both physical and physiological signs. Based on the results obtained, we concluded that a multi-sensor can be reduced to a duo-sensor based classification system without meaningful lose in terms of performance accuracy. However, the work presented in this chapter was only tested on a limited dataset engaging few activities of locomotion type. To further validate our proposed method, we aimed to test it on a bigger dataset in terms of number of subjects and number of activities. Consequently, we developed a new pragmatic protocol as detailed in the previous chapter. Our pragmatic protocol seeks to provide naturalistic data, which will help to address these gaps. We believe that by enlarging the range of activities in question and by affording a naturalistic context in which the activities are performed, this presents a fruitful area for further investigation. In the next chapter, we first evaluate the two developed methods (SpD and SF method) on seven categories of daily life activities: household, walking, running, cycling, going up/down stairs and playing basketball performed in a controlled manner from session S1 dataset. Then, the performance of these methods in predicting the activities of the dataset from the free-living session S3 is examined. The acceleration and physiological signals recorded from four different devices were tested throughout this study. The optimal placement of the sensors on the body and the contribution of the HR and BR parameters are also assessed. Finally, an approach for adapting the recognition system to the real-life conditions is proposed.

Chapter 5
FREE-LIVING ACTIVITY CLASSIFICATION

In the previous chapter, we developed two features extraction methods and applied classification algorithms to identify five sedentary and ambulatory activity types from acceleration and HR data. Both SpD and SF methods proved their accuracy in detecting the activities in a controlled scenario. However, how these algorithms will perform in free-living scenario is undetermined. This is critical since their final implementation is aimed towards identifying sedentary behaviors and activities in real-world applications. Contrary to the restrained experimental protocols, in a naturalistic environment, people perform less constrained activities differently for unknown time intervals. Consequently, we hypothesized that the accuracy of the classification algorithms in controlled settings does not accurately reflect their performance under free-living settings. It is important to evaluate the reliability of these algorithms in naturalistic conditions and to adapt the classification models to such conditions. This is the main objective of this chapter.

A step forward towards achieving an advanced analysis of PA is presented in this chapter. We first introduce a platform that we developed in order to manipulate in depth the dataset generated by the pragmatic protocol. The platform, called RACHA (Real-time Application for the Classification of Human Activities), serves as a complete interface embedding all the analysis stages required for the recognition of human activity. It provides a representation of the 20 participants' profile in terms of the performed activities during the triple-session protocol. The objective of this interface includes visualizing and analyzing the collected signals, and then exploiting step-by-step the construction of the activity recognition model.

Then, a necessary adjustment to the SpD method is proposed and described in this chapter. Using the pragmatic triple-session protocol, we exposed the preliminarily results obtained from manipulating the acceleration, gyroscope, magnetometer, HR and BR signals recorded by the Zephyr, Shimmer and ActiGraph devices, positioned on different parts of the body (hip, wrist, ankle and chest) (rf. Chapter 3). The responses of the classification models in detecting the activities performed in each of the standardized, semi-standardized and free-living sessions of the protocol are reported. We also presented the results of adding a set of free-living data to the training dataset and studied its impact on the performance of the system in recognizing the free-living activities.

In literature, different attempts were conducted in order to derive $\dot{V}_E$ and then EE as shown in Chapter 2 (section 2.5.2.3.3). Interestingly, we conducted, at the end of the chapter, a preliminary study for measuring VE using a chest strap device and we showed the limitations to overcome in manipulating this approach.

5.1 RACHA Platform: Real-time Application for the Classification of Human Activities

The Graphical User Interface (GUI) - RACHA was developed using the MATLAB software. The platform includes two front-end interfaces, one for data visualization and the second for activity classification.

5.1.1 Objectives of RACHA Platform

- Examine raw data signals from the different variables that were collected from each sensing device, from each sensor's location on the body and from each participant.
- Visualize the continuous time-signals in each session of data acquisition and detect the signal variations among the different activities by selecting and enlarging specific time windows.
- Extract a set of features from the selected variables and devices and compare between the two methods of features extraction (SpD and SF).
- Create and learn a classifier using training data from session S1, S2, S3 or a combination of sessions and evaluate the classifier while varying the test-subject.

- Study the performance of the developed classification models in classifying activities of test-sets from all three sessions S1, S2 and S3.
- Customize parameters in the classification process in order to test different scenarios of classification and explore their effects on the model performance.

5.1.2 Data Visualization Interface

The data visualization interface includes the ability (i) to select a subject from a list of subjects, (ii) to choose the session and (iii) to specify the sensor (such as Zephyr or ActiGraph) and the parameter (such as HR or AccelX (x-axis acceleration)) to be displayed. Based on these choices, the interface can then load and plot the selected signal and simultaneously mark the succession of the related activities. In Figure 5.1, a sample signal of AccelX collected during session S1 from the Zephyr system worn by the subject referred to as 'SABAL' is loaded into the plot zone of the interface. The activities can be marked on the plotted signal simultaneously. The plotted blue line in this figure refers to the succession of the activities executed in function of time. Each number represents an activity from Table 5.1. For further exploration, a zoom in tool is available to navigate between windows and activities.

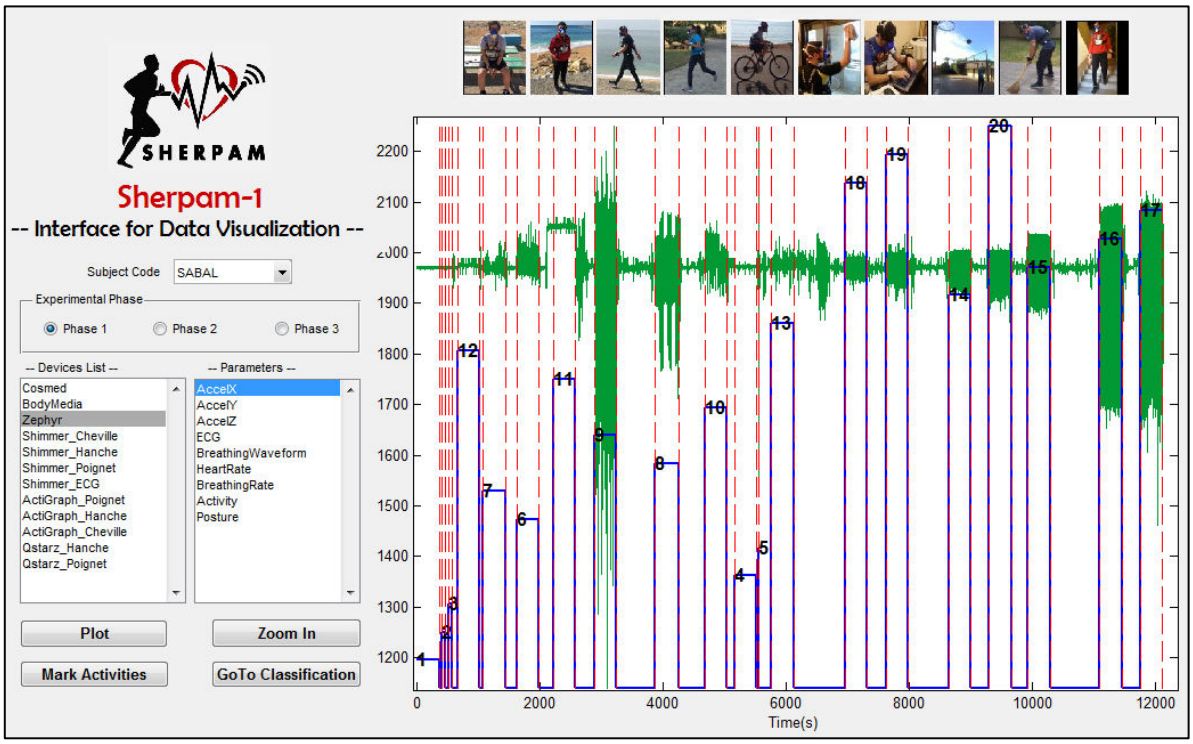


Figure 5.1 - Data Visualization interface. A sample x-axis acceleration (AccelX) signal collected during session S1 from the Zephyr chest strap worn by the subject referred to as 'SABAL' is loaded into the plot zone of the interface. The blue line refers to the succession of the performed activities. Each number represents an activity (rf. Table 5.1).

Table 5.1 - The activity labels and their corresponding numbers as considered in RACHA platform.

Activity N°	Activity Label	Activity N°	Activity Label
1	Sitting	**11**	Lying
2	Controlled breaths 40bpm	**12**	Working on PC
3	Controlled breaths 70bpm	**13**	Car riding
4	Standing	**14**	Slow walking
5	Brisk Jump	**15**	Fast walking
6	Sweeping inside	**16**	Slow running
7	Cleaning the windows	**17**	Fast running
8	Going up/down stairs	**18**	Slow cycling
9	Playing basketball	**19**	Fast cycling
10	Sweeping outside	**20**	Normal walking

5.1.3 **Activity Classification Interface**

The activity classification interface (Figure 5.3) can be directly accessed via the data visualization interface. It serves for the design, the training and the evaluation of the activity recognition model. Up till now, data from 9 modules were integrated in the classification panel of this interface, the signals that are configured to serve this matter are: 3-axis accelerations (AccelX, AccelY and AccelZ) from each of the 3 ActiGraph modules mounted on the ankle, wrist and hip; pressure signal, 3-axis signals from the magnetometer and gyroscope, 3-axis accelerations from the low noise accelerometer and 3-axis accelerations from the wide range accelerometer from each of the 3 Shimmer modules mounted on the ankle, wrist and hip; 3axis-acceleration, HR, BR, activity and posture signals from the Zephyr chest strap device; and speed from the 2 Qstarz modules that were mounted on the wrist and hip. Thereby, it is possible to build the matrix of features based on the chosen signals from this list. The two methods that were extensively explained in Chapter 3 (the SpD and the SF methods) are also implemented in this platform. As a first step, the user can select the desired method and specify the window length of the observations before building the matrix of features. The interface allows visualizing each feature vector from the features matrix for the 3 sessions and the 20 subjects. A sample of one of the features extracted from the HR signal is plotted in Figure 5.2.

Following the feature extraction process, a classification panel is available integrating the possibility to choose both the classifier and the training/testing sets from the three data acquisition sessions. The interface makes use of the built-in KNN and the DT classifiers, as they had been voted best in existing activity recognition studies (Altun et al., 2010; Bao et al., 2004; Maurer et al., 2006) and in our previous study as well (Chapter 4). Using the computed matrix of features and based on the inclusion criteria of data in the training set, the program executes the training process. Once the model is created, the analysis of its responses on the new test-set of data is executed. It is also possible to define the test-set from the

list of subjects' data to evaluate the response of the developed model on the activities of a specific subject. The confusion matrix, the accuracy, the sensitivity and the F-score for all the three sessions are displayed and the predicted vs. real classes of all sessions are plotted at the end of the process.

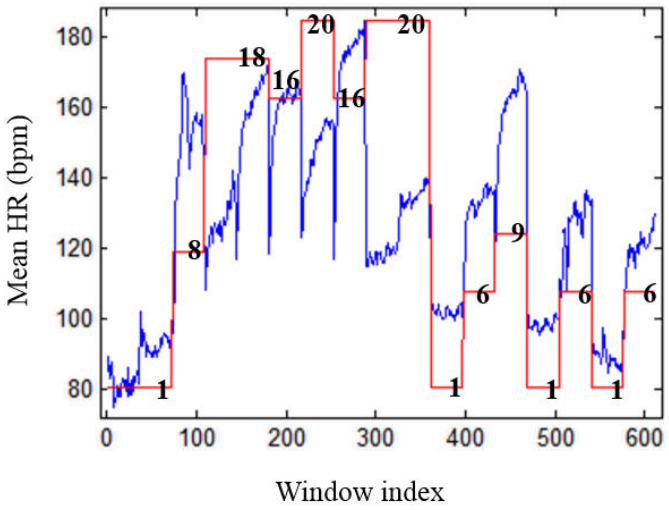


Figure 5.2 - Example of a HR feature extracted from one subject performing session S1 plotted in blue.
The red line represents the series of performed activities, where each number refers to a category of activity (rf. Table 5.2).

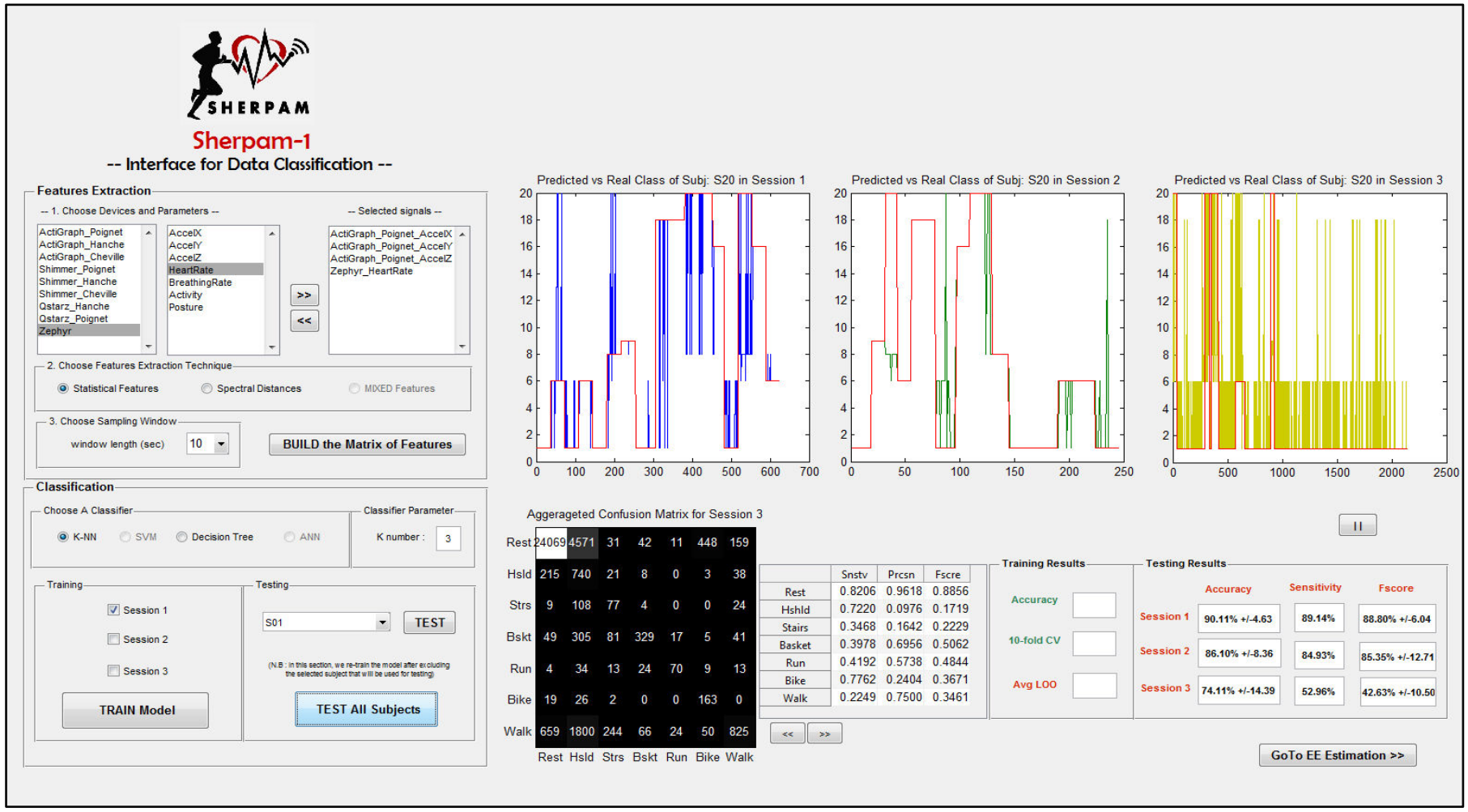


Figure 5.3 - Activity Classification Interface showing an example of classification results.
From the classification panel, it can be observed that the signals selected and used in this classification example are 3D-acceleration from the wrist-mounted ActiGraph module along with the HR signal from the Zephyr chest strap. The Statistical Features are the chosen features in this case, the feature set is divided on 10-s window length. The KNN classifier (with k = 3) is trained using S1-training set. The aggregated confusion matrix resulted from this classification, the accuracy, sensitivity and F-score are displayed and the predicted vs real classes of all sessions are plotted at the end of the process.

5.2 Details on the pragmatic dataset

5.2.1 Labelling the activities groups

Examples of 10s-window signals for each of the 17 activities is presented in Figure 5.4. The analysis in this study is made on the recognition of 7 categories that grouped the 17 considered activities as presented in Table 5.2. (*Note that the activities 2, 3 and 5 (rf. Table 5.1) were excluded from this study*).

Table 5.2 - The 7 categories created from the 17 activities.

Category N°	Category Label	Activity Label
1	Sedentary (Sdn)	01. Sitting 04. Standing 11. Lying 12. Working on PC 13. Car riding
6	Household (Hsd)	06. Sweeping inside 07. Cleaning the windows 10. Sweeping outside
8	Stairs (Str)	08. Going up/down stairs
9	Basketball (Bkt)	09. Playing basketball
16	Running (Run)	16. Slow running 17. Fast running
18	Cycling (Cyl)	18. Slow cycling 19. Fast cycling
20	Walking (Wlk)	14. Slow walking 15. Fast walking 20. Normal walking

5.2.2 S3-photo analysis and annotation protocol

The recorded images in S3 were annotated during two stages to avoid ambiguity in labeling the activities. The first stage occurred at the end of the S3-days after retrieving the devices from the participants. Firstly, the participants were free to delete undesired photos they did not wish us to see. Then a visual inspection was carried out together with the participants so that they can inform us of what they were doing in each photo. Although one can easily guess the performed activities in most cases, sometimes it was recommended that the participants indicate exactly what they were doing to ensure correct annotation. An example of the annotated photos is presented in Figure 5.5. A set of commented photos for each participant was constructed. On a second stage, only the annotated activities that were included in the 7 investigated categories in this study, were selected to be taken into consideration. Using the 'Autographer' manufacturer supplied software, we were able to obtain the timestamp of each photo allowing us to identify the start and end time of each activity with a 1-s resolution.

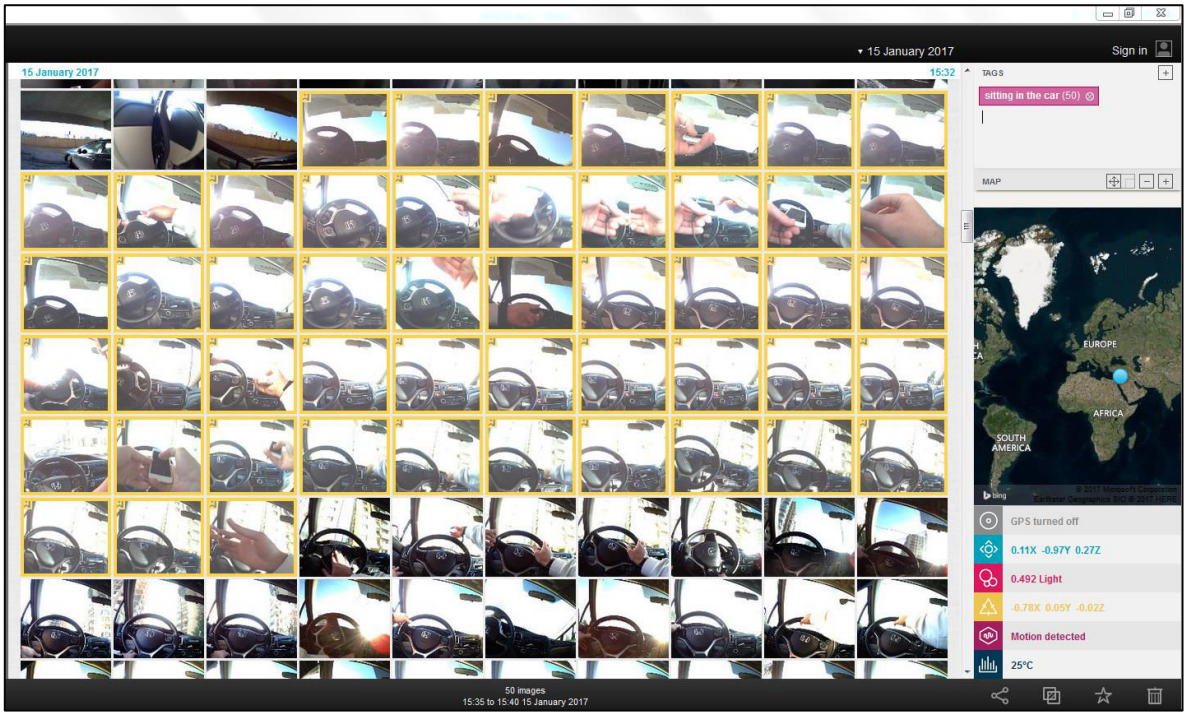


Figure 5.4 - A view of the Autographer camera interface developed to download, visualize and annotate the captured photos.

5.2.3 **Activities composition of the sessions**

As detailed in Chapter 3 (section 3.4.2), the pragmatic data collection protocol is composed of 3 sessions. In S1, the standardized activities were executed for nearly the same duration. However, this was not the case in the semi-standardized session S2 or the free-living session S3. Thereby, this section details the set of activities performed during all three sessions.

- In session S1:

Each subject performed all 17 activities for a duration of approximately 6-min each. For 20 subjects, we have thus a total of 6 min x 20 subjects = 120 min for each activity. The 1st column of the Table 5.3 shows the number of minutes of data that were collected in S1. For instance, the sedentary category groups 5 types of activities, the number of minutes that includes in S1 is thus around 600 (= 120 x 5) min.

- In session S2:

Since the subject can freely choose the duration of each activity to perform, we don't have the same number of minutes between activities. Going up/down stairs activity occurred the least, while the sedentary activities were performed more often.

- In session S3:

98 hours of activities are extracted and manipulated in the experiments (around 5 hours per subject). The sedentary group account for 82% (4889 min out of the 98 hours) of all activities recorded. The activity of walking constitutes a proportion of 10.3% and the remaining 2.9%, 2.3%, 0.6%, 0.6%, 0.5% belongs to the household, basketball, stairs, cycling and running groups respectively.

Table 5.3 - The distribution of minutes recorded of categories in each session.
Categories - Sdn: sedentary, Hsd: household, Str: going up/down stairs, Bkt: basketball, Run: running, Cyl: cycling, Wlk: walking.

	Triple-Session Protocol (min)		
	S1	**S2**	**S3**
Sdn	620	317	4889
Hsd	357	155	171
Str	123	59	37
Bkt	123	62	138
Run	227	59	28
Cyl	252	80	35
Wlk	355	105	611
Overall	**34h**	**14h**	**98h**

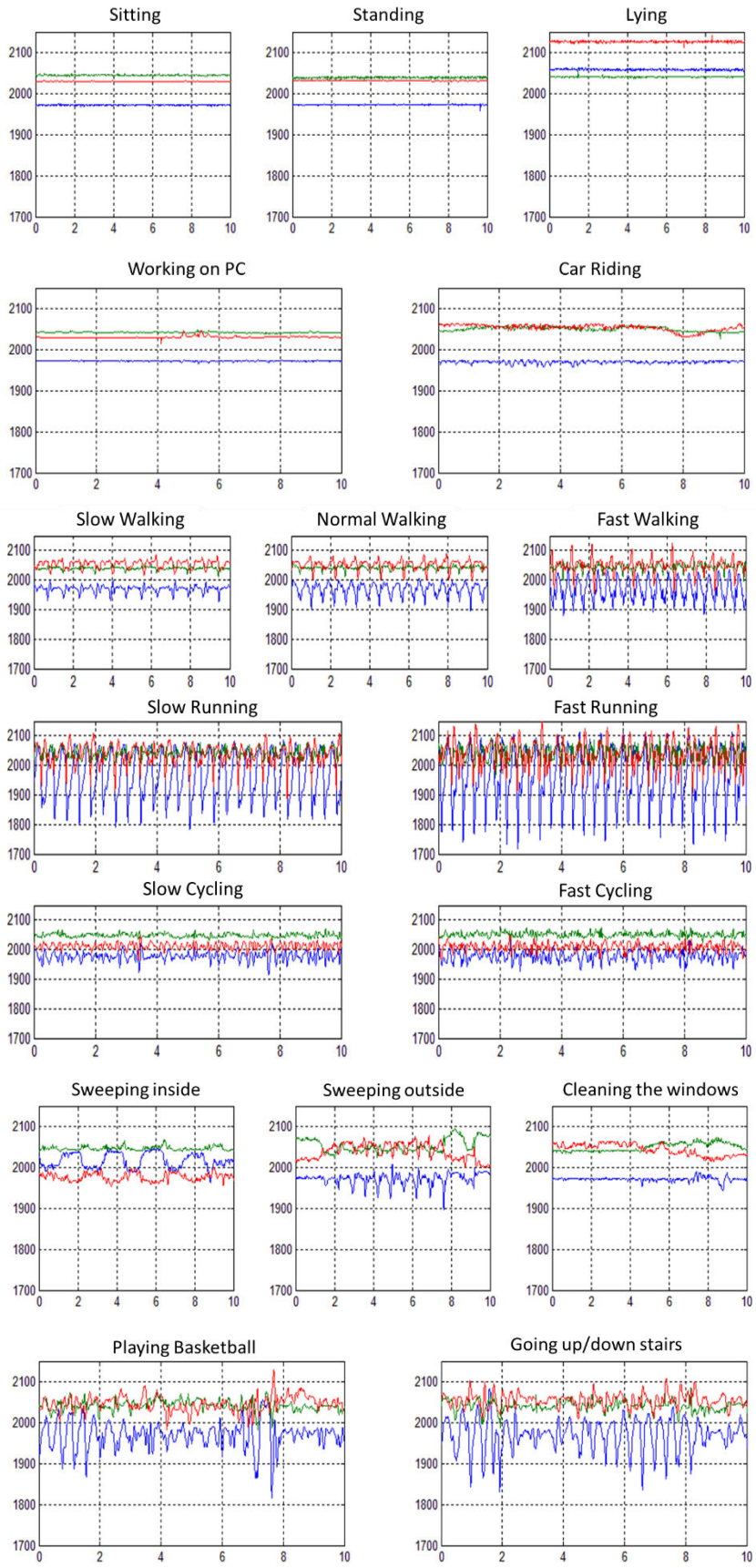


Figure 5.5 - Acceleration signals from 3D accelerometers of the Zephyr chest strap during the 17 performed activities.

5.2.4 **Extracted Features**

In order to validate both SF and SpD methods on this new dataset, the extracted features from the physical signals (3D-accelerometers, 3D-magnetometers and 3D-gyroscopes) were the same SF and SpD features as in Chapter 4.

Unfortunately, when applying the method as detailed in the previous chapter, the results were less than satisfactory which motivated us to explore the reason behind that. In this section, we brought a necessary modification to the equation used to compute the spectral distances in SpD method so that it would be better applied on the new pragmatic dataset.

- **Modification to SpD method:** As detailed in Chapter 4 (section 4.2.1), the Itakura-Saito distance dIS is the spectral distance considered in our study. It is computed between the PSD of a random process $s_1$and that of the given model s_2. The equation of d_{IS} was given by:

$$d_{IS}(\boldsymbol{s_1}, \boldsymbol{s_2}) = \frac{1}{\sigma_2^2}\boldsymbol{a_2^T} R_p(\boldsymbol{s_1})\boldsymbol{a_2} - \log\frac{\sigma_1^2}{\sigma_2^2} - 1 \qquad (3)$$

where Rp(s) is the square Toeplitz matrix of dimension (p+1)*(p+1) whose (k, j) element is **r** (k - j), 0 ≤ k, j ≤ p, r(n), n ∈ N is the discrete autocorrelation function and **a** = $(a_0, a_1, \ldots, a_p)^t$ the column vector composed of the autoregressive model coefficients.

When we computed the d_{IS} distance on the pragmatic dataset, the seven activity models were created on the basis of the 6-min activity signals collected in session S1 from all 20 subjects. Therefore, the total energy of the models created in this manner is much higher compared to the energy of the 10-s windows from test-sets.

In (Eq.3), the 2nd term $\log\frac{\sigma_1^2}{\sigma_2^2}$ that represents the fraction of the 10-s signal's energy to the model's energy being dominant in $d_{IS}(\boldsymbol{s_1}, \boldsymbol{s_2})$, the 1st term that represents the distribution of the energy $\frac{1}{\sigma_2^2}\boldsymbol{a_2^T} R_p(\boldsymbol{s_1})\boldsymbol{a_2}$ has no influential weight in this equation. In this case, the distance in (Eq. 3) represents the ratio of energies rather than their spectral distribution.

To concentrate uniquely on the spectral distribution, we proposed to modify the computation of the distance. The modified $d_{IS}{}^*(\boldsymbol{s_1}, \boldsymbol{s_2})$ can be written as follows:

$$d_{IS}{}^*(\boldsymbol{s_1}, \boldsymbol{s_2}) = \frac{1}{r_{1(1)}}\boldsymbol{a_2^T} R_p(\boldsymbol{s_1})\boldsymbol{a_2} + \frac{1}{r_{2(1)}}\boldsymbol{a_1^T} R_p(\boldsymbol{s_2})\boldsymbol{a_1} \qquad (3^*)$$

The distances $d_{IS}{}^*(\boldsymbol{s_1}, \boldsymbol{s_2})$ are calculated between each window and the seven activity reference models of the pragmatic dataset so that each of the windows will be presented as a set of distances values. The resulting feature vector that we calculated for each signal variable consists of 8 (=7 distances + energy σ_1^2) features.

A total of 27 features per signal were computed for each window in case of SF method and of 8 features in case of the modified SpD* method.

- **For the physiological signals:** we extracted the mean value on each window from BR signal. Four features were also derived from the HR signal: mean HR (HRm), HR above rest or net HR ($HRnet$), percentage of maximal HR ($HRabs$) and percentage of HR reserve ($HRrv$) values from each window of HR data, as follows:

$$HRnet = HRm - HRrest$$

$$HRabs = \frac{HRm}{HRmax} \times 100$$

$$HRrv = \frac{HRm - HRrest}{HRmax - HRrest} \times 100$$

Where $HRrest$ is the resting HR value and $HRmax = 207 - 0.7 \times age$, related to each subject (Roy et al., 2015). These HR features have been previously used in EE estimation studies (de Müllenheim et al., 2017; Strath et al., 2000). They allow to quantify the HR level of the subject compared to the rest (i.e. activity or inactivity) and maximal values (i.e. maximal effort).

5.3 Results

Throughout this study, the reported results are issued from a leave-one-subject out validation process. Classifiers were trained on activity data belonging to all subjects but one, which was used for testing. This process was repeated 20 times such that the observations of each of the 20 subjects are used once as the test-set data. The leave-one-subject-out is a challenging, yet strong validation process. It allows to create and validate a generalized system independent from the subjects' variability. It permits to identify the response of the algorithm if unseen subject's observations are tested.

There are a lot of different performance measures that can be used to evaluate the algorithms' robustness, as explained in Chapter 2 (section 2.3.5.3). The performance measure that is mainly used in this study is the average F-score (defined in Chapter 2 (section 2.3.5.3)) instead of the commonly-used overall accuracy. Given that the seven categories are imbalanced, specifically in session S3 (rf. Table 5.3), the accuracy measure in this case would be biased and influenced by the dominant class (the sedentary category) that has a big number of observations relatively to other classes. For this reason, the F-score is favored over the accuracy measure to evaluate the classification performance in this case. The confusion matrices, the sensitivity and precision measures were also presented in some sections for deeper analyses of the results. It is important to note that the term accuracy that is used in the rest of the chapter refers to the F-score measure.

In sections 4.1 to 4.4, the training sets that were used to learn the classifiers to identify the 7 categories of activities are the controlled data of session S1, since our aim was to test the validity of the controlled-

based algorithms that are implemented in state-of-the-art PAR systems as a reference. Another point to consider is that features were extracted in 10-s windows with no overlap between consecutive windows since the transitions in this study are being neglected. Five exploratory analyses were conducted in the following sections that aim at comparing the recognition performance among sessions, window lengths, sensors and sensor positions, variables and finally among different types of training sets.

5.3.1 **Comparison between sessions**

In this section, we compare the classification rates of the KNN and DT classifiers on test-set from standardized session S1 versus test-sets from semi-standardized session S2 and free-living session S3 using both SpD* and SF methods. The mean and SD of the F-score values obtained from the 3-axis acceleration of the Zephyr chest strap, the ankle-ActiGraph and the ankle-Shimmer modules[3] are grouped in Table 5.4 for the KNN classifier and Table 5.5 for the DT classifier. Overall KNN classification accuracy for the S1 test-set was between 95.4% (Zephyr) to 89.9% (ActiGraph_ankle) using the SF method and between 89.4% (Zephyr) to 82.4% (ActiGraph_ankle) using SpD* method. Overall DT classification accuracy for the S1 test-set ranged from 88.7% (Zephyr) to 85.5% (ActiGraph_ankle) using the SF method and from 86.8% (Zephyr) to 79.4% (ActiGraph_ankle) using the SpD* method. The results of S2-classification are slightly better than those obtained for S1-classification for all the three devices, using either KNN or DT classifiers and using either SF or SpD* method. The Zephyr accelerometer marked an overall accuracy of 96.29% in S2 against the 95.44% in S1 when using the SF method with a KNN classifier. The KNN classifier reported better results than the DT classifier, yet when tested on the S3 dataset, the performance of both classifiers decreased dramatically using either SpD* or SF methods and this was noticed for all the three devices. For the S3 test-set, the KNN classification accuracy using SF and SpD* methods dropped to 50.7% (Zephyr), 49.8% (ActiGraph_ankle) and 46.9% (Zephyr), 45.6% (ActiGraph_ankle) respectively.

Since the KNN classifier was more consistent than the DT classifier for the classification of S1 and S3 test-sets (Table 5.4 compared to Table 5.5), only the results from the KNN classifier were retrained to be presented in the following sections. In order to compare our results with those of the existing state-of-art systems, the results that we present next are the ones obtained from the classification method taken from literature (SF method).

[3] We have only presented the results from the ankle modules to reduce the load of information in one table, yet a comparison between the different modules is discussed afterwards.

Table 5.4 - Overall F-score measures (%) ± standard deviation of the KNN classifier using SF and SpD* methods on the 3D-acceleration signals of Zephyr, ankle-worn ActiGraph (AtG_A) and ankle-worn Shimmer (ShM_A) modules, for the standardized S1, semi-standardized S2 and free-living S3 test-set.

	S1	S2	S3	
Zephyr	95.4 ± 3.7	96.3 ± 2.8	50.7 ± 16.1	SF
	89.4 ± 5.7	92.6 ± 5.0	46.9 ± 13.2	SpD*
AtG_A	89.9 ± 5.7	90.7 ± 6.0	49.8 ± 15.1	SF
	82.4 ± 6.9	85.9 ± 5.7	45.6 ± 13.0	SpD*
ShM_A	91.8 ± 8.7	94.1 ± 5.8	55.0 ± 15.8	SF
	86.1 ± 7.1	87.5 ± 4.5	50.8 ± 12.2	SpD*

Table 5.5 - Overall F-score measures (%) ± standard deviation of the DT classifier using SF and SpD* methods on the 3D-acceleration signals of Zephyr, ankle-worn ActiGraph (AtG_A) and ankle-worn Shimmer (ShM_A) modules, for the standardized S1, semi-standardized S2 and free-living S3 test-set.

	S1	S2	S3	
Zephyr	88.7 ± 4.2	90.3 ± 6.1	47.5 ± 13.1	SF
	86.8 ± 5.4	90.2 ± 4.6	44.3 ± 12.9	SpD*
AtG_A	85.5 ± 7.5	89.8 ± 6.6	50.3 ± 14.2	SF
	79.4 ± 6.9	82.9 ± 6.3	44.8 ± 12.4	SpD*
ShM_A	87.7 ± 12.3	92.0 ± 6.2	50.0 ± 14.8	SF
	83.2 ± 6.2	83.9 ± 6.4	49.1 ± 13.1	SpD*

5.3.2 Comparison between window lengths

The performance of the classifier was tested using window lengths varying from 5 to 60s. For S1 test-set classification, increasing the lengths from 5 to 60s resulted in improved classification accuracy of 4.2% for Zephyr and ActiGraph_ankle and 3.5% for Shimmer_ankle (Figure 5.9). For S3 test-set, a reduction occurred in the 60s-length classification accuracy (-2%) for Shimmer_ankle and (-3.2%) for Zephyr compared to 30s-length (Figure 5.10). Increasing the windows from 5s to 10s lengths resulted in an increase in the classification accuracies for both S1 and S3 test-sets and for all three devices. Although there is no sufficient statistical evidence to decide that any of the window length is better than the others, we have chosen to fix the window length to 10s in the remaining sections. We find that a 10-s length represents a compromise between the large windows that cause a delay in the system's response and the small windows that cannot properly describe the activity.

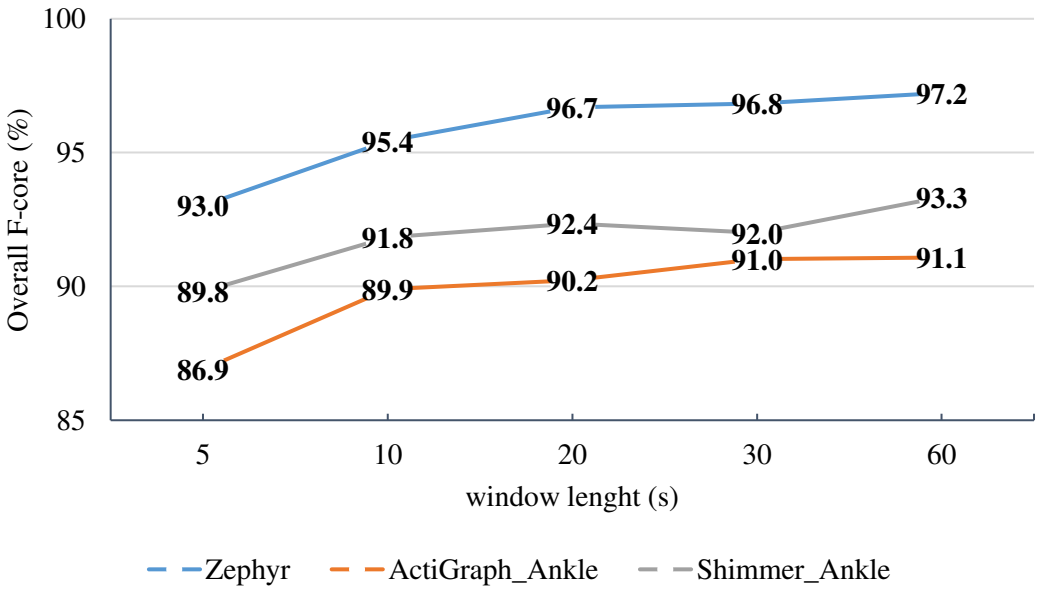


Figure 5.6 - Overall performance (%) of Zephyr, ActiGraph_ankle and Shimmer_ankle with different window lengths for S1 test-set.

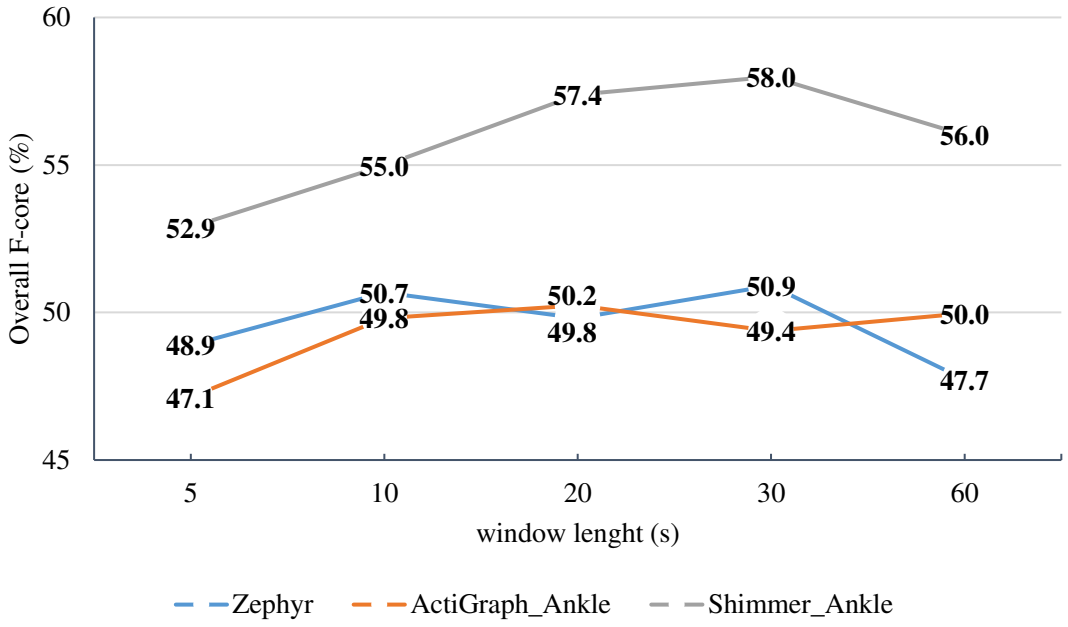


Figure 5.7 - Overall performance (%) of Zephyr, ActiGraph_ankle and Shimmer_ankle with different window lengths for S3 test-set.

5.3.3 Comparison between sensors positions

We compared the performance of the 3D accelerometers of the ActiGraph, Shimmer and Zephyr devices on the four positions: chest, wrist, hip and ankle. Table 5.6 and 5.7 display the performance of KNN classifier for each position and for each activity group in detecting the activities of S1 and S3 respectively. On average, the ankle modules performed the best when compared to the wrist and hip

modules for both the Shimmer and the ActiGraph. This is true when applied on both S1 and S3 test-sets. For S1 test-set, the overall F-score values of the KNN classifier for the wrist, hip and ankle modules were 88.4%, 89.0% and 89.9% respectively for the ActiGraph and 89.5%, 89.9% and 91.8% respectively for the Shimmer.

Table 5.6 - F-score values (%) per category of activities for the different modules and for standardized S1 test-set.
Categories - Sdn: sedentary, Hsd: household, Str: going up/down stairs, Bkt: basketball, Run: running, Cyl: cycling, Wlk: walking.

	ActiGraph			**Shimmer**			**Zephyr**
	Wrist	**Hip**	**Ankle**	**Wrist**	**Hip**	**Ankle**	**Chest**
Sdn	92.7	94.4	92.8	94.1	95.6	93.7	97.7
Hsd	83.5	87.7	85.5	86.8	87.9	85.7	93.6
Str	68.4	81.7	82.3	68.3	83.2	89.0	92.0
Bkt	91.4	83.8	82.9	97.0	83.6	85.4	90.8
Run	99.1	97.0	96.8	98.3	96.2	98.2	99.4
Cyl	94.4	84.9	94.3	94.7	88.8	95.8	95.7
Wlk	89.2	93.8	94.7	87.6	94.2	95.1	98.9
Overall	**88.4**	**89.0**	**89.9**	**89.5**	**89.9**	**91.8**	**95.4**

Table 5.7 - F-score values (%) per category of activities for the different modules and for free-living S3 test-set.
Categories - Sdn: sedentary, Hsd: household, Str: going up/down stairs, Bkt: basketball, Run: running, Cyl: cycling, Wlk: walking.

	ActiGraph			**Shimmer**			**Zephyr**
	Wrist	**Hip**	**Ankle**	**Wrist**	**Hip**	**Ankle**	**Chest**
Sdn	86.8	88.9	84.0	85.9	88.6	87.4	91.5
Hsd	15.8	19.8	15.3	15.8	17.5	16.1	21.3
Str	23.3	35.5	54.6	18.5	35.8	66.7	40.9
Bkt	47.9	50.5	52.8	47.3	46.9	52.5	54.6
Run	48.0	48.5	50.5	40.5	53.7	53.6	55.5
Cyl	32.8	15.9	37.9	40.9	25.5	52.0	38.6
Wlk	35.4	48.7	53.8	32.1	46.8	56.5	52.4
Overall	**41.4**	**44.0**	**49.8**	**40.1**	**44.9**	**55.0**	**50.7**

In correctly classified S1 test-set, the accelerometer of the Zephyr chest strap outperformed (95.4%) the ActiGraph and Shimmer accelerometers of the wrist, hip and ankle positions while in the classification of S3 test-set, the Shimmer_ankle marked the highest accuracy of 55% then the Zephyr with 50.7% and then the ActiGraph_ankle with 49.8%. The ActiGraph_hip and Shimmer_hip modules identified sedentary and household activities better than the ActiGraph_ankle and the Shimmer_ankle modules in

both S1 and S3 test-sets. For household activity, classification accuracy was 87.7% and 87.9% for ActiGraph_hip and Shimmer_hip and 85.5% and 85.7% for ActiGraph_ankle and Shimmer_ankle for S1. An important enhancement is observed in the classification of stairs, walking and cycling groups using the ankle module vs. the wrist and hip modules for both Shimmer and ActiGraph. For S3-stairs activity, classification accuracy was 18.5% and 35.8% for Shimmer_wrist and Shimmer_hip and 66.8% for Shimmer_ankle. The lowest accuracy from all the devices was for the S3-household group and the highest was for the S3-sedentary group.

Table 5.8 - Aggregated confusion matrix, sensitivity and precision (%) for Zephyr module and for S3 test-set.
Categories - Sdn: sedentary, Hsd: household, Str: going up/down stairs, Bkt: basketball, Run: running, Cyl: cycling, Wlk: walking.

	Sdn	**Hsd**	**Str**	**Bkt**	**Run**	**Cyl**	**Wlk**
Sdn	**25199**	3550	28	24	0	494	36
Hsd	245	**753**	1	6	0	11	9
Str	0	22	**189**	8	0	1	2
Bkt	57	238	49	**378**	27	46	32
Run	2	11	12	29	**84**	0	29
Cyl	7	8	1	0	0	**194**	0
Wlk	270	1451	422	112	25	49	**1339**
S (%)	**85.9**	**73.5**	**85.1**	**45.7**	**50.3**	**92.4**	**36.5**
P (%)	**97.8**	**12.5**	**26.9**	**67.9**	**61.8**	**24.4**	**92.5**

Table 5.8 shows the aggregated confusion matrix (from all 20 subjects) in addition to the sensitivity and the precision values for the Zephyr module, which exhibited the highest correct classification for the household activity (21.3%) among the studied devices. From this confusion matrix, we can see that the KNN classifier correctly classified the household activities 73.5% of the time (753 samples) and confused 245, 11, 9 and 6 samples as sedentary, cycling, walking and basketball activities respectively. The poor precision of the household activity (12.5%) is however explained by the big number of samples mainly for the sedentary (3550), walking (1451) and basketball (238) activities that were misclassified as household activity. The sensitivity and the precision of the walking activity were 36.5% and 92.5% respectively. The reason behind this significant difference is that the classifier correctly classified 1339 walking samples and only confused 36 sedentary samples, 32 basketball and 29 running samples as walking activity. However, against the 1339 correctly classified walking samples, 1451 walking samples were misclassified as household, 422 as stairs and 270 as sedentary.

To examine the impact of the combination of more than one accelerometer, we present Figures 5.6 and 5.7 that show the performance of the ActiGraph modules when taken together in the classification process of S1 and S3 test-sets. An increase in the overall classification accuracies of S1 and S3 test-sets is observed when coupling two accelerometers. As an example, the wrist and hip modules that achieved

88.4% and 89% separately, achieved 93.8% when coupled together. Altogether, the classification from the three ActiGraph modules reached 95.5%. However, this combination of the three modules slightly improved the S3 classification: an increase of only 1% is achieved after adding samples from the wrist and hip modules to the ankle module (50.8%) compared to the performance of the ankle module taken alone (49.8%).

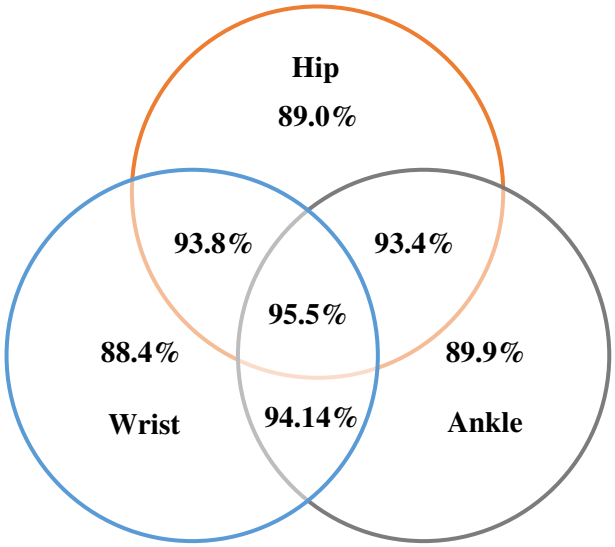


Figure 5.8 – Illustration of the KNN classifier performances for S1 test-set when coupling two and three ActiGraph modules together.

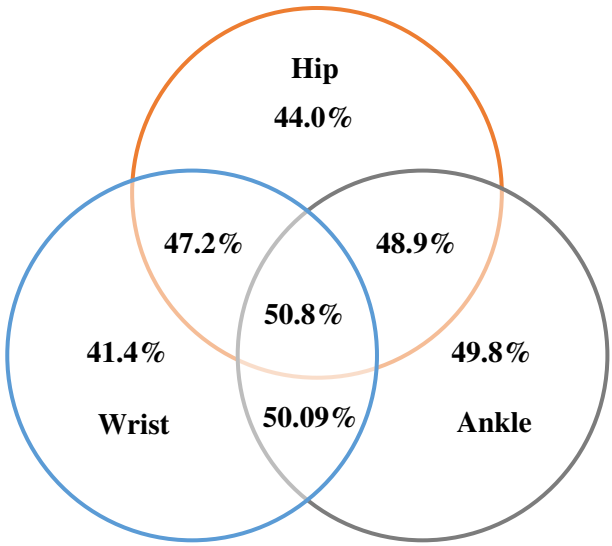


Figure 5.9 - Illustration of the KNN classifier performances for S3 test-set when coupling two and three ActiGraph modules together.

5.3.4 Comparison between sensors' variables

To evaluate the effectiveness of the information brought by the Shimmer gyroscope (Gyr) and magnetometer (Mag) in the recognition, we extracted the corresponding SF features from the 3-axis Gyr and Mag signals separately. The average accuracies are shown in Table 5.9. The classification was enhanced by around 2% after incorporating Gyr at the input of the KNN classifier in both S1 and S3. The magnetometer didn't show any improvement, instead a decrease of 5% was noticed in the S3 results.

Table 5.9 - Overall F-score values (%) of the accelerometer (Accel) taken alone, taken with gyroscope (Accel + Gyr) and taken with magnetometer (Accel + Mag) from Shimmer_ankle module for S1 and S3 test-sets.

	Accel	Accel + Gyr	Accel + Mag
S1	91.8 ± 8.7	93.7 ± 5.6	91.8 ± 5.3
S3	55.0 ± 15.8	56.2 ± 16.0	49.0 ± 15.6

Table 5.10 summarizes the overall classification accuracy achieved by incorporating the vital signals (HR and BR) of the Zephyr device. The average improvement in detecting the S3-activities achieved by the HR and BR data was 0.5% and 1.0% respectively. Although the overall accuracy only slightly increased, a significant increase in detecting stairs activity can be observed after taking HR into account (+ 6.8%) in Figure 5.8.

Table 5.10 - Overall F-score values (%) of the accelerometer (Accel) taken alone, taken with HR (Accel + HR) and taken with BR (Accel + BR) from Zephyr chest strap module for S1 and S3 test-sets.

	Accel	Accel + HR	Accel +BR
S1	95.4 ± 3.7	95.4 ± 3.2	95.8 ± 3.4
S3	50.7 ± 16.1	51.2 ± 16.0	51.7 ± 15.8

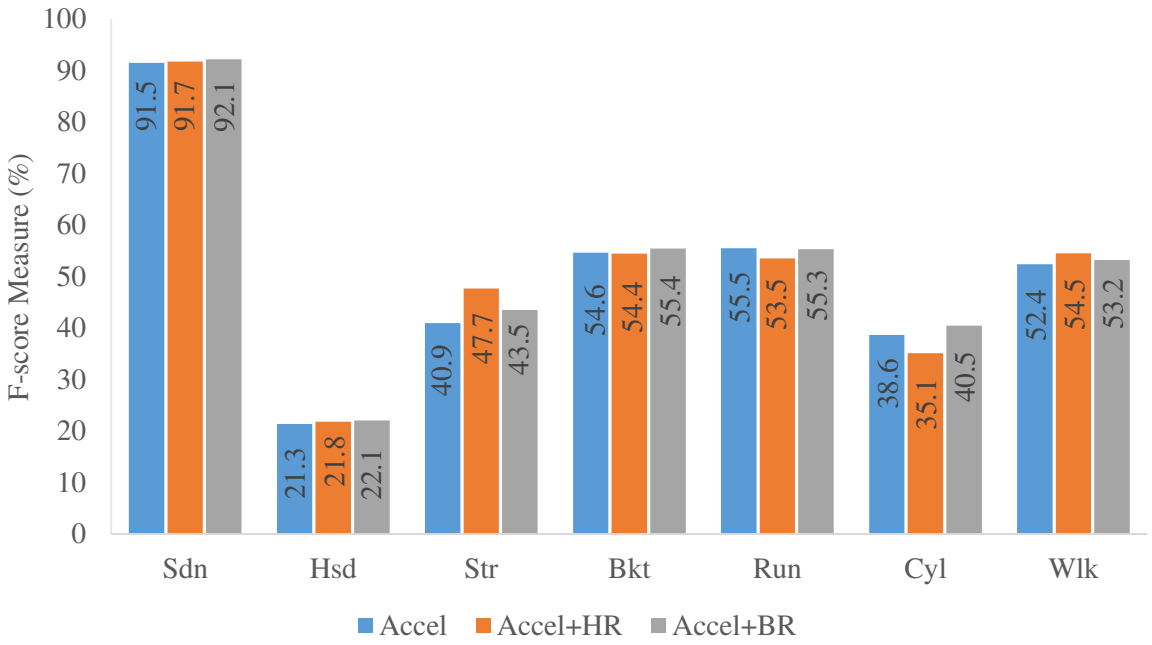


Figure 5.10 - Average accuracy of acceleration (Accel), (Accel + HR) and (Accel + BR) taken from Zephyr module per category in S3 test-set.
Categories - Sdn: sedentary, Hsd: household, Str: going up/down stairs, Bkt: basketball, Run: running, Cyl: cycling, Wlk: walking.

5.3.5 Approach for Real-life Application

To address the poor performance of the algorithms in detecting activities of S3 test-set, we modified the type of data that were used in training the algorithms. As stated in the beginning of section 5.3, all the analysis that we previously presented were issued from classification algorithms trained using the standardized session S1. In this section, we investigate the use of the S3 real-life dataset in the training process. The graph chart in Figure 5.11 presents the overall accuracies of the three following strategies: when using only S1 training-set (TrS1), when using only S3 training-set (TrS3) and when using training-sets from both S1 and S3 (TrS1&S3). On average, the results showed a significant improvement of the accuracy of S3 test-set after adding S3 data to the training-set. The Zephyr, ActiGraph_ankle and Shimmer_ankle witnessed an increase of 5.18%, 4.9% and 2.13% respectively when using TrS3, and 11.47, 9.68% and 7.46% respectively when using TrS1&S3. In this way, the best accuracies that the recognition models could achieve in detecting the seven categories of daily living activities are 62.15% for the Zephyr chest strap, 59.49% for the ActiGraph_ankle module and 62.44% for the Shimmer_ankle module. These values, although moderate, were attained only when controlled and free-living datasets were used in the training process.

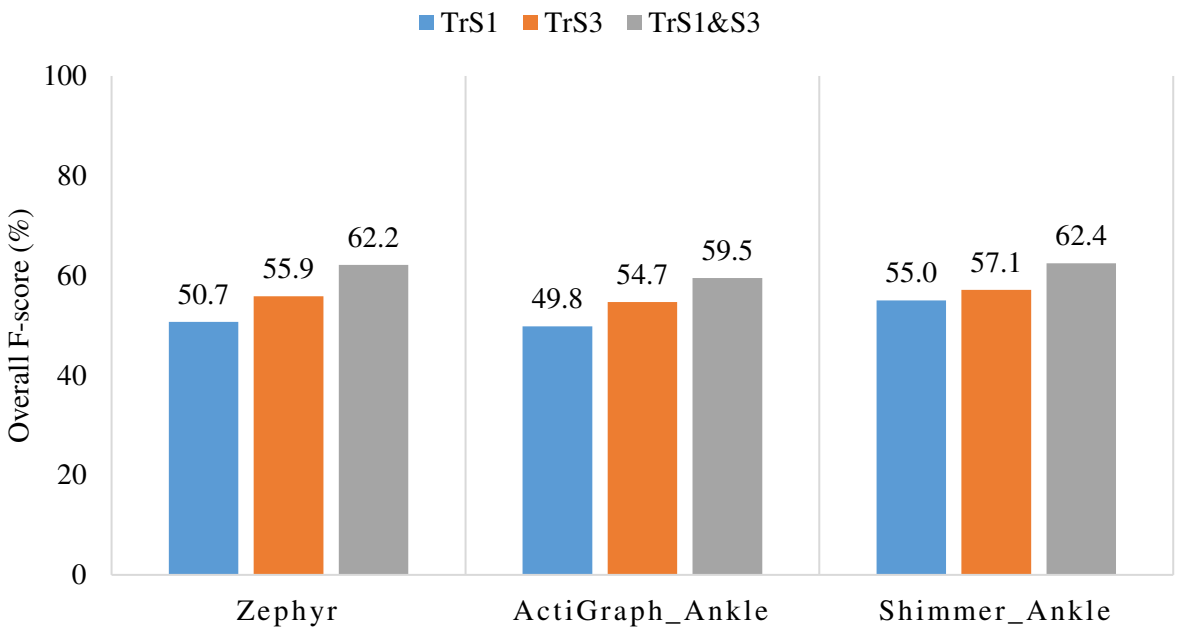


Figure 5.11 - Overall classification accuracy of S3 test-set for Zephyr, ActiGraph_ankle and Shimmer_ankle modules when S1, S3 and S1 with S3 sets are used for training.

5.4 Discussion

To clearly derive our analysis, the discussion is done based on the analyzed aspects, according to the different tests performed during this study:

- Validation of SpD* and SF method on S1 and S3 test-sets:

The first analysis in this study aimed at evaluating the effectiveness of our developed features extraction method (modified SpD* method) on an extended dataset of 20 subjects performing 7 groups of daily life activities. The overall F-score measure, after a leave-one-subject-out validation process, revealed a good performance (between 82.38% and 92.64%) for this method when applied on S1 and S2 test-sets. However, when applied on S3 test-set, the overall performance of our developed method and even that of the literature (SF method) dramatically dropped (-40% approximately), although the S3-sedentary group was accurately identified (84% - 91.4%) (Table 5.7) by all different devices that were tested. **Consequently, the obtained results indicate that both methods are not reliable in detecting six out of the seven categories in real-world scenario.** Apart from the sedentary group, the classification results of the household, stairs, basketball, running, walking and cycling groups were insufficient (between 15% and 66%, lower than 80%). Similar to previous studies, we consider 80% of correct classification as an acceptable accuracy level for the algorithms (Sasaki et al., 2016; Staudenmayer et al., 2009). The lowest classification accuracy marked by the different devices was for the household activity. These findings match those observed in earlier studies (Foerster et al., 1999; Sasaki et al., 2016)

that demonstrated the substantial degradation of the accuracy of the algorithms developed on laboratory data and tested on free-living data. The results of (Foerster et al., 1999) witnessed an increase from 4.5% to 33% of incorrect classification rates when switching from a standard protocol to an ambulatory monitoring (semi-standard protocol). The decays in the recognition performance obtained by (Sasaki et al., 2016) was 40-46% for discriminating between recreational, household, locomotion, sedentary and standing activity groups when performed in free-living environment. In accordance with our present results, the highest accuracy for their algorithm was in identifying the sedentary behavior (86.9%). Another study that was particularly interested in validating the ability of a thigh-accelerometer in detecting the sitting posture in free-living context (Skotte et al., 2012), revealed 98% and 93% of sensitivity and specificity respectively. It can therefore be assumed that the recognition algorithms that are trained on controlled dataset are merely reliable in detecting controlled activities. They are not operational in a real-world context, though, they are able to distinguish the sedentary behavior with a good accuracy. It seems logical to state that these unsatisfactory results are due to the fact that the signals recorded in free-living conditions and those recorded in laboratory conditions are considerably different. This difference can be explained by what is so called: the Hawthorne effect (Parsons, 1974), the PA of the participants are being performed in an unnatural manner in the controlled settings that does not reflect their real behavior in real-world scenario. For these reasons, it is preferable to train the recognition model using data collected under the same conditions in which it will be tested, as it will be discussed later on.

- The effect of window length:

In this study, we also included an exploratory analysis of the window lengths of the observations. We were interested in investigating the effect of varying the duration of the observations on the recognition performance. **As can be observed, our results showed an increase of the classification accuracy of the S1 test-set with the increase of the window length, yet this is not the case with the S3 test-set.** A reasonable explanation of this would be that the controlled datasets involve fixed duration of activities, meaning that a 6-min interval will perfectly match with six clear classification events of a 1-min controlled activity. This accuracy level becomes critical in free-living conditions since the activities in these conditions are not performed for extended periods and are not performed in known durations (Sasaki et al., 2016). Moreover, in free-living behavior, the activities reveal supplementary overlapping characteristics in the data compared to activities performed in restrained conditions.

- Role of semi-standardized session S2:

Our findings concerning the recognition of S2 test-set have shown that the classification accuracy rates are on par with those of S1 test-set. This result can be explained by the fact that after joining the 17 activities into 7 categories, the S2 dataset played the role of a replication of the S1 dataset on a reduced period with no intensity and duration constraints. The intensity and the duration of the S2-activities are

based on the subject preferences, as detailed in Chapter 3 (section 3.4.2.2). Since in this study, the different intensities of the cycling, walking and running activities are grouped in a same category, the training dataset of S1-walking category includes the slow, normal and fast cadency of walking behavior. The performed walking activity by the subject in S2 is thus a recurrence of his S1-walk on one intensity or another. Meaning that in this particular case, we have two measurement days that have almost similar experimental data for a same subject. The S2 dataset would thus serve to examine the repeatability of the classification responses obtained in S1 in another experimental day. By this way, the role of S2-session here is to test whether the recognition accuracy remain constant on different occasions. The results between S1 and S2 were good enough and found to be comparable. The performances of the Zephyr device, as an example, achieved 95.44% and 96.29% for the test-sets of S1 and S2 respectively (KNN classifier/SF method). **This demonstrates the effectiveness of the test-retest reliability of these methods in standardized and semi-standardized contexts.** The results in S2 are slightly better than in S1, a possible explanation of this might be that the number of the observations of S2 test-set is lower than those of S1 test-set, hence reducing the number of misclassifications (rf. Table 5.3). These findings from session S2 are comparable with those of the study in (Bao et al., 2004) since their dataset is collected also in a semi-naturalistic session. They showed that the high classification accuracies they obtained for walking, running, climbing, stairs, standing still, sitting, lying down, working on a computer, bicycling, and vacuuming are as good as the classification results obtained from laboratory recordings.

- Best voted sensor and sensors position:

Another intention from the analysis that we were conducting here is to determine which sensor, variable and position result in the best classification accuracy. The comparison of the overall F-score values of the accelerometer sensors on the S1 test-set revealed that the Zephyr was the finest device (95.4%), after that the Shimmer (91.8%) and then the ActiGraph (89.9%). However, by comparing the overall F-scores of S3 test-set, the Shimmer_ankle marked better accuracy than the Zephyr strap. By considering the three positions of the Shimmer and ActiGraph modules, the ankle position reported better overall accuracy, after that the hip and then the wrist. Nevertheless, by investigating the accuracies function of the activity groups, it can be noticed that the locomotion activity types (walking, cycling and stairs) are better detected by the ankle module, yet the sedentary and household activities are better detected by the hip module. In this study, we also analyzed the effect of coupling two or more accelerometers on the recognition performance. The obtained results showed a performance superior to that achieved by protocols employing a single accelerometer. This is in line with previous works that demonstrated an enhancement of the classification with data fusion from different sensors placement. Several attempts to determine the optimal sensor's positions have been made (Bao et al., 2004; Cleland et al., 2013; Leutheuser et al., 2013). **Consequently, the selection of the sensor's number and placement is**

question of the intended application. A classic example is that of a PAR system that is conceived for the assessment of walking pattern of peripheral arterial disease (PAD) patients, the accelerometer in this case should be placed on the ankle. To study the behavior of Parkinson disease (PD) patients, one sensor on the ankle might not be sufficient and additional sensors on the hip and thigh might be also needed to accurately detect dyskinesia (Jalloul et al., 2015). Yet, increasing the number of sensors couldn't increase the recognition performance of S3 test-set. The ActiGraph modules, for instance, only revealed 1% of increase after adding the wrist and hip modules to the ankle module. **It became clear that the accurate detection of free-living test-sets is not possible regardless of the sensor position or number.**

- Analysis of the gyroscope and magnetometer:

We analyzed the impact of adding the information of the gyroscope and the magnetometer that are already embedded in the Shimmer modules. There exists in literature different works in designing portable magnetic systems combined with inertial sensors for tracking human motion and estimating the orientation of body's segments (Roetenberg et al., 2007; Zhu et al., 2004). Yet, prior studies that have used the magnetometer with the accelerometer in their PAR system, didn't evaluate the contribution of each variable separately (Altun et al., 2010). **Our results showed that coupling the magnetometer to the accelerometer couldn't offer any enhancement in the performance, whereas, the gyroscope added approximately 2% of precision to the recognition.** An expanded analysis of the role of gyroscope in activity recognition can be found in (Shoaib et al., 2013). The authors of this study stated that the gyroscope can complement the accelerometer and can achieve reasonable performance even if taken alone.

- Analysis of the HR and BR:

For the detection of the seven categories, only slight improvement is witnessed by taking the Zephyr's HR or BR variables into consideration in the activity recognition process. **Among the considered activities, ascending/descending stairs was mostly influenced by the HR data that was able to increase its accuracy with up to 6.8%.** (Lara et al., 2012b) made use of the acceleration coupled with 6 vital signs from the Zephyr chest strap: the HR, BR, breath amplitude, skin temperature and ECG amplitude. In accordance with our results, in (Lara et al., 2012b) the 'ascending' activity reported the most significant improvement (between 10% and 13%). (Parkka et al., 2006) also studied HR and respiration in the activity recognition problem and found out that, in their setups, these physiological signals didn't accurately alter the performance. In our study, running, cycling and basketball activities weren't affected by the joint of HR signal, yet a slight decrease was observed compared to the case where acceleration information was taken alone. According to these findings, it can be concluded that the practicality of the physiological variables depends on the types of activities to be recognized. For

instance, if the sedentary and ambulatory types were the only targeted activities, then the acceleration information would deliver sufficient accuracy. On the other hand, if 'going up stairs' is also investigated then the vital signals would certainly yield better performance. We find it interesting to further investigate the impact of the vital signs on identifying the intensity of the activity alternatively to its type. We believe that if our running, walking, and cycling groups were split again between slow, normal and fast intensities (Table 5.2) then the vital signs would provide the PAR system with more reliable outputs in discriminating between these intensities. Besides, there exists many published studies that investigated and proved the strength of the vital signals in the EE estimation field rather than just in the activity recognition field (Butte et al., 2012; Gjoreski et al., 2015).

Taken together, these results suggest that selecting the sensor's variables depend on the purpose of the investigation and the activities under study. The reported findings can be however used as proof of concept in such a way that the developed methodology would be easily exploited in the final implementation of the PAR system in the large range of applications.

- Controlled vs Free-living training-sets:

Finally, in light of previous studies, we applied a recently investigated strategy in the aim of improving the recognition performance of the free-living activities. This strategy is based on not only using controlled data but also including free-living data in the training process of the recognition model. This similar procedure was adopted by (Ermes et al., 2008; Sasaki et al., 2016) in which the manipulated data were issued and compared between two supervised and unsupervised protocols. In the study of (Ermes et al., 2008), unsupervised data were collected from subjects that self-annotated the performed activities on a PDA and the unsupervised data collection in the study of (Sasaki et al., 2016) was based on the direct observation technique. Their results showed similar trends to what was obtained here. **In our study, it was reported that using free-living data in the training process is 7.5% to 11.5% more reliable than using only controlled data.** By considering this same strategy, the study (Ermes et al., 2008) showed an increase of 17% in the performance and (Sasaki et al., 2016) observed an improvement of 9% to 14%. Altogether, these results conclude that in order to be used in real-world applications, the recognition model must be trained on free-living datasets. Yet, the validity of the free-living based recognition model in our study is not sufficient since the overall classification accuracy observed is between 59.5% and 62.4% (lower than the acceptable accuracy level of 80%). Contrary to these findings, the correct classification of free-living activities of (Ermes et al., 2008) reached 90% when supervised and unsupervised data were used in training. We believe however that our free-living data from S3 are more naturalistic and thus more challenging than those collected by (Ermes et al., 2008) which yield better accuracies in their results.

5.5 $\dot{V}_E$ extraction from Zephyr-Breathing Waveform

- Background:

Restating our initial objective in SHERPAM project, which was explained at the beginning of the dissertation: we aim to design a robust system not only to classify daily activities, but also to monitor vital signs of the individual and estimate EE related to each activity. We presented in the previous sections, a number of tests comparing the response of different off-the-shelf modules and studying their validity in activity recognition. We now seek to conceive a system that also renders service with regards to EE. The device that particularly interested us, among the different modules that we tested in ours experiments, is a multi-sensor based module: the Zephyr BioHarness 3 chest strap. This module delivers different physical and physiological signals from multiple sensors integrated in the same chip (refer to section 3.4.1.2). In addition to the acceleration, HR and BR signals, this Zephyr strap offers further possibilities to explore a "breathing waveform" (BW) signal, that we believe will help us to achieve $\dot{V}_E$ and then EE estimation. We have already debated the EE estimation from $\dot{V}_E$ in the review presented in Chapter 2 (section 2.5.2.3.3), in which we discussed the importance of the $\dot{V}_E$ variable and showed the different tools and techniques available in literature to measure it.
For these reasons, and in order to contribute to this domain, we thought of studying this BW signal from the chest strap positioned sensor and evaluate the possibility of extracting from it appropriate measures of pulmonary $\dot{V}_E$ in attempt to estimate EE.
The Zephyr chest strap integrates a pressure sensor that can detect chest expansion and contraction due to the breathing movement. The expansion and contraction are represented by the upward and downward deflections in the sine wave drawn by the BW signal amplitude (Figure 5.12(a)). The principle of the breathing detection by Zephyr can be assimilated to the RIP technology which is used in $\dot{V}_E$ measurement (rf. Chapter 2 (section 2.5.2.3.3)). Although, RIP technology relies on the use of 2 coils placed around the rib cage and the abdomen, yet in our experiments, we seek to study the accuracy of one strap Zephyr placed on the chest. To the best of our knowledge, there is no study that assess the reliability of the Zephyr-BW variable; all reported works that made use of the Zephyr module were interested in studying the HR or BR signals (Bianchi et al., 2013; Chow et al., 2013; Kim et al., 2013). (Johnstone et al., 2012c) determined the reliability and validity of the Zephyr module using a field based protocol where the physiological signals: HR and BR are analyzed. In our study, we were interested in evaluating the validity of the BW and BR signals of Zephyr compared to the outputs of the Cosmed K4b2, as reference.

- Experimental Breathing-Protocol:

In trial 1, we equipped a subject with the Zephyr chest strap and the gas exchange system Cosmed K4b2 and asked the subject to perform controlled breathing cycles of 1 min each. The subject started by

performing controlled breathing following the rhythms of a metronome of 7.5 breaths per minute (bpm) (T_1), then retrieved his normal breath (T_2). Then the subject controlled his breaths again on 7.5 bpm while trying to reach his maximum tidal volume (T_3). After resuming a normal breathing cycle (T_4), the subject then accelerated his breaths to a higher frequency of 15 bpm with a maximum tidal volume (T_5). We repeated this procedure two more times under two different conditions: (i) in Trial 2, we fitted the strap tightly than it is in Trial 1 and (ii) in Trial 3 we slightly lowered the strap. The aim of these tests is to study how the strap tension and its location on the body affect the BW amplitude. The raw unfiltered breathing waveform reported at 25 Hz, along with the Cosmed K4b2 output file were collected at the end of the experiment.

Table 5.11 - Description of the different tests in the experimental breathing protocol during the three trials.

Tests (1min)	Breathing Mode
T_1	Controlled breaths; BR = 7.5bpm.
T_2	Normal breaths
T_3	Controlled breaths; BR = 7.5bpm + VT_{max}.
T_4	Normal breaths
T_5	Controlled breaths; BR = 15bpm + VT_{max}.
T_6	Normal breaths

- Methods:

In order to extract $\dot{V}_E$ measures from the BW signal, we detected the peaks (t_i^{Max},V_i^{Max}) (in green) and (t_i^{Min},V_i^{Min}) (in red) for i = (1…n) with n = number of respiration cycles (see Figure 5.12 (a)), and then computed tidal volume (VT) and total respiratory cycle time (Ttot) signals from which $\dot{V}_E$ can be calculated:

$$VT_i = V_i^{Max} - V_i^{Min}$$

$$Ttot_i = T_i^{Min} - T_{i-1}^{Min}$$

$$VE_i = \frac{VT_i}{Ttot_i}$$

- Analysis:

Figure 5.12(a) illustrates BW signal collected during the 6 breathing modes in three trials. From this figure, the following preliminarily findings can be deduced. First of all, the change in the signal amplitude between the trials can be clearly inspected. When the strap was securely tightened (Trial 1), the range of amplitude was larger. For the same breathing modes, the module was thus detecting a more important movement of the chest in Trial 1 than the case of the loose strap (Trial 2). On the other hand, the displacement of the strap caused a great shift in the DC component of the BW signal. Lowering the

position of the module produced a lowering average of the signal (Trial 3 compared to Trial 2). These first results can show the non- reproducibility of the BW signal, its amplitude is highly sensitive to the tension of the strap and its position on the body.

Secondly, although the relative correlation between VT signals of Zephyr and Cosmed can be visually inspected, yet the correction coefficient is not a constant between the breathing modes within a same trial (Figure 5.12(c)), making it difficult to study $\dot{V}_E$, regardless the units conversion. Another interesting finding is that the BR that we extracted from the BW of Zephyr (BRzephyrE) using Ttot equation, seems to be better than the BR given directly at the output of Zephyr (BRzephyrM), by comparison against the reference values BR of Cosmed (Figure 5.12(d)). In fact, the BRzephyrM value, as indicated in the Zephyr manual, is 'heavily filtered and processed' based on the manufacturer algorithms before being provided. Nevertheless, our results showed that the BRzephyrE that we extracted from the raw unfiltered signal is more reliable than the filtered Zephyr variable and must be adopted instead.

Even though the results that we present are preliminarily, this simple validation experiment would however allow us to infer straight decisions. It can be assumed that in its current state, the BW output of Zephyr cannot be used as indicator of $\dot{V}_E$. Since it is not provided in real physical units and is altered by how tight the strap is fastened and at which position the sensor is adjusted. In addition, we have demonstrated that it is not straightforward to determine a conversion coefficient in order to give this measure a physical sense of ventilation.

Furthermore, the poor quality of VE extraction from BW-Zephyr can be justified by the results obtained with previous studies concerned by RIP technology. It was stated that it is not possible to obtain accurate respiratory volumes with a single-band due to differences in posture and thoraco-abdominal respiratory synchronization (Clarenbach et al., 2005; Konno et al., 1967). Dual-bands modules outperforms single band modules since they provide a complete covering of the breathing mechanics. This statement limits the quantification of the BW variable as $\dot{V}_E$ measure and accordingly, it would be required to add a second strap to the Zephyr in order to produce reliable measures. As such, breathing mechanics cannot be inferred from the BW signal as is, and different approaches must be found to accurately measure $\dot{V}_E$.

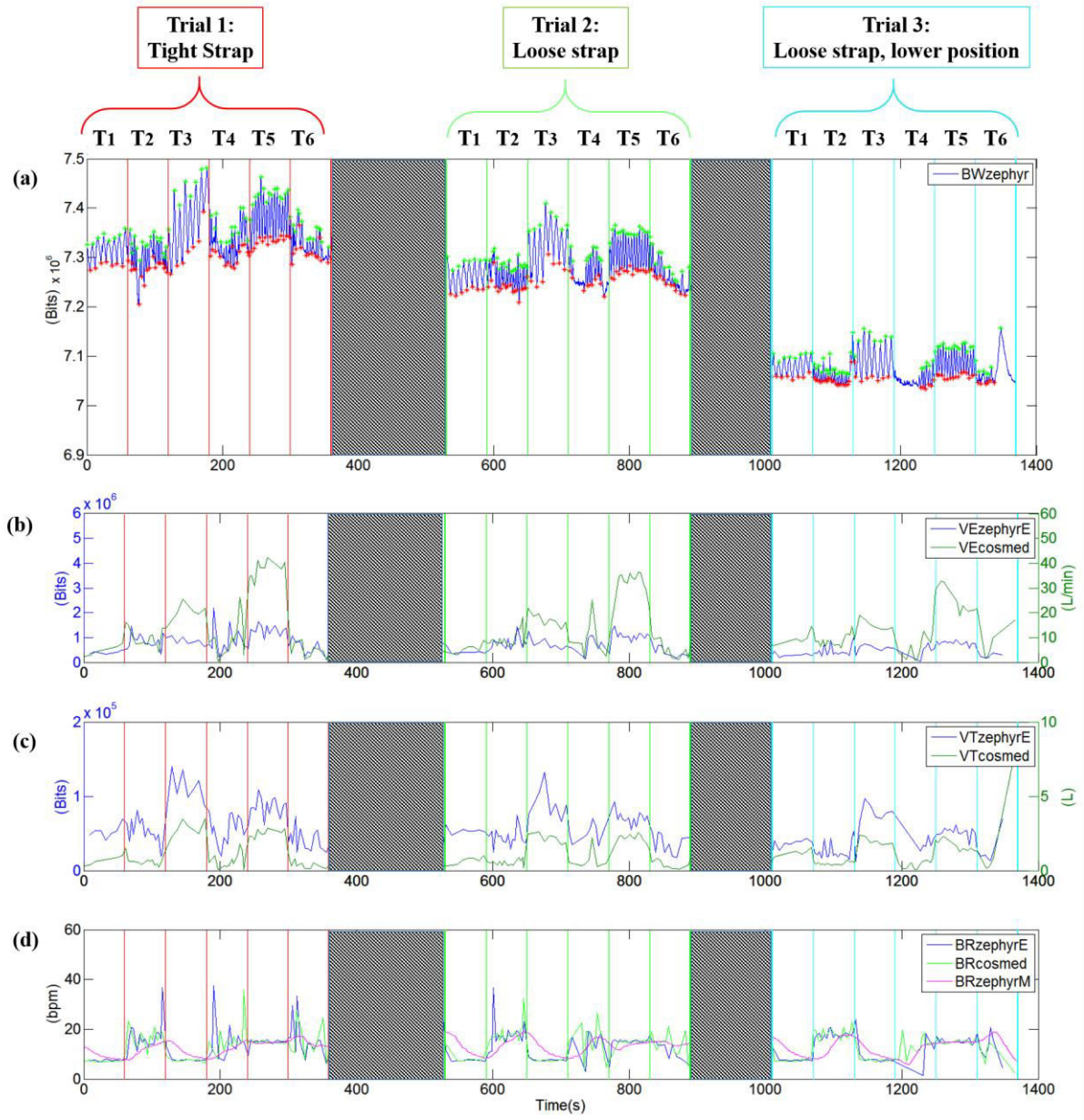


Figure 5.12 - Ventilation ($\dot{V}_E$) extraction process from Zephyr Chest strap.

(a) raw unfiltered breathing waveform (BW) signal of Zephyr, showing the detected peaks. (b) $\dot{V}_E$ extracted from BW zephyr (VEzephyrE) compared to VE given by Cosmed (VEcosmed). (c) Tidal Volume (VT) extracted from zephyr (VEzephyrE) compared to VT given by Cosmed. (d) Breathing Rate (BR) extracted from zephyr (BRzephyrE) compared to BR given by Cosmed and BR given by Zephyr.

5.6 Conclusion

The conducted studies throughout this chapter provided an important opportunity to advance the understanding of PAR systems' performance in real-world conditions. Based on the obtained results, it appears that the transition from controlled to real-life dataset subverted the performance of the recognition models. While both SF and SpD* methods could successfully detect the seven categories (rest, household, stairs, Basket, walk, run and cycle) collected during session S1 (up to 95.4%), the

results on data from session S3 had shown a substantial decrease in the classification rates (46 – 50%). However, once the classifier is trained using the free-living S3 dataset instead of using only the controlled S1 dataset, an improved performance by up to 12% was detected when tested on free-living dataset. Consequently, using only controlled data in learning recognition models wouldn't be sufficient. Collecting dataset under free-living context is necessary to develop a reliable PAR system. Although, some types of activities (such as household, ascending/descending stairs) weren't accurately classified using our developed algorithms, yet an important accuracy level in detecting sedentary behavior in free-living context was obtained. Accordingly, the developed PAR system can accurately draw the user's activity profile by successfully detecting inactivity and activity periods. Once implemented in real-world scenarios, they are able to show the user his current distribution of sedentarity and activities during his day.

Furthermore, we investigated, throughout this study, the different sensors, sensors positions and variables in order to select the optimal skeleton for a reliable PAR system. We demonstrated that this vote depends heavily on the final intended application of the system. For a reliable recognition of a range of locomotion, household and sports activities, a joint of an ankle and a chest-worn accelerometers would be needed. In addition, for a discrimination between walking and ascending/descending stairs, HR information would be effective in this case. In our specific case, the SHERPAM project's monitoring system that is under construction is conceived for implementation in real health care applications (e.g. HF disease). Thereby, it would be important not only to identify the daily life activities but also to monitor the physiological signals (HR, BR, $\dot{V}_E$) of the patient and estimate the EE related to each performed activity. We were thus motivated in selecting a reliable module that integrates both physical and vital sensors. By this means, we suppose that the Zephyr multi-sensor device must serve this matter. In a separated study, we investigated the breathing waveform BW signal of Zephyr. We proposed an estimator of the breathing rate BR from the unfiltered BW signal and validated this measure against the reference device Cosmed K4b2. We also showed the number of challenges that are to be resolved in order to accurately extract $\dot{V}_E$ and then EE measures from a commercially available module. The findings obtained during this study are though preliminary and further analysis are still needed to maintain the relevance of PAR system in free-living context and the measure of $\dot{V}_E$ and EE.

Chapter 6
GENERAL CONCLUSION AND FUTURE PROSPECTS

Despite the fact that the field of human activity recognition has extensively progressed, there remain high demands in raw processing, pattern extraction, classification and performance evaluation under realistic conditions. The main objective of this dissertation is to develop and validate algorithms for the recognition of everyday activities from commercially available wearable sensors. Our work focused on addressing the lack of the existing algorithms in evaluating their performance in real-world scenarios. This chapter is divided in two sections. In the first section, we concluded the most relevant contributions and findings of this dissertation and in the second section, topics for future investigations are proposed and divided into two sub-sections. Firstly, we deployed a number of in-depth analysis to refine the activity recognition problem in light of the pragmatic dataset and in a second phase, we proposed a new perspective towards a better estimation of EE through $\dot{V}_E$ analysis.

6.1 Contributions and findings

In this section, we won't explain over again the derived results from our studies. These results can be found throughout the chapters. We would prefer here to emphasize the main contributions that we provided within the framework of this dissertation:

- We firstly provided a complete review on the activity recognition process, the used experimental tools, the current studies and advances in the field and showed the main gaps in existing PAR algorithms that lead us to carry out our studies.
- We investigated the existing features extraction and features selection methods used in the activity classification algorithms, and explained their limitations in PAR applications. For this reason, we developed and evaluated a new pattern recognition method that extracts spectral distances features from acceleration signals. We re-produced the usual time-and-frequency-domain features in literature as a reference for our proposed set of features and proved the pertinence of our features extraction method. The spectral distances method was able to automatically distinguish between different types of activities and sedentary behaviors while reducing the computation and the complexity of the classification algorithms.
- We studied the impact of physiological signals in the performance of activity recognition by adding HR and BR information to the classification process. We compared our findings with similar studies, and showed the applications in which the use of HR and BR would be essential.
- The literature limitations in developing PAR algorithms using controlled activity datasets without ensuring their validation in free-living context motivated us to conceive a novel and pragmatic data collection protocol. We collected a large and varied dataset based on sophisticated triple-experimental sessions for the purpose of validating commercially available sensors and evaluating the developed recognition algorithms under free-living scenarios. The collected dataset constitutes a potential benchmark in the PAR field, allowing a clearer understanding of human behaviors and disclosing the limitations of previously developed classification algorithms.
- To better visualize and interpret the free-living classification process, we created RACHA, a MATLAB graphical platform which integrates a data visualization interface and a complete panel to design, train and evaluate the activity recognition model.
- We have demonstrated that the algorithms that are trained on a controlled dataset would accurately classify controlled samples but would not perform equally well on free-living samples. When applied on seven categories of activities, this finding was verified using all the different types of sensors. Six categories were correctly classified in the S1-controlled session but not sufficiently in the S3-free-living session. The sedentary behavior was however accurately detected in both the controlled and the free-living context.

- We also demonstrated the effectiveness of the test-retest reliability of the developed methods in standardized and semi-standardized contexts.
- We analyzed the use of free-living training-set in learning the classification model and showed the important enhancement in the algorithms performance in detecting free-living activities compared to the use of controlled training-set.
- We performed different tests that would affect the PAR performance and particularly analyzed the impact of multi-sensor joint and multi-variables sensors on the algorithms' accuracies. We concluded that the choice of sensor's variables, sensor's number and placements on the body is question of the intended application and on the range of activities under study. In order to select the optimal monitoring system, we highlighted the benefits from using commercially available motion devices that integrate various physical and physiological sensors on a single chip (such as Zephyr BioHarness3).

In conclusion, if implemented in real-world scenarios, our algorithms and literature algorithms can accurately detect the sedentary from activity behaviors of the user during his day. Yet, they are, in their current state, not accurate enough for discriminating between different free-living activities opening rooms for further research as it will be discussed in next section.

6.2 Future prospects

6.2.1 Towards a robust PAR system in free-living context

The well-founded protocol that we conceived in our studies still allows other exploitation aspects. We have provided in this dissertation a global understanding of what is expected from a PAR application and proved that a reliable classification model must be funded on free-living datasets collected in real-world settings. Yet, the results reported here are only preliminary and refinements of the recognition models are still required to further improve the accuracy of the results and the recognition model as a whole. The full potential of this protocol would allow to tackle the following aspects:

- User-Specific application:

In our conducted study, the leave-one-subject-out validation process may have limited the accuracy of the classifiers. A robust activity classifier can be created based on user-specific training data. In real-world environment, movement behavior can be different between subjects inducing individual signal variations and thus affecting the activity patterns. This indicates that the classifier must be adapted to the activity profile of each subject and a user-specific classifier based on the user data should avoid inter-subject variability. Prior work has studied the power of subject-dependent versus the subject-independent models and presented the strengths and drawbacks of each approach. (Bao et al., 2004) explained the complication in discriminating between sitting and standing postures using a hip-worn

accelerometer by the fact that the sensor orientation on the waist is influenced by the subject's body shape. The study of (Ermes et al., 2008) that compared pre-trained and user-specific classifiers showed that the accuracy of pre-trained classifiers is limited to specific types of activities such as ambulatory walking and running, whereas, other activities that are subject to large variations between individuals such as stretching, require user-specific training. Even though a generalized recognition model trained once is stronger, yet a PAR system would be more effective in real-world if fitted to each subject's environment. In the final application, a group of reference patterns must thus be anteriorly obtained for each subject and for each specific activity to be investigated.

- Hierarchical recognition process:

Another important point of discussion that we intend to investigate via this protocol is adopting a multi-steps recognition process that is flexible between the activities. In a first step, a classifier can distinguish between large categories of activities then another classifier refine the recognition of similar activities within the same category. An example of a hierarchical classifier application in the literature is presented in the study of (Leutheuser et al., 2013) that distributed 13 daily life activities into groups and subgroups of activities and then a classifier is considered at each level of classification. It is interesting to note that the authors of this study defined a different feature set for the resting group than that for the other groups since the signals characteristics in the two groups differ: the orientation of the body is important for the discrimination of sitting, lying and standing within the resting group but won't add any information for the discrimination of the activities within the walking group. Furthermore, by adequately implementing several variables in the hierarchical recognition process this can reinforce the success of this procedure. For instance, the GPS device (from which data are not investigated in this study but exists in the data collection system) can provide contextual information on whether the activity is performed indoor or outdoor at a first level of detection. HR signal can also tell about the intensity of the activity, and separate between the sedentary group and other groups. On the second level, a more sophisticated classifier can be used next to identify which activity from the same category was performed. As an example from our study, to recognize going up/down stairs activity that was confused most of the time with the walking activity, we propose to examine the atmospheric pressure signal. This signal is given by the altimeter sensor that is integrated in the Shimmer modules along with the accelerometers, the gyroscopes and the magnetometers. The pressure signal, sensitive to the altitude, can easily detect the ascending/descending stairs from the walking activity contrary to the acceleration signal, as presented in Figure 6.1.

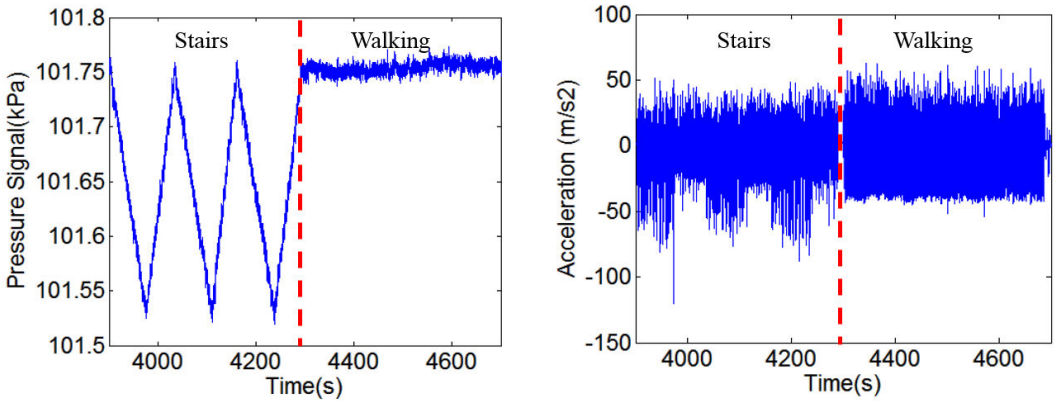


Figure 6.1 - Shimmer-pressure (on the left) and Shimmer-X-axis acceleration (on the right) signals variations in going up/down stairs and walking activities. Each represented activity is performed on 6min. In the going up/down stairs, the subject was ascending and then descending a 5-floors building on 3 consecutive times.

- Hybrid recognition process:

Generally in the classification process, misclassification occurs on a specific number of observations and not on the whole activity. Figure 6.2 shows an example of correctly classified and misclassified observations by the KNN classifier. It can be noticed that on the range of a given activity observations, the misclassifications occurred solely while being neighbored by correctly classified observations. Applying filtering techniques to rectify the misclassified observations by using information from neighboring correctly classified observations within the same activity represents a fruitful alternative. There exists different filtering techniques that might help dealing with misclassified instances: The sequence labeling method, the time smoothing or the calibration process all based on transition probabilities have emerged and were used to support the basic classification algorithms. These approaches require further attention to prove their robustness in the challenging real-world context. In a number of studies, the probabilistic Hidden Markov Models (HMM) were used as a calibration phase in combination with the random forest classifier (RF) or a decision tree, creating by so a hybrid classification process (Kerr et al., 2016; Reddy et al., 2010; Rosenberg et al., 2016; Trabelsi et al., 2013). For instance, (Kerr et al., 2016) used HMM to do time-smoothing at the final stage of the learning process at the output of the RF classifier. In such a case, the algorithms are adapted on the temporal data taking into account the sequential appearance of the activities. The hidden states of the HMM are trained using the probability of transitions between activities in order to produce a final sequence of predicted behaviors (e.g., they can learn that it is more likely that a subject changes from lying to sitting than directly to walking). Although these algorithms might seem complex, yet they were smoothly implemented on mobile phones (Reddy et al., 2010), demonstrating their feasibility in real-world applications.

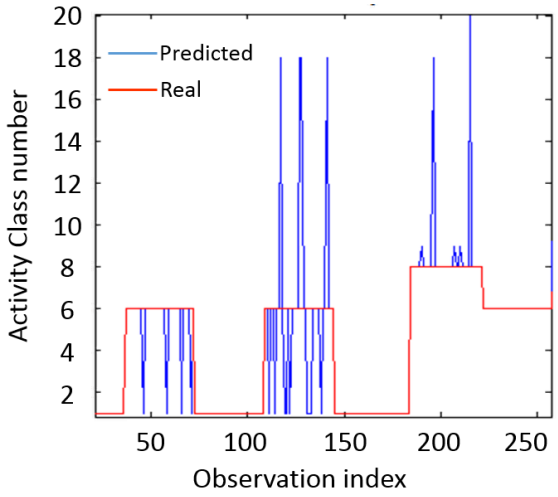


Figure 6.2 - Example of misclassification samples among activities. 4 samples from the activity n°6 (representing household category in our pragmatic protocol) were separately misclassified as n°1 activity (sedentary). If each of these samples was smoothed with its neighbors' samples that were correctly classified, a better performance would be reached.

- Photo captures inspection:

The protocol that we deployed employs a sophisticated observational system for ground truth annotation: the wearable camera for automatic photo captures. A protocol of this nature offers a great opportunity to screen and inspect the source of misclassifications by fetching the corresponding photos. This strategy would allow to identify and better understand the behavior that causes the ambiguity to the classification model. Photos analysis can thus lead to an improved activity recognition application much needed in real-world scenarios.

- EE Estimation:

The activity recognition algorithms, once refined and operational, will subsequently play an important role in the estimation of EE, which is grounded on the use of the recognized activities. There exists over 30 published prediction approaches relating PA to EE (Kozey et al., 2010). An overview of the EE studies was conducted in the Chapter 2 of this dissertation. Although wearable activity monitors proved their ability in measuring EE of some locomotion activities, numerous challenges face the development of equations for activities of daily living with which EE can be accurately predicted. It is thus important to further collect, calibrate and investigate data to update the quality of the prediction equations. The protocol that we presented here offers a great opportunity to validate the existing approaches by providing empirical data from a broad types of activities and range of intensities collected in the field setting outside the laboratory. The criterion measure for EE estimation is ensured by the portable metabolic system, the Cosmed K4b2 that was used in both S1 and S2 sessions of this protocol. By manipulating this available dataset, we intend as well to evaluate our heuristic rules of extracting EE from $\dot{V}_E$ variable, as it will be presented in next section.

- SHERPAM Application:

The PAR work presented here is conducted as part of the SHERPAM project, which is introduced in the beginning of the dissertation. The entire chain of data acquisition, signal processing and pattern recognition constitutes an important building block for the SHERPAM mobile-health platform. The developed algorithms involving PA and EE monitoring will be embedded on the prototype platform, for instance, the SpD features extraction method (presented in Chapter 4) in 2 steps, creating the activity models and then computing the spectral distances between the observations and the models. In the running SHERPAM application, the heavy models creation will be deployed on the server, whereas the distances will be computed locally on the smartphone. Implementation tests of this procedure are still in progress. On the other hand, SHERPAM platform is oriented towards health-care perspectives. It will be addressed to allow cross clinical applications in healthy or disabled populations particularly heart failure (HF) and PAD patients. In this sense, extending the work towards clinical protocols to evaluate its effectiveness on these populations is a further step in our project. In all its explored applications, SHERPAM would be flexible enough to tackle the challenges raised by m-Health. In addition, the issue of acceptance of SHERPAM module, involving the selected chest strap device (Zephyr^{TM}) will be evaluated on both front-end (the users) and back-end (the clinicians). It is essential to understand the usability constraints on the technical, medical and societal aspects by experimenting the platform on the intended population. Eventually, an m-health SHERPAM platform will be made available as an open versatile software and hardware system, once a beta version is reached.

6.2.2 Towards a better estimation of EE: $\dot{V}_E$ analysis perspective

Recent attempts to estimate EE in less cumbersome and more efficient manner are still undergoing. The association of EE and $\dot{V}_E$ had been demonstrated over the last decades where $\dot{V}_E$ measurement is considered a direct reflection of EE (Bernard et al., 1979; Boutellier et al., 1985; Durnin et al., 1955). As detailed in Chapter 2 (section 2.5.2.3.3), $\dot{V}_E$ estimation can be based on the use of magnetic sensor belts placed on the thorax to measure the movements of the rib cage and abdomen (Gastinger et al., 2010a; McCool et al., 1986). Yet, the calibration process relied, up till now, on regression equations and machine learning algorithms. These techniques showed relatively good accuracies in estimating $\dot{V}_E$, however, they are blindly applied: they consist of black boxes created in an incomprehensible way suffering from the lake of transparency[4]. In this context, researchers became unable to find the source of error in the estimation and thus unable to improve the performance. Nevertheless, in one of our studies

[4] **"We know we can put in some inputs and get some outputs, and the networks seem to work wondrously well, but how do we ensure that these magical boxes work for every given set of circumstances?"**
- Quote from Clark Barrett, a computer science researcher at University of Stanford, explaining the problem of machine learning techniques.

that was dedicated for the extraction of $\dot{V}_E$ from a pressure sensor housed in a chest strap (rf. Chapter 5 (section 5.5)), we showed how such approach lacks in achieving the desired outcome. A new technique than can explore this relationship between activity and $\dot{V}_E$/EE under new conditions must be projected. Accordingly, our thoughts are directed towards creating physiological models of $\dot{V}_E$, in which the mechanism of the respiratory system is known and can dynamically affect the output of the model. Furthermore, due to the complexity and nonlinearity of the human respiratory system, further investigations are still expected to better understand and estimate $\dot{V}_E$.

In the intention of finding a physiological model of pulmonary ventilation, we made a brief overview of the literature in this field. Interestingly, endeavors were deployed formulating respiratory system components, modelling lung mechanics, describing gas exchanges and simulating ventilation distribution down the lungs (Hoppensteadt et al., 2012; Luca et al., 2012; Steimle et al., 2009; Steimle et al., 2011). However, the application of such studies concern patients in intensive care units where mechanical ventilation is required. Such models cannot be applied in the SHERPAM project since: i) they are modelled based on the mechanical ventilation as input, ii) they model the ventilation system in supine resting positions stimulating only static mechanics.

For all these reasons, we believe that it is fundamental to create new physiological models for ventilation mechanics (e.g. capacity of lungs and alveoli) that can respond to the movement variation. Relying on existing physiological knowledge in respiratory system (Hoppensteadt et al., 2012), it would be required to conceive a model calibrated depending on subjects characteristics and taking into account the effects of exercise intensity. This approach although promising is yet challenging, exposing a number of issues that must be solved, such as the granularity level of the model and the identification of the pertinent and most sensitive parameters. Previously published studies have so far tackled these issues on other fields (Le Rolle et al., 2016; Ojeda et al., 2016). It is now required for these solutions to be adapted on our proposed approach in order to strengthen the performance in $\dot{V}_E$ and EE estimation.

APPENDIX-A
DETAILS ON PRIOR STUDIES

In this appendix, we depicted a number of the most recent and pertinent works that we found interesting in literature. We highlighted for each study, its aim, the range of activities, the experimental procedure and equipment, the methodology and its main conclusions. This review sheds the light particularly on these reported works since they were concerned by the activity recognition in a real-life context. Which allowed us to conduct our experimental protocols and scientific analysis beyond state-of-the-art.

Study 1: Activity recognition from user-annotated acceleration data (Bao et al., 2004)	
Purpose	Evaluate the accuracy of recognition algorithms with multiple accelerometers on 20 activities using datasets annotated by the participants themselves.
Activities	20 Activities: walking, sitting and relaxing, standing still, watching TV, running, stretching, scrubbing, folding laundry, brushing teeth, climbing stairs, vacuuming, etc.
Subjects	20 subjects from the academic community: 13 males and 7 females. Age Range: 17 to 48 years.
Equipment	5 ADXL210E accelerometers. Placement: thigh, ankle, arm, wrist, hip.
Experiment	Semi-naturalistic data collection: Subjects run an obstacle course composed of a number of activities listed on a worksheet. The subjects performed the activities in the given order. They self-annotated start and end times for activities, and wrote down any relevant note about the activity. Specific activity data collection: Subjects performed, in the laboratory, a random sequence of 20 activities in the given order. They self-annotated the start and end times of each activity.
Features and Methods	Mean, energy, frequency-domain, entropy, and correlation.
Window	Sliding windows of 6.7 seconds with 50% overlap.
Learning	Decision Table, K-Nearest Neighbors (KNN), C4.5 Decision Tree (DT) and Naïve Bayes (NB) classifiers
Flexibility	User-specific and leave-one-subject out training protocols. A comparison between the two approaches was made.
Application	Healthy subjects/ In real world circumstances.
Metrics Evaluation	Accuracy, confusion matrices.

Findings	Recognition accuracy is highest for DT and then KNN classifiers. The thigh-accelerometer was the most accurate in recognizing the set of activities. The leave-one-subject-out validation process reported better results compared to the user-specific. Classification accuracy rates of between 80% to 95% for walking, running, climbing stairs, standing still, sitting, lying down, working on a computer, bicycling, and vacuuming collected outside the laboratory setting are comparable with recognition results using laboratory data.

Study 2: Performance of Activity Classification Algorithms in Free-Living Older Adults (Sasaki et al., 2016)	
Purpose	Develop and test the performance of laboratory-based algorithms and free-living based-algorithms in detecting the activity type in free-living older adults
Activities	Standing, sedentary, locomotion, household, and recreational categories (with finer activities in each category).
Subjects	35 older adults. Age Range: over 65 years old.
Equipment	3 ActiGraph GT3X+. Placement: hip, wrist and ankle.
Experiment	Normal laboratory protocol: The subjects performed one from two activity routines. Each activity on 5 mins. Free-living protocol: A single daytime block of 2–3 h of direct observation (DO) was carried out for each subject. A personal digital assistant (PDA) programmed for continuous focal sampling direct observation (CFS-DO) are used to code the activities performed.
Features and Methods	Time domain features: 10th, 25th, 50th, 75th, and 90th percentiles of acceleration signals, mean acceleration, and standard deviation of acceleration. Frequency-domain features: 10th, 25th, 50th, 75th, 90th percentiles of signal frequency, range of frequency distribution, total signal power, mean frequency, 1st and 2nd dominant frequency, power of 1st and 2nd dominant frequency, dominant frequency between 0.6 and 2.5 Hz (df625), power of df625, entropy, entropy density, and ratio noise/signal.
Window	Sequential windows of 5-30 s, with no detection of activity transitions.
Learning	Support Vector Machine and Random Forest classifiers.
Flexibility	Leave-one-sample-out validation technique.
Application	Healthy subjects/ In real world circumstances.
Metrics Evaluation	Accuracy, kappa, sensitivity, specificity, confusion matrices.
Findings	Laboratory-based algorithms performed poorly in free-living conditions, whereas algorithms developed with free-living accelerometer data improved the activity recognition rates. However, none of the tested algorithms achieved the preset acceptable accuracy level of 80%.

Study 3: Classifiers for Accelerometer-Measured Behaviors in Older Women (Rosenberg et al., 2016)	
Purpose	Develop and test a computational algorithm to detect walking and sedentary time in older adults. The algorithm was developed on data collected across multiple free-living days and validated in a completely independent cohort of older adults that were not involved in the algorithm development phase.
Activities	Sitting, riding in a vehicle, standing still, standing moving, walking/running.
Subjects	Cohort 1: 39 older women. Age range: 55-96 years. Cohort 2: 222 older women. Age range: 67-100 years.
Equipment	1 ActiGraph GT3X+ (hip), SenseCam camera, Qstarz BT1000X GPS.
Experiment	Study 1: Free-living subjects wearing the devices during waking hours. Study 2: participants from cohort 2 completed a 400 meter walk. (This study serves for the validation of walking activity)
Features and Methods	Basic statistical descriptors of the Vector Magnitude (VM): mean, standard deviation, coefficient of variation, minimum, maximum, 25th, 50th, 75th percentile. 1-s lag autocorrelation, correlation between each axis, roll, pitch, and yaw angles of the direction of acceleration. Fast Fourier transform: the resulting power spectrum, Fmax, Pmax, Fmaxband, Pmaxband, entropy, FFT1 – FFT15.
Window	1-minute window.
Learning	Random Forest + Hidden Markov Model.
Flexibility	Leave-one-participant-out cross validation technique.
Application	Healthy subjects/ In real world circumstances.
Metrics Evaluation	Sensitivity, specificity and balanced accuracy (mean of sensitivity and specificity).
Findings	High performance of 82.2% in predicting five behaviors in free-living older women. High sensitivity of 87.9% in identifying walking behavior during a 400 meter walk filed test. The developed algorithms were developed and trained in a separate dataset and then applied to an independent free-living validation dataset.

Study 4: Validity of PALMS GPS scoring of active and passive travel compared with SenseCam (Carlson et al., 2014)	
Purpose	Assess validity of the Personal Activity Location Measurement System (PALMS) for deriving time spent walking/running, bicycling, and in vehicle, using image captures as comparison.
Activities	Walking/running, bicycling, in a vehicle + discriminating between indoor or outdoor.
Subjects	40 participants (cyclists). Mean age: 36 years.
Equipment	Qstarz BT-Q1000XT. Placement: hip + SenseCam camera around the neck.
Experiment	Recordings during waking hours for 3–5 days, including some weekend days.

Features and Methods	Speed of displacement (km/h), level of vertical acceleration (classified from 0.5–9), and step frequency (steps/min)
Window	2-mins resolution.
Learning	PALMS + threshold cut-point
Flexibility	subject independent (all data merged)
Application	Healthy subjects (active commuters and cyclists)/ In real world circumstances.
Metrics Evaluation	Confusion matrices, IntraClass Correlation Coefficients (ICC).
Findings	PALMs algorithms, which were validated against SenseCam image captures, showed an accuracy of 93% in detecting trips, comparable to widely used cut-point algorithms. Mode-specific agreement between PALMS and SenseCam was over 70% for vehicle time and bicycling, but lower for walking/running.

Study 5: Assessment of Physical Activity and Energy Expenditure by GPS Combined With Accelerometry in Real-Life Conditions. (Nguyen et al., 2013)	
Purpose	Develop a model to assess types and durations of common PA and EE in free-living conditions.
Activities	Burst walking (3 times walking for 30 sec followed by 30 sec sitting), preferred speed walking at 10% slope (6 min), preferred speed walking at –10% slope (5 min), walking slowly on the level (5 min), walking on the level at their preferred speed (5 min), above their preferred speed (5 min), 7) running at a moderate pace (5 min), and cycling 2 laps of a horizontal circuit at a moderate speed (8–12 min).
Subjects	41 healthy subjects. Age range: 25 ± 4.
Equipment	2 accelerometers and a GPS on: waist, ankle and shoulder respectively and a portable indirect calorimeter (Metamax 3B, Cortex, Germany).
Experiment	Protocol 1: The subjects first performed a calibration trial in a circuit outdoors, under supervision. Then they performed 8 different supervised and timed PA consisting of walking, running and cycling bouts (as listed above). Protocol 2: The subjects were asked to perform randomly and spontaneously 5 different activities indoors or outdoors over 2 hours: walking (outdoor), shopping, running (outdoor), watching TV/Desk working (indoor), and lying (outdoor or indoor).
Features and Methods	Speed of displacement (km/h), level of vertical acceleration (classified from 0.5–9) and the step frequency (steps/min), constant threshold values, the lowest and highest SF, speed, and levels of intensity for each PA category.
Window	Minute-by-minute basis.
Learning	NS
Flexibility	NS
Application	Healthy subjects/ In real world circumstances.

Metrics Evaluation	CI, p-value.
Findings	All types of PA, performed during protocol 1, were correctly classified into the PA categories. Predicted duration for burst walking was 100% correct. Predicted durations for uphill, downhill, level walking and running were lower than the actual duration with less than 1 min. Mean error in estimating EE was 15%–25% depending on the type of PA.

Study 6: Using the SenseCam to Improve Classifications of Sedentary Behavior in Free-Living Settings (Kerr et al., 2013)	
Purpose	Assess the validity of 100cpm cut-point technique in identifying sedentary behaviors by using the SenseCam camera as reference.
Activities	Sedentary behaviors: TV watching, administrative activity, eating, other screen use, riding in car, self-care, manual labor, etc. (Classify different types of activities as sedentary or no)
Subjects	40 participants (cyclists), mean age: 36 years.
Equipment	Qstarz BT-Q1000XT. Placement: hip + SenseCam camera around the neck.
Experiment	Recordings during waking hours for 3–5 days, including some weekend days.
Features and Methods	Minutes in each behavior, minutes under the 100-cpm threshold, and mean counts in each behavior type.
Window	Minute level.
Learning	Threshold-based learning: 100 cpm accelerometer cutpoint.
Flexibility	NS
Application	Healthy subjects (active commuters and cyclists)/ In real world circumstances.
Metrics Evaluation	Sensitivity and specificity.
Findings	The results suggest that the difference in daily estimates of sedentary behaviors was 30minutes. For many sitting behaviors, such as non-TV screen time, administrative activity, eating, and watching TV, an accelerometer cutpoint of 100 cpm is 90% accurate. Classification accuracies were better in the field than in the laboratory when compared to the 50% of misclassifications in previous studies conducted in laboratory.

Study 7: Detection of Daily Activities and Sports With Wearable Sensors in Controlled and Uncontrolled Conditions. (Ermes et al., 2008)	
Purpose	Assess the recognition of daily activities and sports performed by the subjects in unsupervised settings compared to supervised settings.
Activities	Lying down, sitting and standing, walking, running, cycling with an exercise bike, rowing with a rowing machine, playing football, Nordic walking, and cycling with a regular bike.
Subjects	12 subjects: 10 males and 2 females. Age Range: 27.1 ± 9.2

Equipment	3-D ADXL202 accelerometers. Placement: Hip and wrist. Garmin eTrex Venture GPS receiver.
Experiment	Supervised session: A supervisor accompanied the subjects and used a (PDA) to annotate the type and context of activities. Unsupervised session: The subject used the PDA to annotate the activities by himself. They were left to freely choose what they want to do: go to work, attend lectures, go home and take a nap, perform different activities such as bowling, driving a car, walking to different places like library, cottage, etc.
Features and Methods	Mean, variance, median, skew, kurtosis, 25% percentile, and 75% percentile, estimation of power of the frequency peak and signal power in different frequency bands. Speed: GPS location data.
Window	1-sec windows.
Learning	Custom decision tree, automatically generated decision tree, artificial neural Network, and hybrid model: combining a tree structure containing a priori knowledge and artificial neural networks.
Flexibility	Leave-one-subject-out cross validation.
Application	Young healthy subjects/ In real world circumstances.
Metrics Evaluation	Accuracy and confusion matrices.
Findings	The total accuracy of the activity recognition using both supervised and unsupervised data was 89%. The accuracy decreased by 17% when only supervised data were used for training and only unsupervised data for validation. The hybrid model classifier provided better results than the reference classifiers.

Study 8: Detection of Posture and motion by accelerometry: a validation study in ambulatory monitoring. (Foerster et al., 1999)	
Purpose	Validate the accelerometry assessment against behavior observation and to examine the retest reliability.
Activities	Sitting, standing, lying supine, sitting and talking, sitting and operating PC keyboard, walking, stairs up, stairs down, cycling.
Subjects	24 male university students. Age range: 21±34 years.
Equipment	4 accelerometers (IC Sensor Model 3031). Placement: the sternum, wrist, thigh, lower leg. Throat micro on larynx. Not used in this study: sensor for the vertical movement of the head on the left ear and ECG sensor for heart rate measures.
Experiment	A standard data collection protocol was performed in the laboratory and repeated twice. A free-living data collection protocol was performed outside the laboratory: Participants, accompanied with the observer performed various activities on a 50-mintues duration. They were free to choose the types of activities (e.g. coffee shop, cafeteria, library, reading a newspaper, conversation, etc.) yet the observer was suggesting some kinds of activities in order to obtain a wide range of postures and motions (e.g. upstairs, downstairs, sitting, and lying).
Features and Methods	Distances of Halmos for AC and DC components.

Window	20-secs windows.
Learning	Distances-based classification.
Flexibility	NS
Application	Young healthy subjects/ In real world circumstances.
Metrics Evaluation	Misclassification rate and confusion matrices.
Findings	The 9 activities were highly detected in the standard protocol yet in the free-living protocol, the misclassification rate was 33%. The reduction from nine to five classes led to 4.7% misclassifications. Speech activity, although a promising parameter in behavior assessment, appeared to be less valid in this present study.

Study 9: Detection of Physical Activity Types Using Triaxial Accelerometers (Skotte et al., 2012)	
Purpose	Validate thigh-accelerometer in detecting sitting posture during free living by comparison with recordings of pressure sensor worn in the hip pockets. Evaluate the recognition of everyday activity types recorded in a controlled protocol.
Activities	Walking, running, cycling, walking stairs, sitting, and standing still.
Subjects	17 subjects: 10 females and 7 males. Age range: 34 ± 11 years.
Equipment	GT3X+ accelerometers: Placement: hip and thigh + a pressure sensor in the hip pocket.
Experiment	Standardized 30-minutes protocol: Walking, running, and cycling carried outdoors in 2 self-paced speeds: "moderate" and "brisk". Sitting activity was performed by sitting on an ordinary office chair in a computer workplace, and for the standing activity, the subjects were asked to stand still. 9-hour protocol (for detection of sitting posture during free living): Subjects were instructed to carry out their everyday life during a 9-hour measurement period, which included working hours (mainly office work) and off-duty hours, including travel time on the way home from work. The subjects were asked to fill out a diary, specifying the start and end of working hours, travel time and time spent lying.
Features and Methods	Mean acceleration, standard deviation, inclination of x-axis, forward/backward angle θ of the thigh (to discriminate between walking and walking stairs).
Window	2-secs windows with 50% overlap.
Learning	Custom decision tree, manually developed.
Flexibility	NS
Application	Healthy subjects/ In real world circumstances.
Metrics Evaluation	Specificity and sensitivity.

Findings	Sensitivity for the detection of physical activity types in the standardized trials were 99%–100% and 95% for walking stairs. Specificity was higher than 99% for all activities. During free living, sensitivity and specificity for detection of sitting posture were 98% and 93%, respectively.

Study 10: Optimal Placement of Accelerometers for the Detection of Everyday Activities (Cleland et al., 2013)	
Purpose	Determine the best position on the body for accelerometers in detecting a range of everyday activities.
Activities	Walking, running on a motorized treadmill, sitting, lying, standing and walking up and down stairs.
Subjects	8 males. Age Range: 24 to 33.
Equipment	6 modules of Shimmer 3. Placement: chest, lower back, hip, thigh, wrist and foot.
Experiment	Duration of each activity was 2 minutes with the exception of climbing stairs, which was carried out on 10 flights of stairs (~80 steps) and repeated, after a one minute pause. For treadmill activities, the subjects self-selected the speed. Data were labeled by an observer.
Features and Methods	Mean, average mean over 3 axes, standard deviation, average standard deviation, skewness, average skewness, kurtosis, average kurtosis, energy, average Energy, correlations.
Window	512 samples per window with 256 samples overlapping.
Learning	C4.5 Decision Tree, Naïve Bayes, Multilayer Perceptron and Support Vector Machine.
Flexibility	10-fold cross validation.
Application	Healthy subjects/ In real world circumstances.
Metrics Evaluation	F-measure, accuracy.
Findings	The Hip-accelerometer delivered the best accuracy among the different tested positions. The accuracy in detecting finer activities was improved by considering multiple accelerometers.

Study 11: Hierarchical, Multi-Sensor Based Classification of Daily Life Activities: Comparison with State-of-the-Art Algorithms Using a Benchmark Dataset. (Leutheuser et al., 2013)	
Purpose	Provide a publicly available dataset of daily life activities (DLA) to be used as a benchmark for new algorithms in the future. Propose a hierarchical multi-sensor based classification system for DLAs which was compared to existing systems.
Activities	13 activities: postures (sitting, lying, standing), household activities (washing dishes, vacuuming, sweeping), walking behaviors (normal walking, treadmill running, ascending stairs, descending stairs), and sports activities (bicycling on ergometer 50W and 100W, rope jumping).
Subjects	23 subjects: 10 female and 13 male. Age range: 27±7 years.
Equipment	4 modules Shimmer 3. Placement: wrist, hip, chest and ankle.

Experiment	A researcher accompanied the subject during the whole data acquisition. A labeling application on the mobile phone was used to label start time and end time of single activities concurrently to data collection. Durations of the activities: 1 or 2 mins.
Features and Methods	For the classification systems BASE, HOUSE, WALK, and BICYCLE: 152 features were calculated: min, min, mean, variance, spectral centroid, bandwidth, energy. Computed for each of the accelerometer and gyroscope axis. For the classification system REST: Gravitational component of the acceleration signal computed by a 3rd order elliptic low pass filter with an infinite impulse response and a cut-off frequency at 0.25 Hz.
Window	5s-windows with 50% overlap.
Learning	AdaBoost (ADA), classification and regression tree (CART), k-Nearest Neighbor and Support Vector Machine with a RBF kernel.
Flexibility	Leave-one-subject out validation technique.
Application	Healthy subjects/ In real world circumstances.
Metrics Evaluation	Mean class dependent classification rate, the overall mean classification rate and confusion matrices.
Findings	The proposed method reached the overall mean classification rate of 89.6%. By using more sensors more complex activities can be classified.

Study 12: Centinela: A human activity recognition system based on acceleration and vital sign data (Lara et al., 2012b)	
Purpose	Propose a model for activity recognition using acceleration and vital signs and evaluate the accuracy of the algorithms under different parameter configurations.
Activities	running, walking, sitting, ascending, and descending
Subjects	8 individuals: 7 males and 1 female. Age Range: 24 (Min: 9 years/ max: 34 years).
Equipment	Zephyr BioHarness BT chest sensor strap.
Experiment	Naturalistic fashion: no specific instructions were given to the participants. The speed, intensity, gait, and other environmental conditions were arbitrarily chosen by the subjects.
Features and Methods	24 statistical features from 3-acceleration signals: Mean, variance, standard deviation, correlation between axes, interquartile range, mean absolute deviation, root mean square and energy. 54 structural features from 6 vital signs: 9 coefficients from the polynomials of degree one, two, and three 12 transient features from 6 vital signs: trend, and magnitude of change
Window	5s, 12s, and 20s window sizes.
Learning	Eight classification algorithms: Naïve Bayes (NB), Bayesian Network (BN), C4.5 decision tree (DT), Multilayer perceptron, additive logistic regression (ALR), bagging ensemble of 8 NB, bagging ensemble from 8 BN, and bagging ensemble of 10 DT.
Flexibility	5x10 fold cross validation
Application	Healthy subjects.

Metrics Evaluation	Average accuracies, per-class mean percentage accuracy, confusion matrices.
Findings	Highest mean accuracy achieved was 95.7% for the ALR algorithm with a window size of 12s and considering both vital signs and acceleration data. Activities of running, sitting and ascending are improved whereas descending and walking weren't influenced when using vital signs in the recognition process.

REFERENCES

ActiGraph. What are counts? : 2016. Available from: https://actigraph.desk.com/customer/portal/articles/2515580.

Aggarwal, J.K., Ryoo, M.S., 2011. Human activity analysis: A review. ACM Computing Surveys (CSUR). 43**,** 16.

Ainslie, P.N., Reilly, T., Westerterp, K.R., 2003. Estimating human energy expenditure. Sports medicine. 33**,** 683-698.

Ainsworth, B.E., et al., 1993. Compendium of physical activities: classification of energy costs of human physical activities. Medicine & Science in Sports & Exercise. 71-80.

Ali, A., King, R.C., Yang, G.-Z., 2008. Semi-supervised segmentation for activity recognition with Multiple Eigenspaces. In: Medical Devices and Biosensors, 2008. ISSS-MDBS 2008. 5th International Summer School and Symposium on. Vol., ed.^eds. IEEE, pp. 314-317.

Altun, K., Barshan, B., Tunçel, O., 2010. Comparative study on classifying human activities with miniature inertial and magnetic sensors. Pattern Recognition. 43**,** 3605-3620.

Anliker, U., et al., 2004. AMON: a wearable multiparameter medical monitoring and alert system. IEEE Transactions on information technology in Biomedicine. 8**,** 415-427.

Atwater, W.O., Benedict, F.G., 1905. A respiration calorimeter. Washington: Carnegie Institute of Washington, Pub. 2.

Bao, L., Intille, S.S., 2004. Activity recognition from user-annotated acceleration data. In: Pervasive computing. Vol., ed.^eds. Springer, pp. 1-17.

Bao, M.-H., 2000. Micro mechanical transducers: pressure sensors, accelerometers and gyroscopes, Vol. 8, Elsevier.

Basseville, M., 1989. Distance measures for signal processing and pattern recognition. Signal processing. 18**,** 349-369.

Bernard, T.E., Kamon, E., Franklin, B.A., 1979. Estimation of oxygen consumption from pulmonary ventilation during exercise. Human Factors: The Journal of the Human Factors and Ergonomics Society. 21**,** 417-421.

Bianchi, W., et al., 2013. Revitalizing a vital sign: improving detection of tachypnea at primary triage. Annals of emergency medicine. 61**,** 37-43.

Bonomi, A., Westerterp, K., 2012. Advances in physical activity monitoring and lifestyle interventions in obesity: a review. International Journal of Obesity. 36**,** 167-177.

Bonomi, A.G., et al., 2009. Improving assessment of daily energy expenditure by identifying types of physical activity with a single accelerometer. Journal of Applied Physiology. 107**,** 655-661.

Bourke, A.K., et al., 2017. A Physical Activity Reference Data-Set Recorded from Older Adults Using Body-Worn Inertial Sensors and Video Technology—The ADAPT Study Data-Set. Sensors. 17**,** 559.

Boutellier, U., Arieli, R., Farhi, L.E., 1985. Ventilation and CO 2 response during+ Gz acceleration. Respiration physiology. 62**,** 141-151.

Bouten, C.V., et al., 1997. A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. Biomedical Engineering, IEEE Transactions on. 44**,** 136-147.

Bussmann, J., et al., 2001. Measuring daily behavior using ambulatory accelerometry: the Activity Monitor. Behavior Research Methods, Instruments, & Computers. 33**,** 349-356.

Butte, N.F., Ekelund, U., Westerterp, K.R., 2012. Assessing physical activity using wearable monitors: measures of physical activity. Medicine and science in sports and exercise. 44**,** S5-12.

Carlson, J.A., et al., 2014. Validity of PALMS GPS Scoring of Active and Passive Travel Compared with SenseCam. Med. Sci. Sports Exerc.

Celka, P., et al., 2005. Wearable biosensing: signal processing and communication architectures issues. Journal of Telecommunications and Information Technology. 90-104.

Chen, K.Y., Bassett, D.R., 2005. The technology of accelerometry-based activity monitors: current and future. Medicine and science in sports and exercise. 37**,** S490.

Chen, Y.-P., et al., 2008. Online classifier construction algorithm for human activity detection using a tri-axial accelerometer. Applied Mathematics and Computation. 205**,** 849-860.

Chow, L., Bambos, N., 2013. Real-time physiological stream processing for health monitoring services. In: e-Health Networking, Applications & Services (Healthcom), 2013 IEEE 15th International Conference on. Vol., ed.^eds. IEEE, pp. 611-616.

Clarenbach, C.F., et al., 2005. Monitoring of ventilation during exercise by a portable respiratory inductive plethysmograph. CHEST Journal. 128**,** 1282-1290.

Cleland, I., et al., 2013. Optimal Placement of Accelerometers for the Detection of Everyday Activities. Sensors. 13**,** 9183-9200.

CominLabs. SHERPAM: 2014. Available from: www.sherpam.cominlabs.ueb.eu.

Corbin-Jallow, S., Moore, R., 1999. Projections of Expenditures for Long-Term Care Services for the Elderly. Vol., ed.^eds. CONGRESSIONAL BUDGET OFFICE (US CONGRESS) WASHINGTON DC.

Crouter, S.E., Churilla, J.R., Bassett Jr, D.R., 2006a. Estimating energy expenditure using accelerometers. European journal of applied physiology. 98**,** 601-612.

Crouter, S.E., Clowers, K.G., Bassett, D.R., 2006b. A novel method for using accelerometer data to predict energy expenditure. Journal of applied physiology. 100**,** 1324-1331.

Cvetkovic, B., Milic, R., Lustrek, M., 2015. Estimating Energy Expenditure with Multiple Models using Different Wearable Sensors.
de Müllenheim, P.-Y., et al., 2017. Using GPS, accelerometry and heart rate to predict outdoor graded walking energy expenditure. Journal of Science and Medicine in Sport.
De Vries, S.I., et al., 2011. Evaluation of neural networks to identify types of activity using accelerometers. Med Sci Sports Exerc. 43**,** 101-7.
DeVaul, R., Dunn, S., 2008. Real-Time Motion Classification for Wearable Computing Applications, Project paper (2001). Vol., ed.^eds.
Doheny, E.P., et al., 2013. Falls classification using tri-axial accelerometers during the five-times-sit-to-stand test. Gait & posture. 38**,** 1021-1025.
Doherty, A.R., et al., 2013. Using wearable cameras to categorise type and context of accelerometer-identified episodes of physical activity. Int J Behav Nutr Phys Act. 10**,** 1186.
Druyan, M.E., et al., 2012. Clinical guidelines for the use of parenteral and enteral nutrition in adult and pediatric patients: applying the GRADE system to development of ASPEN clinical guidelines. Journal of Parenteral and Enteral Nutrition. 36**,** 77-80.
Dugas, L.R., et al., 2005. A novel energy expenditure prediction equation for intermittent physical activity. Medicine and science in sports and exercise. 37**,** 2154.
Dugdill, L., Crone, D., Murphy, R., 2009. Physical activity and health promotion: evidence-based approaches to practice, Vol., John Wiley & Sons.
Dumond, R., et al., 2017. Estimation of respiratory volume from thoracoabdominal breathing distances: comparison of two models of machine learning. European Journal of Applied Physiology. 1-23.
Dunnewold, R., Jacobi, C., Van Hilten, J., 1997. Quantitative assessment of bradykinesia in patients with Parkinson's disease. Journal of neuroscience methods. 74**,** 107-112.
Durnin, J., Edwards, R., 1955. Pulmonary ventilation as an index of energy expenditure. Quarterly journal of experimental physiology and cognate medical sciences. 40**,** 370-377.
Ellis, K., et al., 2014. Identifying active travel behaviors in challenging environments using GPS, accelerometers, and machine learning algorithms. Public Health. 2**,** 39-46.
Ermes, M., et al., 2008. Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. IEEE transactions on information technology in biomedicine. 12**,** 20-26.
Fang, H., Srinivasan, R., Cook, D.J., 2012. Feature selections for human activity recognition in smart home environments. Int. J. Innov. Comput. Inf. Control. 8**,** 3525-3535.
Farkas, I., Doran, E., 2011. Activity Recognition from Acceleration Data Collected With a Tri-Axial Accelerometer. Lucrare publicata la Acta Tehnica Napocensis Electronics and Telecommunication, vol52. 38-43.
Figo, D., et al., 2010. Preprocessing techniques for context recognition from accelerometer data. Personal and Ubiquitous Computing. 14**,** 645-662.
Foerster, F., Smeja, M., Fahrenberg, J., 1999. Detection of posture and motion by accelerometry: a validation study in ambulatory monitoring. Computers in Human Behavior. 15**,** 571-583.
Franke, J., Halim, S., 2007. Wild bootstrap tests for comparing signals and images, Vol., Fraunhofer-Institut für Techno-und Wirtschaftsmathematik, Fraunhofer (ITWM).
Freedson, P.S., Melanson, E., Sirard, J., 1998. Calibration of the Computer Science and Applications, Inc. accelerometer. Medicine and science in sports and exercise. 30**,** 777-781.
Freedson, P.S., Miller, K., 2000. Objective monitoring of physical activity using motion sensors and heart rate. Research quarterly for exercise and sport. 71**,** 21-29.
Fruin, M.L., Rankin, J.W., 2004. Validity of a multi-sensor armband in estimating rest and exercise energy expenditure. Medicine and science in sports and exercise. 36**,** 1063-1069.
Gastinger, S., et al., 2010a. Estimates of ventilation from measurements of rib cage and abdominal distances: a portable device. European journal of applied physiology. 109**,** 1179-1189.
Gastinger, S., et al., 2010b. A comparison between ventilation and heart rate as indicator of oxygen uptake during different intensities of exercise. Journal of sports science & medicine. 9**,** 110.
Gastinger, S., et al., 2014. A review of the evidence for the use of ventilation as a surrogate measure of energy expenditure. Journal of Parenteral and Enteral Nutrition. 38**,** 926-938.
Gjoreski, H., et al., 2015. Context-based ensemble method for human energy expenditure estimation. Applied Soft Computing.
Gradl, S., et al., 2012. Real-time ECG monitoring and arrhythmia detection using Android-based mobile devices. In: Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE. Vol., ed.^eds. IEEE, pp. 2452-2455.
Grant, P.M., et al., 2006. The validation of a novel activity monitor in the measurement of posture and motion during everyday activities. British journal of sports medicine. 40**,** 992-997.
Hall, M., et al., 2009. The WEKA data mining software: an update. ACM SIGKDD explorations newsletter. 11**,** 10-18.
Helmrich, S.P., et al., 1991. Physical activity and reduced occurrence of non-insulin-dependent diabetes mellitus. New England journal of medicine. 325**,** 147-152.
Hoppensteadt, F.C., Peskin, C.S., 2012. Modeling and simulation in medicine and the life sciences, Vol. 10, Springer Science & Business Media.
Hu, F.B., et al., 2004. Adiposity as compared with physical activity in predicting mortality among women. New England Journal of Medicine. 351**,** 2694-2703.
Hu, M., et al., 2016. Refining Time-Activity Classification of Human Subjects Using the Global Positioning System. PloS one. 11**,** e0148875.

Huss, A., et al., 2014. Using GPS-derived speed patterns for recognition of transport modes in adults. International journal of health geographics. 13**,** 40.
Huynh, T., Schiele, B., 2006. Towards less supervision in activity recognition from wearable sensors. In: Wearable Computers, 2006 10th IEEE International Symposium on. Vol., ed.^eds. IEEE, pp. 3-10.
Jalloul, N., et al., 2015. Detection of Levodopa Induced Dyskinesia in Parkinson's Disease patients based on activity classification. In: Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE. Vol., ed.^eds. IEEE, pp. 5134-5137.
Jatobá, L.C., et al., 2008. Context-aware mobile health monitoring: Evaluation of different pattern recognition methods for classification of physical activity. In: Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE. Vol., ed.^eds. IEEE, pp. 5250-5253.
Jia, Y., 2009. Diatetic and exercise therapy against diabetes mellitus. In: Intelligent Networks and Intelligent Systems, 2009. ICINIS'09. Second International Conference on. Vol., ed.^eds. IEEE, pp. 693-696.
Johnstone, J.A., et al., 2012a. Bioharness™ multivariable monitoring device: part. II: reliability. Journal of sports science & medicine. 11**,** 409.
Johnstone, J.A., et al., 2012b. BioHarness™ multivariable monitoring device: part. I: validity. Journal of sports science & medicine. 11**,** 400.
Johnstone, J.A., et al., 2012c. Field based reliability and validity of the BioHarness™ multivariable monitoring device. Journal of sports science & medicine. 11**,** 643.
Jovanov, E., et al., 2005. A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation. Journal of NeuroEngineering and rehabilitation. 2**,** 6.
Kate, R.J., et al., 2016. Comparative evaluation of features and techniques for identifying activity type and estimating energy cost from accelerometer data. Physiological measurement. 37**,** 360.
Katz, S., et al., 1970. Progress in development of the index of ADL. The gerontologist. 10**,** 20-30.
Kelly, P., et al., 2013. An ethical framework for automated, wearable cameras in health behavior research. American journal of preventive medicine. 44**,** 314-319.
Kerr, J., et al., 2013. Using the SenseCam to improve classifications of sedentary behavior in free-living settings. American journal of preventive medicine. 44**,** 290-296.
Kerr, J., et al., 2016. Objective assessment of physical activity: classifiers for public health. Medicine and science in sports and exercise. 48**,** 951.
Khusainov, R., et al., 2013. Real-time human ambulation, activity, and physiological monitoring: Taxonomy of issues, techniques, applications, challenges and limitations. Sensors. 13**,** 12852-12902.
Kim, J.-H., et al., 2013. Measurement accuracy of heart rate and respiratory rate during graded exercise and sustained exercise in the heat using the Zephyr BioHarnessTM. International journal of sports medicine. 34**,** 497-501.
Konno, K., Mead, J., 1967. Measurement of the separate volume changes of rib cage and abdomen during breathing. Journal of applied physiology. 22**,** 407-422.
Kozey, S.L., et al., 2010. Accelerometer output and MET values of common physical activities. Medicine and science in sports and exercise. 42**,** 1776.
Krumm, J., Horvitz, E., 2006. Predestination: Inferring destinations from partial trajectories. UbiComp 2006: Ubiquitous Computing. 243-260.
Kuncheva, L.I., 2004. Combining pattern classifiers: methods and algorithms, Vol., John Wiley & Sons.
Kunze, K., et al., 2006. Towards recognizing tai chi-an initial experiment using wearable sensors. In: Applied Wearable Computing (IFAWC), 2006 3rd International Forum on. Vol., ed.^eds. VDE, pp. 1-6.
Lagerros, Y.T., Lagiou, P., 2007. Assessment of physical activity and energy expenditure in epidemiological research of chronic diseases. European journal of epidemiology. 22**,** 353-362.
Laporte, R.E., Montoye, H.J., Caspersen, C.J., 1985. Assessment of physical activity in epidemiologic research: problems and prospects. Public health reports. 100**,** 131.
Lara, O.D., Labrador, M.A., 2012a. A mobile platform for real-time human activity recognition. In: Consumer Communications and Networking Conference (CCNC), 2012 IEEE. Vol., ed.^eds. IEEE, pp. 667-671.
Lara, O.D., Labrador, M.A., 2013. A survey on human activity recognition using wearable sensors. Communications Surveys & Tutorials, IEEE. 15**,** 1192-1209.
Lara, Ó.D., et al., 2012b. Centinela: A human activity recognition system based on acceleration< i> and</i> vital sign data. Pervasive and mobile computing. 8**,** 717-729.
Le Faucheur, A., et al., 2008. Measurement of walking distance and speed in patients with peripheral arterial disease. Circulation. 117**,** 897-904.
Le Masurier, G.C., Tudor-Locke, C., 2003. Comparison of pedometer and accelerometer accuracy under controlled conditions. Medicine and Science in Sports and Exercise. 35**,** 867-871.
Le Rolle, V., et al., 2016. Recursive Model Identification for the Evaluation of Baroreflex Sensitivity. Acta biotheoretica. 64**,** 469-478.
Leitzmann, M.F., et al., 2007. Physical activity recommendations and decreased risk of mortality. Archives of internal medicine. 167**,** 2453-2460.
Leonard, W.R., Ulijaszek, S.J., 2002. Energetics and evolution: an emerging research domain. American Journal of Human Biology. 14**,** 547-550.
Leonard, W.R., 2003. Measuring human energy expenditure: What have we learned from the flex-heart rate method? American journal of human biology. 15**,** 479-489.
Leutheuser, H., Schuldhaus, D., Eskofier, B.M., 2013. Hierarchical, multi-sensor based classification of daily life activities: comparison with state-of-the-art algorithms using a benchmark dataset. PloS one. 8**,** e75196.
Liao, L., et al., 2007. Learning and inferring transportation routines. Artificial Intelligence. 171**,** 311-331.

Lippmann, R.P., 1989. Pattern classification using neural networks. IEEE communications magazine. 27**,** 47-50.
Lo, B., Yang, G.-Z., 2014. Appendix A: Wireless Sensor Development Platforms. Body Sensor Networks. 527.
Lo, B.P., et al., 2005. Body sensor network–a wireless sensor platform for pervasive healthcare monitoring, Vol., na.
Lorig, K.R., et al., 2001. Effect of a self-management program on patients with chronic disease. Effective clinical practice: ECP. 4**,** 256-262.
Luca, C., Salceanu, A., 2012. On the physiological influence of electromagnetic waves considering an electrical model of pulmonary ventilation. In: Word Energy System Conference (WESC). Vol., ed.^eds., pp. 28-30.
Lyden, K., et al., 2011. A comprehensive evaluation of commonly used accelerometer energy expenditure and MET prediction equations. European journal of applied physiology. 111**,** 187-201.
Lyden, K., et al., 2014. A method to estimate free-living active and sedentary behavior from an accelerometer. Medicine and science in sports and exercise. 46**,** 386.
Lynch, J., et al., 1996. Moderately intense physical activities and high levels of cardiorespiratory fitness reduce the risk of non-insulin-dependent diabetes mellitus in middle-aged men. Archives of internal medicine. 156**,** 1307-1314.
Mannini, A., Sabatini, A.M., 2010. Machine learning methods for classifying human physical activity from on-body accelerometers. Sensors. 10**,** 1154-1175.
Martin, S., et al., 2006. Self-monitoring of blood glucose in type 2 diabetes and long-term outcome: an epidemiological cohort study. Diabetologia. 49**,** 271-278.
Mathie, M.J., et al., 2004. Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement. Physiological measurement. 25**,** R1.
Maurer, U., et al., 2006. Activity recognition and monitoring using multiple sensors on different body positions. In: Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. International Workshop on. Vol., ed.^eds. IEEE, pp. 4 pp.-116.
McCool, F., et al., 1986. Estimates of ventilation from body surface measurements in unrestrained subjects. Journal of Applied Physiology. 61**,** 1114-1119.
McLaughlin, J., et al., 2001. Validation of the COSMED K4 b2 portable metabolic system. International journal of sports medicine. 22**,** 280-284.
Melanson Jr, E.L., Freedson, P.S., Blair, S., 1996. Physical activity assessment: a review of methods. Critical Reviews in Food Science & Nutrition. 36**,** 385-396.
Moeslund, T.B., Granum, E., 2001. A survey of computer vision-based human motion capture. Computer vision and image understanding. 81**,** 231-268.
Moore, S.T., MacDougall, H.G., Ondo, W.G., 2008. Ambulatory monitoring of freezing of gait in Parkinson's disease. Journal of neuroscience methods. 167**,** 340-348.
Mukhopadhyay, S.C., 2015. Wearable sensors for human activity monitoring: A review. IEEE sensors journal. 15**,** 1321-1330.
Müllenheim, P.Y., et al., 2016. Clinical Interest of Ambulatory Assessment of Physical Activity and Walking Capacity in Peripheral Artery Disease. Scandinavian journal of medicine & science in sports. 26**,** 716-730.
Murphy, S.L., 2009. Review of physical activity measurement using accelerometers in older adults: considerations for research design and conduct. Preventive medicine. 48**,** 108-114.
Myers, J., et al., 2004. Fitness versus physical activity patterns in predicting mortality in men. The American journal of medicine. 117**,** 912-918.
Nguyen, D., et al., 2013. Assessment of physical activity and energy expenditure by GPS combined with accelerometry in real-life conditions. Journal of physical activity & health. 10.
Ojeda, D., et al., 2016. Sensitivity Analysis of Vagus Nerve Stimulation Parameters on Acute Cardiac Autonomic Responses: Chronotropic, Inotropic and Dromotropic Effects. PloS one. 11**,** e0163734.
Owen, N., et al., 2010. Sedentary behavior: emerging evidence for a new health risk. In: Mayo Clinic Proceedings. Vol. 85, ed.^eds. Mayo Foundation, pp. 1138.
Pan, J., Tompkins, W.J., 1985. A real-time QRS detection algorithm. Biomedical Engineering, IEEE Transactions on. 230-236.
Pande, A., et al., 2014. Energy Expenditure Estimation in boys with Duchene muscular dystrophy using accelerometer and heart rate sensors. In: Healthcare Innovation Conference (HIC), 2014 IEEE. Vol., ed.^eds. IEEE, pp. 26-29.
Parkka, J., et al., 2006. Activity classification using realistic data from wearable sensors. Information Technology in Biomedicine, IEEE Transactions on. 10**,** 119-128.
Pärkkä, J., 2011. Analysis of personal health monitoring data for physical activity recognition and assessment of energy expenditure, mental load and stress, Vol., VTT.
Parsons, H.M., 1974. What happened at Hawthorne? Science. 183**,** 922-932.
Pentland, A., 2004. Healthwear: medical technology becomes wearable. Computer. 37**,** 42-49.
Perez, A.J., Labrador, M.A., Barbeau, S.J., 2010. G-sense: a scalable architecture for global sensing and monitoring. IEEE Network. 24.
Pober, D.M., et al., 2006. Development of novel techniques to classify physical activity mode using accelerometers. Medicine and science in sports and exercise. 38**,** 1626.
Powell, K.E., Paluch, A.E., Blair, S.N., 2011. Physical activity for health: What kind? How much? How intense? On top of what? Public Health. 32**,** 349.
Preece, S.J., et al., 2009. Activity identification using body-mounted sensors—a review of classification techniques. Physiological measurement. 30**,** R1.
Psota, T., Chen, K., 2013. Measuring energy expenditure in clinical populations: rewards and challenges. European journal of clinical nutrition. 67**,** 436-442.
Puska, P., Benaziza, H., Porter, D., 2004. Physical Activity, WHO Information Sheet on Physical Activity, WHO. Vol., ed.^eds. WHO.[Online]. Available: www. who. int/media/en/gs_pa. pdf.

Puyau, M.R., et al., 2004. Prediction of activity energy expenditure using accelerometers in children. Medicine and science in sports and exercise. 36**,** 1625-1631.
Ravi, N., et al., 2005. Activity recognition from accelerometer data. In: AAAI. Vol. 5, ed.^eds., pp. 1541-1546.
Ravussin, E., Bogardus, C., 1992. A brief overview of human energy metabolism and its relationship to essential obesity. The American journal of clinical nutrition. 55**,** 242S-245S.
Reddy, S., et al., 2010. Using mobile phones to determine transportation modes. ACM Transactions on Sensor Networks (TOSN). 6**,** 13.
Roetenberg, D., Slycke, P.J., Veltink, P.H., 2007. Ambulatory position and orientation tracking fusing magnetic and inertial sensing. IEEE Transactions on Biomedical Engineering. 54**,** 883-890.
Rosdahl, H., et al., 2010. Evaluation of the Oxycon Mobile metabolic system against the Douglas bag method. European journal of applied physiology. 109**,** 159-171.
Rosenberg, D., et al., 2016. Classifiers for Accelerometer-Measured Behaviors in Older Women. Medicine & Science in Sports & Exercise.
Ross, P.E., 2004. Managing care through the air [remote health monitoring]. IEEE spectrum. 41**,** 26-31.
Rothney, M.P., et al., 2007. An artificial neural network model of energy expenditure using nonintegrated acceleration signals. Journal of applied physiology. 103**,** 1419-1427.
Roy, S., Mccrory, J., 2015. Validation of maximal heart rate prediction equations based on sex and physical activity status. International journal of exercise science. 8**,** 318.
Ryder, J., et al., 2009. Ambulation: A tool for monitoring mobility patterns over time using mobile phones. In: Computational Science and Engineering, 2009. CSE'09. International Conference on. Vol. 4, ed.^eds. IEEE, pp. 927-931.
Sabatini, A.M., 2006. Inertial sensing in biomechanics: a survey of computational techniques bridging motion analysis and personal navigation. In: Computational Intelligence for Movement Sciences: Neural Networks and Other Emerging Techniques. Vol., ed.^eds. IGI Global, pp. 70-100.
Sallis, J.F., Saelens, B.E., 2000. Assessment of physical activity by self-report: status, limitations, and future directions. Research quarterly for exercise and sport. 71**,** 1-14.
Sasaki, J.E., et al., 2016. Performance of activity classification algorithms in free-living older adults. Medicine and science in sports and exercise. 48**,** 941.
Schneller, M.B., et al., 2015. Validation of Five Minimally Obstructive Methods to Estimate Physical Activity Energy Expenditure in Young Adults in Semi-Standardized Settings. Sensors. 15**,** 6133-6151.
Schoeller, D.A., Racette, S.B., 1990. A review of field techniques for the assessment of energy expenditure. The Journal of nutrition. 120**,** 1492-1495.
Sherar, L.B., et al., 2011. International children's accelerometry database (ICAD): design and methods. BMC Public Health. 11**,** 485.
Shoaib, M., Scholten, H., Havinga, P.J., 2013. Towards physical activity recognition using smartphone sensors. In: Ubiquitous Intelligence and Computing, 2013 IEEE 10th International Conference on and 10th International Conference on Autonomic and Trusted Computing (UIC/ATC). Vol., ed.^eds. IEEE, pp. 80-87.
Shoaib, M., et al., 2015. A survey of online activity recognition using mobile phones. Sensors. 15**,** 2059-2085.
Shotton, J., et al., 2013. Real-time human pose recognition in parts from single depth images. Communications of the ACM. 56**,** 116-124.
Skotte, J., et al., 2012. Detection of physical activity types using triaxial accelerometers. Journal of Physical Activity and Health.
St-Onge, M., et al., 2007. Evaluation of a portable device to measure daily energy expenditure in free-living adults. The American journal of clinical nutrition. 85**,** 742-749.
Standing, P., et al., 2001. Changes in referral patterns to cardiac out-patient clinics with ambulatory ECG monitoring in general practice. British Journal of Cardiology. 8**,** 394-399.
Staudenmayer, J., et al., 2009. An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. Journal of Applied Physiology. 107**,** 1300-1307.
Steimle, K.L., et al., 2009. A mathematical physiological model of the pulmonary ventilation. In: The 7th IFAC Symposium on Modeling and Control in Biomedical Systems. Vol., ed.^eds., pp. 222-7.
Steimle, K.L., et al., 2011. A model of ventilation of the healthy human lung. computer methods and programs in biomedicine. 101**,** 144-155.
Storti, K.L., et al., 2008. Gait speed and step-count monitor accuracy in community-dwelling older adults. Medicine & Science in Sports & Exercise. 40**,** 59-64.
Strath, S.J., et al., 2000. Evaluation of heart rate as a method for assessing moderate intensity physical activity. Medicine and science in sports and exercise. 32**,** S465-70.
Sung, M., Marci, C., Pentland, A., 2005. Wearable feedback systems for rehabilitation. Journal of neuroengineering and rehabilitation. 2**,** 17.
Swartz, A.M., et al., 2000. Estimation of energy expenditure using CSA accelerometers at hip and wrist sites. Medicine and Science in Sports and Exercise. 32**,** S450-S456.
Tao, Y., Hu, H., Zhou, H., 2007. Integration of vision and inertial sensors for 3D arm motion tracking in home-based rehabilitation. The International Journal of Robotics Research. 26**,** 607-624.
Tapia, E.M., et al., 2006. The design of a portable kit of wireless sensors for naturalistic data collection. In: Pervasive Computing. Vol., ed.^eds. Springer, pp. 117-134.
Tapia, E.M., et al., 2007a. Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart monitor. In: In: Proc. Int. Symp. on Wearable Comp. Vol., ed.^eds. Citeseer.

Tapia, E.M., et al., 2007b. Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor. In: Wearable Computers, 2007 11th IEEE International Symposium on. Vol., ed.^eds. IEEE, pp. 37-40.
Teng, X.-F., et al., 2008. Wearable medical systems for p-health. IEEE reviews in Biomedical engineering. 1**,** 62-74.
Tessendorf, B., et al., 2011. Recognition of hearing needs from body and eye movements to improve hearing instruments. Pervasive Computing. 314-331.
Thomas Wyss, M., Méder, U., 2010. Recognition of military-specific physical activities with body-fixed sensors. Military medicine. 175**,** 111858.
Thompson, D., et al., 2015. Multidimensional physical activity: an opportunity not a problem. Exercise and sport sciences reviews. 43**,** 67.
Trabelsi, D., et al., 2013. An unsupervised approach for automatic activity recognition based on hidden Markov model regression. Automation Science and Engineering, IEEE Transactions on. 10**,** 829-835.
Trost, S.G., et al., 2012. Artificial neural networks to predict activity type and energy expenditure in youth. Medicine and science in sports and exercise. 44**,** 1801.
Uswatte, G., et al., 2000. Objective measurement of functional upper-extremity movement using accelerometer recordings transformed with a threshold filter. Stroke. 31**,** 662-667.
Vähä-Ypyä, H., et al., 2014. A universal, accurate intensity-based classification of different physical activities using raw data of accelerometer. Clinical physiology and functional imaging.
Valanou, E., Bamia, C., Trichopoulou, A., 2006. Methodology of physical-activity and energy-expenditure assessment: a review. Journal of Public Health. 14**,** 58-65.
van Hees, V.T., et al., 2013. Impact of study design on development and evaluation of an activity type classifier. Journal of Applied Physiology.
Van Sluijs, E.M., McMinn, A.M., Griffin, S.J., 2007. Effectiveness of interventions to promote physical activity in children and adolescents: systematic review of controlled trials. Bmj. 335**,** 703.
Verberk, W.J., et al., 2007. Self-measurement of blood pressure at home reduces the need for antihypertensive drugs. Hypertension. 50**,** 1019-1025.
Viéville, T., Faugeras, O.D., 1990. Cooperation of the inertial and visual systems. In: Traditional and non-traditional robotic sensors. Vol., ed.^eds. Springer, pp. 339-350.
Waldrop, J., Stern, S.M., 2003. Disability Status, 2000, Vol. 3, US Department of Commerce, Economics and Statistics Administration, US Census Bureau.
Wang, H., et al., 2010. Resource-aware secure ECG healthcare monitoring through body sensor networks. IEEE Wireless Communications. 17.
Wang, W.-z., et al., 2011. Analysis of filtering methods for 3D acceleration signals in body sensor network. In: Bioelectronics and Bioinformatics (ISBB), 2011 International Symposium on. Vol., ed.^eds. IEEE, pp. 263-266.
Warburton, D.E., Nicol, C.W., Bredin, S.S., 2006. Health benefits of physical activity: the evidence. Canadian medical association journal. 174**,** 801-809.
Wijndaele, K., et al., 2015. Utilization and Harmonization of Adult Accelerometry Data: Review and Expert Consensus. Medicine and science in sports and exercise.
Winkley, J., Jiang, P., Jiang, W., 2012. Verity: an ambient assisted living platform. IEEE Transactions on Consumer Electronics. 58.
Witt, J.D., et al., 2006. Measurement of exercise ventilation by a portable respiratory inductive plethysmograph. Respiratory physiology & neurobiology. 154**,** 389-395.
Wu, J., et al., 2011. Automated time activity classification based on global positioning system (GPS) tracking data. Environmental Health. 10**,** 101.
Yin, J., Yang, Q., Pan, J.J., 2008. Sensor-based abnormal human-activity detection. IEEE Transactions on Knowledge and Data Engineering. 20**,** 1082-1090.
Zakeri, I.F., et al., 2013. Cross-sectional time series and multivariate adaptive regression splines models using accelerometry and heart rate predict energy expenditure of preschoolers. The Journal of nutrition. 143**,** 114-122.
Zhang, K., Pi-Sunyer, F.X., Boozer, C.N., 2004. Improving energy expenditure estimation for physical activity. Medicine and science in sports and exercise. 36**,** 883-889.
Zhang, T., et al., 2006. Using wearable sensor and NMF algorithm to realize ambulatory fall detection. Advances in Natural Computation. 488-491.
Zheng, Y., et al., 2008. Understanding mobility based on GPS data. In: Proceedings of the 10th international conference on Ubiquitous computing. Vol., ed.^eds. ACM, pp. 312-321.
Zhu, R., Zhou, Z., 2004. A real-time articulated human motion tracking using tri-axis inertial/magnetic sensors package. IEEE Transactions on Neural systems and rehabilitation engineering. 12**,** 295-302.
Zongker, D., Jain, A., 1996. Algorithms for feature selection: An evaluation. In: Pattern Recognition, 1996., Proceedings of the 13th International Conference on. Vol. 2, ed.^eds. IEEE, pp. 18-22.

SCIENTIFIC VALORIZATION

International Journals

1. **Abdul Rahman H**., Di Ge, A. Le Faucheur, J.Prioux, G.Carrault: Advanced classification of ambulatory activities using spectral density distances and heart rate. Biomedical Signal Processing and Control Journal. 2017.
2. **Abdul Rahman H**., Bagot M., Khreiss S., Boulanger J., Dumont R., Gastinger S., Ge D., Guidec F., Le Faucheur A., Launay P., Prioux J., Quinton P., Somat A., Ville N., Carrault G.: SHERPAM: an open, versatile platform for mobile e-Health. Preliminary Results. Sensors. 2017 (submitted).
3. Dumond R., Gastinger S., **Abdul Rahman H.**, Le Faucheur A., Quinton P., Kang H. & Prioux J.: Estimation of Respiratory Volume from Thoracoabdominal Breathing Distances: Comparison of Two Models of Machine Learning. European Journal of Applied Physiology. 2017.

International Conferences

1. **Abdul Rahman H**., G.Carrault, Di G., J. Prioux, A. Le Faucheur, R. Dumond, H. Amoud: Ambulatory physical activity representation and classification using spectral distances approach. 3rd International Conference on Advances in Biomedical Engineering (ICABME'15), September 2015.
2. **Abdul Rahman H**., Le faucheur A., Ge D., Carrault G., Prioux J.: Recognition of dynamic and stationary activity types in out-of-lab conditions using rate and accelerometers data fusion. 21 annual Congress of the European College of Sport Sciences (ECSS'16), July 2016.
3. **Abdul Rahman H**., Le Faucheur A., Ge D., Carrualt G., Prioux J.: From bounded to pragmatic data collection protocol: Validity of state-of-the-art activity recognition in daily life context. 5th International Conference on Ambulatory Monitoring of Physical Activity and Movement (ICAMPAM'17), June 2017.
4. Dumond R., Le Faucheur A., de Mullenheim P.-Y., **Abdul Rahman H.**, Prioux J.: Couplage de l'accéléromètre, du géo-positionnement satellitaire et de la fréquence cardiaque dans la détection d'activités physiques et sédentaires. 2ème congrès international de l'Association Africaine des Sciences du Sport, Mars 2016.
5. Dumond R., Le Faucheur A., de Mullenheim P.-Y., **Abdul Rahman H**., Prioux J. Couplage du géo-positionnement satellitaire, de l'accélérométrie et de la fréquence cardiaque dans la détection de différents types d'activités physiques et sédentaires. Associations des Chercheurs en Activités Physiques et Sportives, Octobre 2015.

Awards

Honorable Mention Award for the student oral presentation in the International Conference on Ambulatory Monitoring of Physical Activity and Movement 2017 (ICAMPAM'17).