Ana Luísa Gonçalves Neves Gomes

Licenciatura em Engenharia Biomédica

Human Activity Recognition with Accelerometry: Novel Time and Frequency Features

Dissertação para obtenção do Grau de Mestre em
Biomedical Engineering

Orientador : Dr. Hugo Gamboa, Faculty of Sciences and Technology, New University of Lisbon

Co-orientador : Dr. Vítor Paixão, Champalimaud Neuroscience Programme, Champalimaud Centre for the Unknown

Júri:

Presidente: Prof. Dr. Mário Secca

Arguente: Prof.ª Dr.ª Carla Quintão

FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

November, 2014

Human Activity Recognition with Accelerometry: Novel Time and Frequency Features

To my parents and sister

Acknowledgements

First of all, I acknowledge Professor Hugo Gamboa, who always supported me and saw potential in my abilities even at the beginning of this thesis. I appreciate the opportunity he gave me to work with him in an interesting area which allowed me to develop my programming skills and to create an out-of-the-box approach over the whole research. I'm also grateful for his constant sharing of knowledge through rigorous orientation and dedicated effort for the success of this research. I also thank to Ricardo Gomes and to the rest of the staff from Plux for their availability and support.

I acknowledge my co-supervisor, Dr. Vítor Paixão, from Champalimaud Foundation for his constant availability, enthusiasm and excellent guidance and advices. I am also grateful for the support and good mood of the researchers from the Rui Costa group lab and all the volunteers of the Foundation.

I acknowledge Inês Machado for her willingness and support at the beginning of my thesis. Her work was very important to get a better and deeper understanding of what was previously performed and the most important topics to develop.

Moreover, I acknowledge Professor Ricardo Matias, from the Politécnico de Setúbal, who gave new insights on the developed work with great enthusiasm and interest and for his availability to apply the achieved knowledge in the clinical environment.

Finally, I acknowledge my boyfriend, Patrizio Di Censi, for his daily understanding, patience and unconditional support. To my friends, Rita Alhandra, Ricardo Silva, M. Marta Santos, Inês Vale, Maria Inês Silva, David Branha and Catarina Cavaco, for these last years full of remarkable and fun moments. Thank you all for supporting me when I needed and to call me to reason when I deserved.

I am especially grateful to my parents and my sister Gabriela, for their support and for always believing in me as a person and professional. I thank them for their understanding, trust and for giving me everything without demanding anything in return.

Last but not least, I thank my grandmother Maria Luísa who recently passed away but had always inspired me during the development of the thesis. She was a person who was proud and took care of me and was my confidant. She was and will always be with me.

Abstract

Human Activity Recognition systems require objective and reliable methods that can be used in the daily routine and must offer consistent results according with the performed activities. These systems are under development and offer objective and personalized support for several applications such as the healthcare area.
This thesis aims to create a framework for human activities recognition based on accelerometry signals. Some new features and techniques inspired in the audio recognition methodology are introduced in this work, namely Log Scale Power Bandwidth and the Markov Models application.
The Forward Feature Selection was adopted as the feature selection algorithm in order to improve the clustering performances and limit the computational demands. This method selects the most suitable set of features for activities recognition in accelerometry from a 423^{th} dimensional feature vector.
Several Machine Learning algorithms were applied to the used accelerometry databases – FCHA and PAMAP databases - and these showed promising results in activities recognition.
The developed algorithm set constitutes a mighty contribution for the development of reliable evaluation methods of movement disorders for diagnosis and treatment applications.

Keywords: Human Activity Recognition, Forward Feature Selection, Log Scale Power Bandwidth, Wavelets, Hidden Markov Models, Clustering Algorithms.

Resumo

Os sistemas de Reconhecimento de Actividade Humana requerem métodos fiáveis e objectivos que possam ser aplicados em diferentes actividades do dia-a-dia e devem oferecer resultados consistentes de acordo com as actividades realizadas. Estes sistemas encontram-se em desenvolvimento e permitem a obtenção de dados objectivos e personalizados, direccionados para diferentes aplicações tais como as áreas do bem-estar e da saúde.

Esta tese teve como objectivo criar um sistema de reconhecimento de actividades humanas baseado em sinais de acelerometria. Novas features e métodos inspirados no reconhecimento áudio são introduzidos neste trabalho, nomeadamente a feature Log Scale Power Bandwidth e a aplicação de Modelos de Markov Escondidos.

O Forward Feature Selection foi adoptado como algoritmo de selecção de features de forma a melhorar a performance dos métodos de clustering e limitar as exigências computacionais. Este método selecciona o conjunto de features mais adequado para o reconhecimento de actividades por acelerometria a partir de um vector de features de 423 dimensões.

Vários métodos de Machine Learning foram aplicados às duas bases de dados adoptadas - base de dados FCHA e PAMAP – e estes apresentaram resultados promissores na identificação das actividades realizadas.

O conjunto de algoritmos desenvolvidos nesta tese poderá contribuir para o desenvolvimento de métodos de avaliação fiáveis para o diagnóstico e tratamento de deficiências motoras.

Palavras-chave: Reconhecimento de Actividade Humana, Forward Feature Selection, Log Scale Power Bandwidth, Wavelets, Modelos de Markov Escondidos, Algoritmos de Clustering.

Contents

List of Figures

List of Tables

Acronyms

ACC Accelerometry

ARI Adjusted Rand Index

BIC Bayesian Information Criterion

CoM Center of Mass

CVA Cerebrovascular Accident

ECG Electrocardiogram

EMG Electromyogram

FFT Fast Fourier Transform

FCHA Foundation Champalimaud Human Activity

HA Hungarian Accuracy

HAR Human Activity Recognition

HMM Hidden Markov Model

LSPB Log Scale Power Bandwidth

MEMs Micro-Electro-Mechanical-Systems

MFCC Mel-frequency Cepstrum Coefficients

MS Multiple Sclerosis

OCD Obsessive Compulsive Disorder

PA Physical Activity

PD Parkinson's Disease

SE Spectral Entropy

SMA Signal Magnitude Area

1

Introduction

1.1 Motivation

Over time, the increasingly demand for objectivity in clinical diagnosis and the continuous pursuit for human wellbeing led to the development of software and hardware for healthcare. The combined efforts of medicine and engineering created and developed techniques that provide large amounts of information and simultaneously allow to interpret the generated data. The acquired information can be in the form of biosignals (signals measured and monitored from biological events), such as Electromyogram (EMG), Electrocardiogram (ECG) or an Accelerometry (ACC) signal. Biosignal processing requires an acquisition stage and a transformation stage with conversion (from digital into acceleration units - g - in the accelerometry case), filtering and extraction of the useful features, which will depend on the aim of the investigation. The feature extraction step becomes very important for activity recognition because it defines what information we will cluster with. The last step is the clustering stage, applied through the obtained features, and the final evaluation of the final results for a proper interpretation of the results. Due to its importance, the clustering development for Physical Activity (PA) recognition is the central focus of this thesis and it is a main task in exploratory data mining to organize datasets into different groups, which is used in many fields, including in pattern identification.

Clustering techniques applied to biosignal's morphology, allow the detection and classification of the different physical positions and everyday movements. The clustering step (vide chapter 2) is crucial for pathology detection and abnormal behavior evaluation [44] due to changes that can be detected in the morphology of the accelerometry signal. Therefore, it is mandatory to acquire enough knowledge and data in order to be

able to distinguish between normal movement patterns and those of certain pathologies. The accelerometer research in human medicine aims to help in the diagnosis and prevention of various diseases in which changes of the PA can be detected, such as Obsessive Compulsive Disorder (OCD), Pakinson's Disease (PD), Multiple Sclerosis (MS) and Cerebrovascular Accident (CVA) [35]. Pathologies such as OCD present difficulties in diagnosis because it requires time-consuming and subjective monitorization techniques. Thus, objective clinical support becomes imperative in this type of pathologies, not only for diagnosis but also for treatment and evaluation by monitorization and measurement of critical information that the signal may present.
Figure 1.1 presents the protocol used in clustering and PA recognition study with an ACC signal as it was discussed earlier. Chapter 2 presents the same organization described in the scheme for an easier understanding of the adopted methodology.

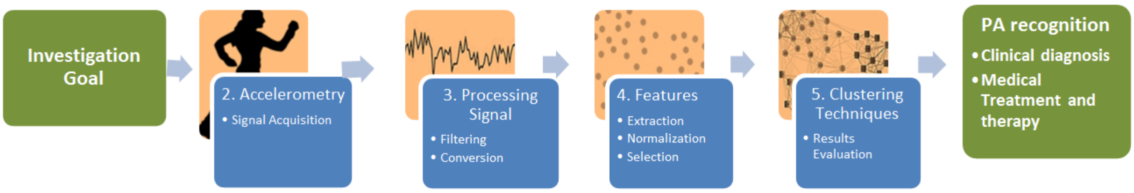


Figure 1.1: ACC signal analysis with clustering approaches: after setting an investigation aim, data is collected (accelerometry signals) followed by signal processing (separation of the gravitational component from the body acceleration and conversion). After feature extraction from the signal and clustering, it is possible to obtain the PA recognition. This acquisition knowledge can contribute for the development of several pathologic condition treatments such as OCD, PD and CVA.

1.2 State of the Art: Brief Historical Perspective

Several techniques for data acquisition and processing have been developed to improve the early diagnosis and clinical treatment of various diseases, such as MS and PD. MS is a neurological and progressive disorder where the subject shows disturbances associated with spasticity, weakness, speed and cadence of his movements. Such physical problems contribute for the Center of Mass (CoM) displacement and thus motion sensors, such as accelerometry, can provide objective and valid information about PA which contributes for the objective diagnosis of [5] .
PD is another disease with a progressive dysfunction of the central nervous system and presents many features associated with the human movement, such as tremor and postural instability [11] , [34] . These features are related to a difficulty in seeking the stability required to balance the center of mass and therefore these individuals with PD show a high risk for falling [34].

More recently, Fazio and co-workers [13] concluded that it is possible to compare parkinsonian patients with healthy individuals through the accelerometry signals and it is possible to evaluate the disease and its progression with the acquired information. In this research, 24 healthy subjects, 24 with ataxic gait and 17 patients affected by PD were studied. It was observed a significant reduction of acceleration parameters in neurological patients when compared with healthy subjects. The mean acceleration was 0.74 ms^{2} for parkinsonian, 1.07 ms^{2} for ataxic and 1.38 ms^{2} for healthy patients. The root-mean square of the acceleration was also used to quantify the attenuations of acceleration, which reflected the importance of the feature extraction and the research of new ones from the accelerometry signal to obtain new information.
In [50] authors applied a PA recognition methodology using different features extracted from the raw accelerometer data. Several activities were analyzed (figure 1.2) using windows of one second (segmentation) with 50% overlap. The usefulness of four features to study specific characteristics of the accelerometer signal has been shown as well in this study: the mean, standard deviation, energy and signal correlation. U. Maurer and co-workers [50] achieved high accuracy results with decision trees: 92%.

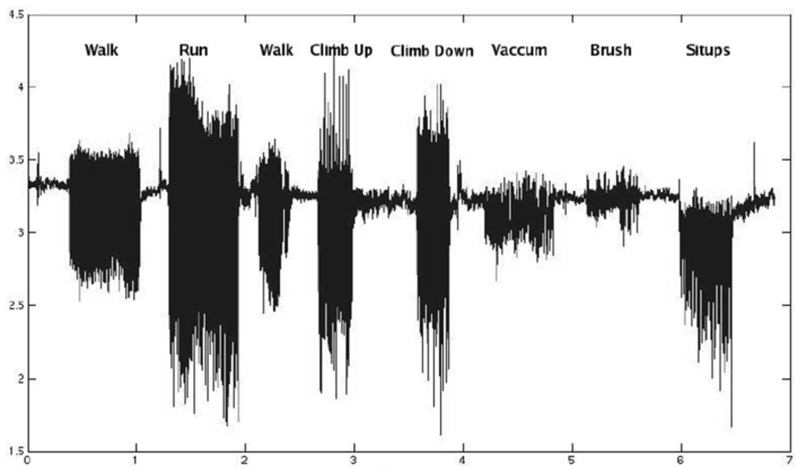


Figure 1.2: Accelerometry signal in X-axis, located in the pelvic region [50], which presents a clear distinction between the physical activities: walking, running, walking, climbing up and climbing down stairs, vacuuming, brushing teeth and sit-ups.

Similarly to [50], I. Machado [35] concluded that four features from time and frequency domains can achieve high recognition accuracy (about 99%). The used features were mean, autocorrelation, root mean square and Mel Frequency Cepstral Coefficients (MFCC). In [35], the author contributed to the PA recognition in accelerometry with the K-means technique and also inferred that one waist-worn accelerometer can identify the physical activities in an adequate manner.
With lower accuracy rates, A. Khan and his team [1] presented some conclusions as to PA recognition with several clustering techniques. In this study, six daily physical activities with a triaxial accelerometer and annotation were performed during 34 hours, inside and

outside the laboratory. For each second of the accelerometry data, Spectral Entropy (SE), the Autoregressive Coefficients and the Signal Magnitude Area (SMA) were calculated. These features illustrated the effectiveness of the method, with an average accuracy of 95%.

The referred features were used in a clustering stage for human activity recognition with K-means algorithm, an unsupervised and iterative clustering method which produces very effective results. On the other hand, this clustering technique shows a long running time for its computation complexity. Therefore, the authors from [42] suggest an improved k-means algorithm retrieving similar clustering results as the standard k-means but with the advantage of having a shorter running time by improving its speed.

Furthermore, in [2] it is shown not only k-means clustering presents a good performance, but also spectral clustering and affinity propagation approaches show high accuracy results.

Moreover, in another article [33], where k-means is exploited and also gives good accuracy rates, a comparison between different similarity measures is made. K-means clustering results, as in other clustering methods, depends on similarity measurements, which are values which translate the distance between two data points to be compared.

Euclidean and Manhattan distance functions were compared and R. Loohach and colleagues [33] concluded that k-means clustering performance increases when Manhattan distance function is used, except for k=4, 5, 7. For these last k-values, Euclidean distance shows a better efficiency. Furthermore, Manhattan distance function requires less computation than the Euclidean one, which improves the time complexity of the clustering technique.

In [55] the Hidden Markov Model (HMM) is used to identify the sequence corresponding to 12 physical activities and the final results lead to 91.4 % as a mean correct classification rate averaged over all observations. They also concluded that the HMM study leads to a better classification rate (84 %) than with k-means algorithm (60 %). In spite of this fact, which highlights the potential benefit of automatic identification of human activity with the HMM approach, transitions between activities are also important and potentially rich in motion information. These transitions are highly relevant and should be taken into consideration during an HMM application. With all available information related to these areas, we propose a study with clustering techniques on accelerometry data with complex motor activities for the PA recognition as close as possible to real movement. The dissertation of this thesis was developed in PLUX – Wireless Biosignals S.A. in collaboration with the Champalimaud Neuroscience Programme, Champalimaud Centre for the Unknown.

1.3 Objectives and Thesis Structure

The developed work present in this thesis aims to create and develop a framework to recognize different human activities and positions by their accelerometry signals using advanced algorithms, such as machine learning techniques and Markov Models. All relevant stages of this framework are represented in the scheme shown in Figure 1.3. The implemented algorithms are design to identify relevant parameters from data and to support movement disorders diagnosis and physiotherapy.

As an introduction, the thesis presents a state-of-the-art with the respective literature review in order to contextualize the experimental obtained data. The topics of this thesis cover features and clustering techniques investigation, accelerometry signals, Markov Models study and the literature revision that provides the motivation for this thesis.

For signals acquisition, a protocol formed by several physical activities was adopted and an ACC system was placed on the waist (near of the CoM) and to monitor individuals for data collection. The protocol for data acquisition was composed by several activities, such as lying down, walking, sitting, running, standing, ascending and descending stairs and allowing the possibility to add jumping and riding a bicycle. Then, it is possible to add or even create new features that may lead to an improvement of PA recognition in several domains and develop algorithms to extract relevant information from the ACC data. The evaluation of the clustering application is a crucial stage of the thesis development to report the effectiveness of the framework and data interpretation. Therefore, it is important to study the influence of features and other parameters on the performance of different clustering algorithms. The Hungarian Accuracy (HA) is also implemented as a metric to evaluate the output labels.

Finally, the framework evaluation in daily routine settings and the knowledge application to a real life case acquired over this work could also be an interesting step to follow, such as the progression of a certain disease or the locomotion impairment as a consequence of a CVA.

1.4 Thesis Overview

This thesis is divided into six chapters.

The first two chapters concern to the state-of-art that supports all the research required for the data processing presented in this study. Chapter 1 shows a brief historical perspective, the objectives and motivation of the present thesis. Chapter 2 presents the required theoretical background for the field of research, including basic operation of the accelerometer, its advantages and operating criteria, the composition of the acquired signals and the existing machine learning techniques to be used.

Chapter 3 is related with the materials and methods used for the acquisition system adopted in this work. Other ACC databases used for testing the developed framework are referred in this section as well and represented in the following figure 1.3.

Chapter 4 and 5 correspond to the results obtained in this work. In particular, chapter 4 presents a set of features similar to those used in [35] (time, statistical and spectral domains) and a new set of implemented features inspired on audio recognition methods, including features from the time-frequency domain.
The choice of features is a fundamental step to apply machine learning methodologies to sensor data and it influences the outcome of any approach. The proposed methodology for feature selection and the influence of several parameters such as window size and HMM application are also described.
Chapter 5 carries out all results achieved from all studies referred before in chapter 4 with clustering performance evaluation methods.
Chapter 6 the main conclusions of the present work are drawn.
Appendix B presents additional results and appendix C consists on two papers: one recently submitted to the 8th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2015), held in Lisbon in January 2015, and other published in Elsevier - Procedia Computer Science.

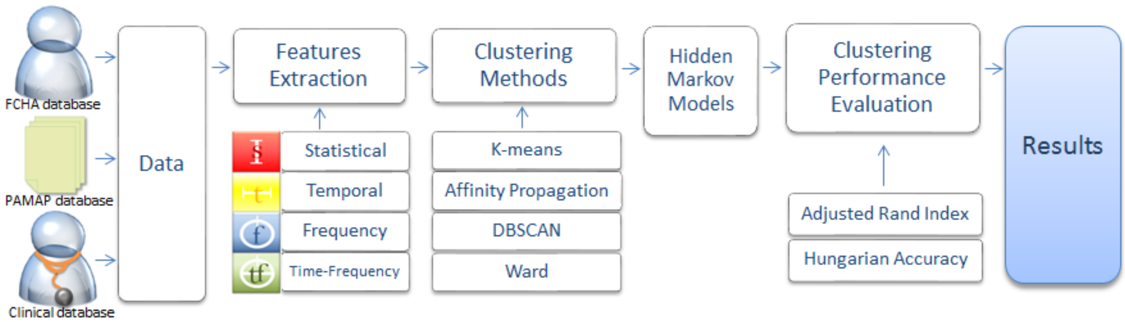


Figure 1.3: Overall structure of the framework developed in the present work for Human Activity Recognition (HAR) systems. The first stage was related to the acquisition of ACC data from three databases: Foundation Champalimaud Human Activity (FCHA) database, PAMAP database and clinical database (vide chapter 3). All databases were used in all stages of the framework: feature extraction, clustering and HMM application. Feature extraction was computed with features from all four domains (statistical, time, frequency and time-frequency) whose output was used as input of the unsupervised learning methods (K-means, Affinity Propagation, DBSCAN and Ward). The Hidden Markov Model was applied to the clustering results and this process is evaluated according to two performance evaluation methods: Adjust Rand Index (ARI) and HA (vide chapter 5).

2

Theoretical Background

2.1 Accelerometry and Accelerometer

Recently, the use of systems based on accelerometers to quantify and characterize human movement has increased significantly. These systems are becoming indispensable in the field of diagnosis, treatment and research by providing qualitative and quantitative data [27].
Overall, there are several types of accelerometers such as fluid, magnetic, strain gauge, piezoresistive, capacitive and piezoelectric [27], [58]. These latter three classes are more commonly applied in the classification and study of the human movement. The general and basic mechanism of the Micro-Electro-Mechanical-Systems (MEMs) will be described in the following section 2.1.1 [27].

2.1.1 The Accelerometer System Function

The underlying basis of operation for the measurement of acceleration is represented as a mass-spring system [19],[18], [27]. In MEMs sensors, the accelerometric system works accordingly to the Hooke's and Newton's second law, where: F = kx (eq. 1) and F = ma (eq. 2), respectively. F denotes the force *(N)*, m is the mass of the system *(kg)*, k is the spring constant, x is the displacement (*m*) and a is the acceleration *(ms^{-2})*. When a compressive or extension force is detected due to a given movement, the spring-mass system will react to produce a proportional force to the initially imposed force [18], [27], [35]. If the system's mass is known, as well as the spring constant, it is possible to determine the acceleration of a given movement, accordingly to its displacement by: (eq. 1)=(eq. 2).

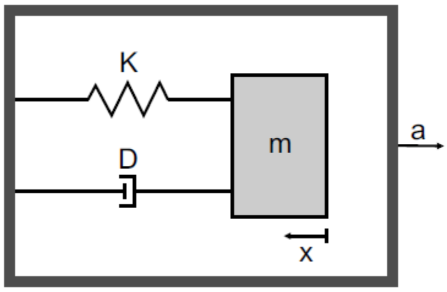


Figure 2.1: Representation of the operating principle of MEMs accelerometer where k is the spring constant, x is the displacement of the system with mass m and a is the acceleration of the mass-spring system, calculated from the equations 1 and 2.

2.1.2 The Accelerometer Signal

The accelerometer signal is composed of a static component, known as the gravitational acceleration (in g), which is always present (Fig. 2.2) and qualifies the accelerometer as an inclinometer. In these conditions, an horizontal sensing axis detects acceleration around 0 g and a vertical sensing axis (aligned with the center of the Earth) can detect an acceleration value around 1 g [27]. Therefore, the gravitational acceleration offers information concerning space orientation of the accelerometer and thus, the subject's posture.

In addition to the gravitational component, accelerometer-based systems show inertial components, in a local coordinate system in the 3D accelerometer case. Each component or plane responds to the frequency and intensity of the movement of the subject [19], [27]. There are uniaxial, biaxial and triaxial devices, for 1D, 2D or 3D accelerometers respectively, and its selection depends on many factors such as the aim of the investigation and its budget.

Due to the characteristics of the accelerometer signal, calibration is an important procedure and it is performed by placing the accelerometer in a known static orientation position relative to the given body under study [23]. The reference point is taken as 0 g in a free-fall situation. The acquired data from the accelerometer depends on several conditions and criteria, which will be further discussed in section 2.1.3.

2.1.3 The Accelerometer Functioning Criteria for Human Activity Analysis

The body acceleration components change continuously during the movement of the subject, regardless of the accelerometer location. The output data will certainly present background noise, as the result of electronic factors, motion artefacts and others [48].

An accelerometer can detect rotational and translational accelerations during movement. The selection of the accelerometer placement becomes very important in this context in order to reduce these motion effects [14], [28], [40]. The placement of the sensor must be

studied in order to ensure the patients' comfort and to assure that the study is performed as planned [35]. In some situations, where the goal is to study 'whole body' movements, the most used placement is close to the center of mass of the human body, for instance the sternum [28], waist [19] or lower back, in the lumbar region [41].
The task performed by the subject will produce variations in the detected accelerations. If the subject is at rest, acceleration is determined by its position according to the gravitational direction and it is possible to determine the posture of the subject. It is also important to consider the orientation of the accelerometer relatively to the body location for better understanding and analysis of the output data.

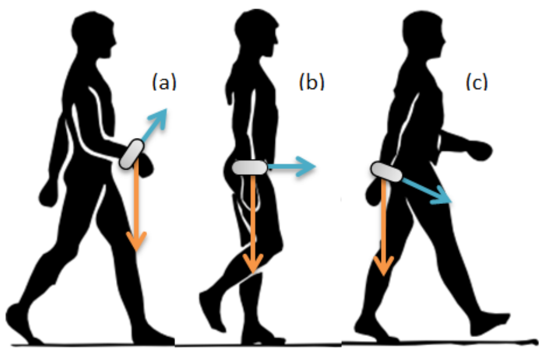


Figure 2.2: Representation of a 1D accelerometer on the wrist and respective sensing axis (in blue) accordingly to the subject's movement. The orange arrows represent the gravitational acceleration. In (b) the output of the sensing axis is aligned with the global horizontal and the gravitational component is null. In (a) and (c) the sensing axis is no longer perpendicular to the gravitational vector (unlike in b) and will show not just the body acceleration but a part of the gravitational component (from 0 to 1 g).

2.1.4 Advantages of Accelerometer System

The MEMs technology presents several advantages over more traditional and subjective analysis, such as:

- Relative low cost and simple to operate compared with traditional optical gait analysis such as optical markers [13], [18], [19], [27]. It can be applied in real life situations such as to backup rehabilitation;
- Portability [54], [58]. It can be used in different environments and infrastructures, not just in a laboratory [10], [13], [23], [27];
- Small size of the accelerometry apparatus allowing the possibility of more test movements without restrictions [13], [19], [27], [58];
- Allows direct measurement of acceleration in three dimensions, reducing error due to displacement and velocity [27], [53];

- The ability to record data for longer periods [19];
- Few requirements for energy, memory and processing power [35].

2.1.5 Applications of the Accelerometer Technology

Based on the accelerometers' operation method and its several advantages, this type of device presents a wide range of applications, such as military, automotive, industrial or science and medical fields [10], [18]. In the two latter areas, accelerometry is especially important due to its several applications and its importance in the clinical diagnosis and treatment. The following list presents some applications of the accelerometry signal in the medical investigation field:

- Understand the stability in walking pattern [27] and evaluate walking impairment, which is a prevalent feature of MS [19], [53];
- Evaluation of PA pattern analysis in many situations, such as skiing [41], control of the elderly [3], [58], fall prediction [3], [27], [58] and neurodegenerative diseases diagnosis, such as PD;
- Sleep-control studies as well as in autism [19] and conditions with mobility-impairment, such as obesity or stroke.

2.2 Accelerometry Signal Processing

For the signal processing step the knowledge of the particular characteristics of the signal as well as specific filtering and conversion techniques are mandatory. This process allows a better analysis of the critical information and the higher reliability of results. In this section, the characteristics of accelerometry signals and a short approach of the necessary conversion, filtering and segmentation techniques applied to ACC signals are described.

2.2.1 Characteristics of the Signal

As mentioned in section 2.1.2, the ACC signal is constituted essentially by the gravitational acceleration and the linear or body acceleration due to the human movement. There is also a third component related with the intrinsic noise of the electronic system and measurement conditions [14], [35].
For a careful analysis of the body acceleration, a separation between the gravitational and linear accelerations is required. However, both overlap in the frequency spectrum, making the filtering step more difficult: the linear acceleration is contained mostly below 15 Hz [48] and the gravitational is extended from 0 up to several Hz [28], [35].

2.2.2 Filtering and Segmentation Process

The raw data generated by the accelerometric device can be recorded or wirelessly transmitted, which is then followed by data conversion and filtering techniques [23], [35], [54]. Data can be converted to acceleration units, ms ^{2} or g [41], according to the device specifications and acquisition type.
Afterwards, an high-pass filter is applied to the raw data signal with a cutting frequency (with different frequency values according to the literature). While F. Foerste and coworkers [14] applied a 0.5 *Hz* cut-off high-pass filter frequency to all the signals, [35] used a 0.25 *Hz* cut-off frequency on a second order Butterworth high-pass to separate both accelerations, body and gravitational components.

Segmentation can be applied throughout several approaches, which help to clearly identify the most relevant properties of the signal and the most important information obtained from the dataset to use in the feature extraction and selection.
Timing and sliding windows may be considered as segmentation methods and the most important questions about these techniques are: What is the most suitable length of each window? Is an overlap percentage required? All these questions should be answered in order to maximize the information and high resolution is obtained from the signal [32], [35], [39], [50].

2.3 Machine Learning

Machine Learning is considered a subarea of Artificial Intelligence since those algorithms make computers learn how to behave more intelligently by somehow generalizing rather than just storing and recovering data items. This field offers a wide variety of methods for PA recognition and many of them may be based on the feature representation of accelerometry signal. There are essentially two types of machine learning algorithms:

- Unsupervised learning or clustering – where the system receives unlabeled dataset with no information about the class to which each point belongs to [44]. In this method, data points are grouped into clusters based on similarity of some metric [12], [17], [25], [35]. The goal of clustering is to determine well-separated and homogeneous groups. In other words, the system must have simultaneously similar entities belonging to the same cluster and different entities belonging to different clusters [4], [42];

- Supervised learning or classification – is a process which identifies unknown and unlabeled data with a training set of data base. This database contains information about all known classes [24], [26], [55] and this method will be introduced in more detail in the section 2.3.1.

There is a hybrid machine learning type named semi-supervised learning, where only a portion of the training dataset are available and the unlabeled data part is also used in the learning process [24], [25].

2.3.1 Machine Learning Techniques

Generally clustering methods aim to identify and organize a dataset into different labeled groups, through their similarity in order to provide meaningful information from the original dataset. These groups are called clusters and they can differ in shape, size and density [25]. Values that belong to the same cluster show a higher similarity between them than to different clusters [25], [31], [35], [44]. The collected data can be analyzed with clustering methods for PA recognition through different extracted features [50].
Clustering is essential when an identification of unlabeled data is required and it can be applied to different types of data, such as binary, categorical, numeral, ordinal, textual, spatial, and temporal, among others [31], [32]. The quality of the final results depends on both the implemented clustering method and the adopted approach [25], [44].
Dataset can be named static if its features do not change with time and the majority of clustering methods have been implemented on static data [31].
Clustering itself is not a specific algorithm but rather a general task that comprises different algorithms and methods. In the following sections, general-purpose clustering algorithms used in time series clustering studies will be briefly addressed and measurements to study the similarity between data points, such as the requested criteria for performance evaluation of the final results [25], [31].

On the other hand, for supervised learning and for an in-depth study of PA recognition, it would also be interesting to use the Markov model application, which aims to present results from the present outcomes through probabilities. This model is used for modelling conditions of the occasions that may happen repeatedly over time or for modelling predictable events that take place over time [15].

The main steps for application of the Markov Model are essentially: enumerate the states (in this case, states are all the activities performed by the subject in the accelerometry study) and define admissible state transitions; identify the probabilities and associate them with the transitions; identify the outcome values and determine the expected results.
Therefore, it is possible to analyze the transition between the most probable activities with all the original data. For instance, activity transition between the seating and running tasks is not possible without the activity "to get up" between both first. The subject needs to seat, after to get up and then to run. Without the get up task in the middle the subject cannot start to run in the seated position. Thus, the "sit-run" transition will show a low probability of happening. All these transitions are crucial for a good Markov

analysis and for credible PA recognition results [16], [17], [43], [55].

2.3.2 Unsupervised Learning Methods

Unlike static data, feature time series contain values which change over time [31].
Time series clustering may be divided into two groups: "whole clustering" and "subsequence clustering", which is performed on several individual time series or on sliding window extractions of a single time series, respectively [35]. Moreover, time series clustering requires a clustering algorithm which divides the dataset into different clusters.
The choice of algorithm depends on the type of dataset available and the goal of the research. It is used initially to convert time series into static data format in order to use the static data algorithms [28], [37]. In the existing literature based on PA recognition with accelerometry data, clustering methods vary according to the intended approach [28].In agreement with [22], clustering methods are classified into five classes: partitioning class; hierarchical class; model-based class; density base class and grid-based class.
Each of these classes is summarized below as well as some of the most important clustering algorithms belonging to the Scikit-learn library. The latter is a machine learning algorithm collection for supervised and unsupervised environments with an useful interface strongly integrated within the Python language [46].

Partitioning Class Initially, the unlabeled data is divided into k partitions (clusters) and each partition contains at least one data object. A cluster has a crisp partition when each data object belongs to just one cluster and this type of partition is present for instance in the k-means algorithm. On the other hand, a fuzzy partition allows one data point to exist in more than one cluster with a different degree [25], [31], [44].

- K-Means - The k value (number of clusters) is defined and k points are chosen at random as cluster centers. All data points are assigned to their closest cluster center according to some metric and the centroid of each cluster is calculated. This process is repeated and for each iteration the centroid position will change due to re-calculation by averaging all the points in the cluster. The relocation of the data points and the computation of each clusters' centroid is stopped when almost no change is observed or when the process is below a defined threshold [12], [17], [31]. Advantages: simplicity, efficiency and high computing velocity for large datasets [24]–[26], [33], [35], [42]. Limitations: The final clusters are sensitive to the initial cluster centers [24] and this method requires the knowledge of the k-value (number of clusters) [4], [25].
- Mini Batch k-means - It is a variant of K-means method and uses mini-batches, which are subsets of the original data, arbitrarily sampled in each iteration and define a local solution. For each data point, each centroid is updated by averaging the sample and all previous samples assigned to that specific

centroid. Advantages: minimize time complexity compared to the original method. Limitations: This algorithm outputs less accurate results than the original k-means method.

- Affinity Propagation - This method clusters a dataset by sending messages between pairs of samples until they converge. These messages are scalar values and represent the suitability for one data point to be the exemplar (representative) of the other, which is updated in response to the values from other pairs. This update procedure continues until convergence and the final clustering is obtained. Annex A of this document includes more information for a detailed analysis of this clustering method, such as all parameters and python code used in this work. Advantages: for small or medium sized datasets, choosing the number of clusters based on the dataset provided [56].

Hierarchical Class Is based on dataset division into a hierarchical decomposition, creating a tree shaped structure. This hierarchical division can be performed in an agglomerative ("bottom up") or in a divisive method ("top down"). The agglomerative technique classifies data objects in individual clusters and then merges clusters into wider groups based on criteria defined by the user. This approach ends when all data objects are in the same cluster or when certain conditions are satisfied. The divisive method is the opposite of the latter because all data objects are defined as a single cluster and then they are separated into different groups, based on specific criteria [25], [31], [44].

- Spectral Clustering - This method involves complex calculations where the first step consists in the similarity matrix construction based on the data objects, followed by the associated Laplacian matrix computation. The eigenvalues calculated from the Laplacian matrix will be finally clustered with k-means [9], [31]. Limitations: Requires the K-value initially and it is appropriated for only few clusters. Advantages: Very efficient for sparse datasets up to several thousands and analyses data as a graph partitioning problem without any supposition on the shape of the clusters.
- Ward - Clusters the dataset with an agglomerative approach. The Ward technique is based on the minimal variance criterion and two clusters will be agglomerated into one when a defined value is not reached. Otherwise, those clusters will be apart [31]. Limitations: Requires the K-value initially.

Model-based Class Defines a model to assign to each cluster and tries to fit all data objects to the specified model [31].

Density-based Class In this method, a cluster grows up until its neighborhood density is lower than a defined threshold. The density refers to the present number of data objects [31].

- DBSCAN - Clusters are areas of high density and are separated by regions of low density and therefore can express any shape in opposition to the k-means algorithm output. A core data point belongs to the regions of high density and therefore each cluster is a set of core data points. Every non-core data point that presents a large distance from any core sample is defined as an outlier. Advantages: It finds the number of clusters automatically.
- Mean Shift - It is a versatile algorithm which estimates the global density reducing computational complexity. This approach studies the feature's space as a probability density function in which the local maximum corresponds to the formed clusters. For each data sample, a window around this maximum point is defined and the sample mean is computed. For each iteration the window is shifted to the mean and this procedure is repeated until a defined value is reached [6]. Advantages: This procedure defines the number of clusters automatically and does not constrain to the clusters shape.

Grid-based Class This method divides the space object into a number of cells. This division forms a structure in a grid shape where clustering techniques can be applied [31].

The clustering distance metric it is also an important component for clustering algorithms. Several distance measurements may be applied, which will influence the final results [31], [44] such as the Minkowski Distance, expressed by:

$$d[k,l] = (\sum_{i=1}^{d}(|x_{ik} - x_{il}|)^s)^{1/s} \text{(eq. 3)}$$

where s represents the order of the function and depending on its value different names for Minkowski distance arise: for s=1, it is called Manhattan or City-Block distance, in which the distance between two data points equals the sum of the difference of their coordinates (for a horizontal or vertical move in adjacent cells it is 1 unit and for crossed move, it will be 2 units [33]). On the other hand, for s=2, the distance is called Euclidean and it is the minimum sum of squares of its coordinates [20].
All clustering methods referred above will be used in this thesis with the Euclidean distance to find the most suitable techniques for PA recognition.

2.3.3 Procedures for K-value Estimation

If some clustering algorithm being used requires the previous knowledge of the number of clusters (k-value) such as k-means and Ward methods, one solution may be to try different possibilities and check which is the most suitable k-value. To find the k value automatically presents a major difficulty and one solution would be the one that minimizes the total squared distance of all points to their cluster center [24], [25]. The following list is a small approach of the most used K-identifying procedures.

Gap statistic - The Gap statistic performs a comparison between the resultant value from the square error criterion, W_k, and a reference distribution with no obvious clustering ($E_n^* log(W_{Kb})$) such as an uniform distribution. Gap statistic computation is represented as:

$$Gap(K) = E_n^* log(W_{Kb}) - log(W_K) \text{ (eq. 4)}$$

where:

$$W_K = \sum_{k=1}^{K} \frac{1}{2n_K} D_K \text{ (eq. 5)}$$

and

$$D_K = \sum_{x_i \in C_K} \sum_{x_j \in C_K} ||x_i - x_j||^2 \text{ (eq. 6)},$$

which represents the sum of intra-cluster distances between points in the cluster C_K containing n_K samples.

To compute $E_n^* log(W_{Kb})$, the average of B copies is computed. Each of these copies is generated by the reference distribution and will show a certain standard deviation, sd_k, used in:

$$s_K = sd_K \sqrt{1 + \frac{1}{B}} \text{ (eq. 7)}$$

Finally, the k-value is defined for the smallest K found, such that

$$Gap(K) \geq Gap(K+1) - s_{k+1} \text{ (eq. 8) [7], [38]}$$

Elbow method - The "elbow" method offers a solution based on intra-cluster variance. The elbow method shows the percentage of variance explained by the ratio of clusters and its number. Initially, this curve increases sharply, reaching a plateau, which corresponds to the percentual difference between different k-values. This value corresponds to k-value [52].

Bayesian Information Criterion -The Bayesian Information Criterion (BIC) is a statistical criterion for model selection, which may also be used to unravel clustering problems for finding the number of clusters [49] . The authors from [49] have proposed a method to identify the number of clusters through the knee point detection of the resulting BIC curve, represented by:

$$BIC = L(\theta) - \frac{1}{2} m log(n) \text{ (eq. 9)}$$

where $L(\theta)$ is the log-likehood function of the accelerometry data , m is the number of clusters and n is the size of the dataset.

2.3.4 Clustering Performance Evaluation

The final and most important goal of clustering is to obtain new information, knowledge, from the original dataset and it is crucial a judicious and careful assessment of the final results. For clustering validation it is important to be aware that different parameters or approaches may result in different clustering results, even with the same algorithm and initial dataset. So, it is important to adopt initially certain criteria to get the results as objective as possible [31].
There are two different evaluation types: known and unknown ground truth, with and without knowledge about the clusters' number, respectively [31]. In the first, after a clustering performance evaluation, it is possible to assess if the separation of the data is similar to the available ground truth set. In this situation it is important to create a data annotation as a ground truth, such as manual annotation in each activity performed in a given moment. With an unsupervised learning, a data annotation is required to evaluate the performance of the adopted algorithm and to verify the labeled results [28], [35], [44]. For clustering performance evaluation two methods were used in the present work:

Hungarian Assignment - With two solution sets, the predicted labels and the ground truth set, it is possible to measure the distance between both for label vector comparison. Unlike the supervised learning process, the labeling of the predicted clusters on the data set must correspond to the ground truth available only up to an unknown permutation of the indices. If two partitions of the data set are equivalent but its labelings are represented in a different form (for instance, for k=2, in the first partition the cluster labeled 1 is labeled 2 in the second one and vice-versa), there is an ambiguity. This problem is inherent to the unsupervised learning due to lack of information about data labeling. To overcome this ambiguity the labelled indices in one predicted solution are permuted in order to increase the agreement between the two solution sets under comparison. This ambiguity can be minimized through the Hungarian method with a matrix construction based on the labels and the ground truth similarity. This evaluation method measures the fraction of differently labeled points through the diagonal of the resulting matrix [29].

Adjust Rand Index - No conjecture is performed on the cluster arrangement and this technique can be used in an unsupervised environment, ignoring permutations. With a ground truth class assignment as C and with k as the number of clusters, we have:

$$RI = \frac{a+b}{C_2^{n_{samples}}} \text{ (eq. 10) and } ARI = \frac{RI-E[RI]}{max(RI)-E[RI]} \text{ (eq. 11)}$$

where a is the number of pairs of data points that are in the same cluster in C and in the same cluster in k and b is the number of pairs of data points that are in different clusters in C and in different clusters in k. $C_2^{(n_samples)}$ in expression (eq. 10) is the number of possible pairs in the dataset. Nevertheless, if the k-value is close to the number of existing data points, it is possible to obtain a Rand Index (RI) value close to zero. Therefore, we can adjust the index to the ARI value, shown in (eq. 11) [9], [35].

3

Acquisition Systems and Procedures

The HAR systems developed and studied in the present work with a dependent subject approach are based on the processing of several ACC datasets. The analyzed data is formed by three databases: the online available PAMAP database, the FCHA database acquired from 9 volunteers in the Champalimaud Foundation and the clinical database.

3.1 Foundation Champalimaud Human Activity database

Daily routine activities require many complex movements and positions and this complexity diverge according to each subject, namely in terms of velocity and time duration. This complexity makes signal acquisition from ACC sensors even more challenging if these physical activities are carried out in a non-controlled environment.
The algorithm created in this thesis aims to recognize and distinguish all activities present in the protocol, which is formed by 3 types of postures of static nature, and 6 movements, of cyclic nature, listed on table 3.1.

All 9 tasks shown in Table 3.1 were performed by 9 volunteers with an age range from 23 to 44 years old. Tri-axial accelerometric sensors were located on the waist and on the wrist (figure 3.1) with a frequency sampling of 800 *Hz* and a resolution of 16 bits. However, only the waist sensor was used for data analysis. The ACC data acquired in this protocol forms the Foundation Champalimaud Human Activity (FCHA) database. Data acquisition was carried out in the Champalimaud Foundation, with the OpenSignals software [51]. The collected data was initially saved in h5 format and then analyzed and processed with the Python language [35].

Table 3.1: Set of activities performed by 9 volunteers in order to collect its ACC data and to create the FCHA database.

Postures	Movements
	Walking
Standing	Running
Sitting	Ascending Stairs
Lying Down	Descending Stairs
	Cycling
	Jumping

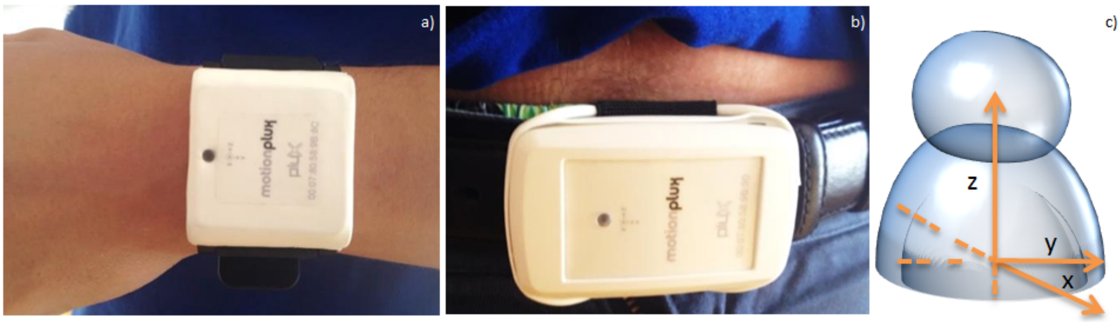
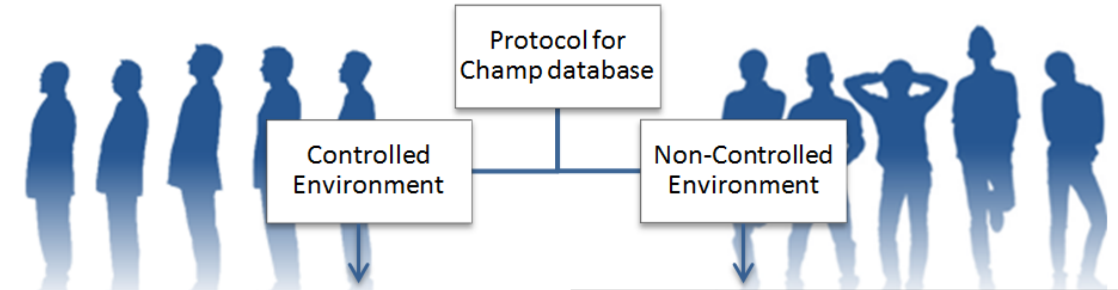


Figure 3.1: a) Wrist tri-axial accelerometer; b) waist tri-axial accelerometer and respective acceleration components' directions; c) Axis orientation of the waist sensor relative to the volunteer's body.

The adopted protocol is composed of two approaches, a controlled and a non-controlled environment stage, in order to test and exploit the algorithm's performance in different acquisition conditions (Figure 3.2). In the first part, several activities such as standing, sitting, walking, running and lying down were performed with a pre-defined order and time, excluding the ascending and descending stairs tasks. The walking and running activities were carried out in a crosswalk with pre-defined velocities (of 4 km/h and 10 km/h respectively). In the non-controlled environment, a list of activities is defined and the volunteers could choose the activities from the list and the order which they wanted to perform them. Task repetitions were also allowed. In this approach the speed and duration time were chosen by the volunteers in order to create as much as possible a realistic acquisition environment, similar to a daily routine. The activities are listed in Figure 3.2.

Figure 3.3. shows the acquired accelerometry signals with the first stage of the adopted protocol, the controlled environment.



Label	Movement	Time (s)
0	Standing	60
1	Sitting	
2	Walking (with crosswalk)	
3	Running (with crosswalk)	
4	Lying Down	
5	Ascending Stairs	Variable
6	Descending Stairs	

Label	Movement	Time (s)
0	Standing	Variable
1	Sitting	
2	Walking	
3	Running	
4	Lying Down	
7	Cycling	
8	Jumping	

Figure 3.2: Recorded Movements and respective labels and time duration for each protocol. The set of signals acquired from this protocol is called FCHA (carried out in the Champalimaud Foundation) database.

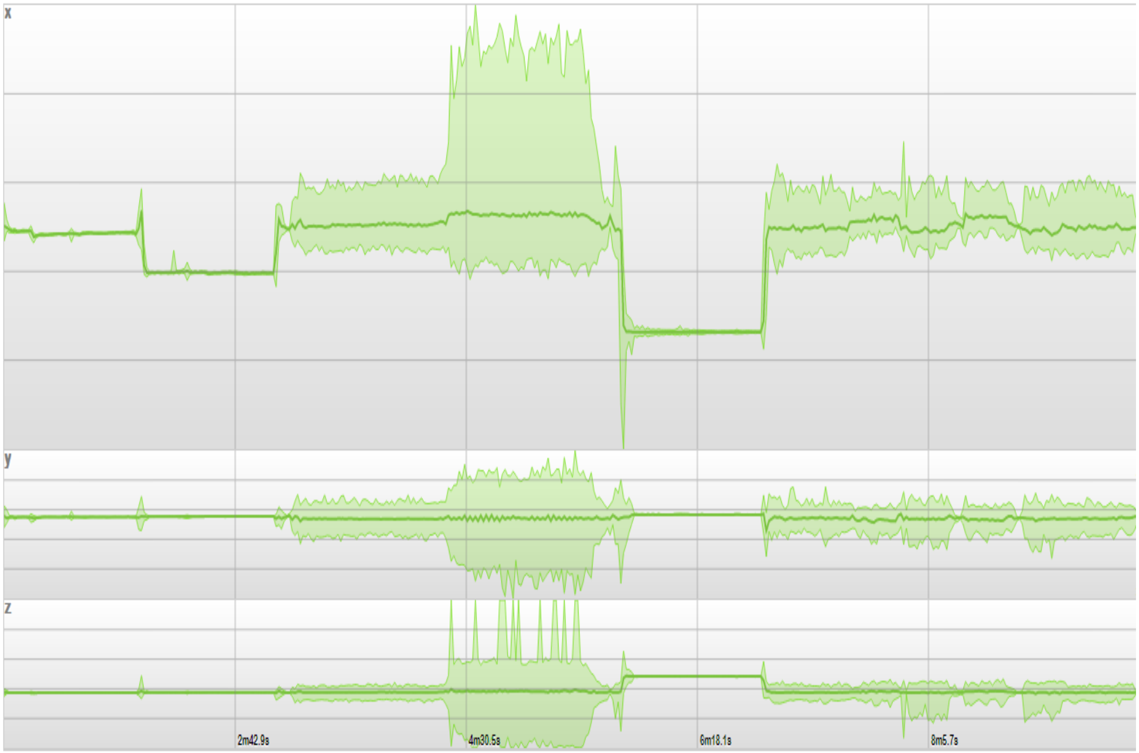


Figure 3.3: Representation of all three components (x,y and z) of the acquired accelerometry signals in the first stage of the protocol, the controlled environment, in the OpenSignals software [51]. The accelerometer signals refer to the subject01s' activities execution.

3.2 PAMAP database

The ACC signals database of this study also included the PAMAP database, available at [45] in order to verify the created HAR system performance with acceleration data at different resolutions and sensor locations. The PAMAP signals were acquired with a sampling frequency of 100 *Hz* and a resolution of 13 *bits*. The same 9 activities were selected from those presented in the PAMAP data, listed in the table 3.2. The data was acquired from 8 volunteers with an age range 24-32 years and the 3D-accelerometric sensor used was placed on the chest. All movement tasks were performed at each subject's pace, in order to acquire data on the most realistic condition as possible [45]. However, task duration was defined as presented on table 3.2 and for this reason the PAMAP database is classified as data acquired in a controlled environment.

Table 3.2: Movements and respective labels and time duration from PAMAP database.

PAMAP Database		
Label	Movement	Time (seconds)
0	Standing	180
1	Sitting	180
2	Walking	180
3	Running	180
4	Lying Down	180
5	Ascending Stairs	60 (twice)
6	Descending Stairs	60 (twice)
7	Cycling	180
8	Jumping	120

The ACC data from the FCHA database corresponds to data acquired in two different environments: a controlled and a non-controlled environment as referred before. Furthermore, ACC data presents different resolutions with the PAMAP inclusion. This data variability allows analysis and adaptability of the developed framework in different situations.

3.3 Clinical database

There is a third database used in this work in order to study the framework application in a clinical environment. This database was provided by the Human Movement Analysis Lab from the Politécnico de Setúbal with 120 *Hz* as sampling frequency. The ACC signals present in this clinical database correspond to the walking task from 9 asymptomatic and to 9 symptomatic volunteers. During this protocol, the sensor was placed in the CoM. Symptomatic patients who participated in the study showed different degrees of neurological dysfunction as a consequence of CVA and other health complications such as diabetic peripheral neuropathy and lymphoma. The dysfunction score presented in Table 3.3 did not reflect just the walking ability but also other parameters such as: the reaction speed, the impairment of reasoning, speech and of other body parts. Therefore, this score was not be used as ground truth.

This third database was used by the framework developed in the present work as first potential clinical application showing its usefulness to evaluate the quality of motion.

Table 3.3: Nine asymptomatic (N0-8) and nine symptomatic (P0-8) volunteers and respective dysfunction type and impairment score. The ND parameter refers to not defined.

Subject	Dysfunction Type	Impairment Score
N0-8	None	ND
P0	CVA	21
P1	CVA	22
P2	CVA	17
P3	CVA	13
P4	Diabetic peripheral neuropathy	20
P5	CVA + lymphoma	22
P6	CVA	15
P7	CVA	10
P8	CVA	14

4

Framework for HAR systems

In this section two distinct stages related to the clustering step in the proposed framework for activity recognition are described, namely: base clustering and meta clustering corresponding to procedures to be carried out before and after the clustering application, respectively.
Several parameters were studied in order to improve as much as possible the clustering results, such as: filtering and window-length and feature free parameters. This study also includes the most relevant feature identification for activity recognition to reject those that reflect irrelevant or even redundant information and to reduce computational demands. These studies are important for feature extraction reflecting the signal's characteristics and information transmission to the clustering algorithm composing ACC data. All these studies are explained in detail in the section 4.1.
Finally, the last section (vide section 4.2) highlights the importance of a ground truth availability to compute several algorithms, such as the Hungarian Assignment and the Hidden Markov Models application.

4.1 Base Clustering

The segmentation stage belongs to the base clustering stage and it will affect the observer perspective regarding the acquired signals according to the window length used. After segmentation, feature extraction consists on the search for the values representing relevant information about raw data. This information can represent the shape of the signal or how it behaves over time or frequency, such as for instance the mean or the standard deviation values.

Feature Selection is also an important step to implement after feature extraction because it influences the final outcome. This stage is essential when there is a massive amount of information with a vast quantity of features, where some of which are not completely essential for the clustering step or provide redundant information [32], [44]. Therefore, with feature selection procedure the time and computational complexity are reduced and the accuracy of clustering performance increases. Features must be selected according to the acquired data and the aim of the investigation. An ideal feature type presents a large variation between different clusters and a small one between data belonging to the same cluster [35], [36], [44].

4.1.1 Basic Features and Wavelets Application

It is possible to group features according to different parameters, such as time, statistical and frequency domains. In the temporal domain, there are several known features, already used in accelerometry, such as zero crossing rate, autocorrelation and variance. Other features such as mean, standard deviation, histogram and root mean square are some examples computed from the statistical domain, that do not depend on the frequency or the time variation of the signal [23], [28], [35], [50], [57]. In the frequency domain, it is possible to compute the maximum and median frequencies, the fundamental frequency, the power bandwidth and the Fast Fourier Transform (FFT) coefficients, through the FFT of the original signal.
In the frequency spectrum analysis through FFT the original data is compared to a family of sine functions at harmonically related frequencies. Nevertheless, the FFT does not provide information about the time at which these frequency components occur, which leads to the need for a tool that allows us to analyze the signal on both domains.
A wavelet is a specific technique for the time-frequency domain and it is an irregular waveform of limited duration with zero average value on the time domain. A wavelet analysis decomposes the original signal into elements. It allows a visualization of the frequency content over time and consequently a better transient event description in an accelerometry signal [16], [19], [39]. The decomposition process may be computed at different levels. A single level decomposition applies two complementary low-pass and high-pass filters to the input signal, which outputs the approximation and the detail coefficients, respectively. The approximation coefficients reflect the main characteristics of the signal and these values are used as feature coefficients in this work [30].

4.1.2 CUIDADO Features and Log Scale Power Bandwidth Implementation

There are other features, called the "CUIDADO features", applied for the first time for audio signals by G. Peeters in [47] and can be applied and useful for accelerometry studies. The variance and temporal centroid are two examples of features used in [47], belonging

to the temporal domain applied in the present work. The variance computes the signal's variance over time and the temporal centroid corresponds to the time averaged over the signal's energy wave.
On the other hand, in the spectral domain, there are several features such as: the total energy which estimates the signal power at a given time; the spectral centroid that may be described as the barycenter of the distribution of frequencies and their probabilities to occur; the spectral spread and skewness which are the variance and the measure of the asymmetry of the distribution defined above, respectively; the spectral kurtosis which measures the distribution's flatness around the mean value; the spectral slope and the spectral decrease, related to the decrease of the spectral distribution through its linear regression computation; the spectral roll-off point which is the frequency where 95% of the signal energy is contained below of this value; and the spectral variation, computed from the normalized cross-correlation between two consecutive amplitude spectra (represented as spectrum variation) [47].
The MFCC belongs also to the CUIDADO set and is widely used for speaker recognition. For a more accurate analysis of all accelerometry signals a meticulous analysis in the frequency and amplitude domains would be mandatory. In the present work, the lower frequencies were studied in more detail than higher frequencies through logarithmic scales. This study was inspired by the audio spectrum and the mel-scale which ultimately lead to the Log Scale Power Bandwidth (LSPB) feature creation. The inputs for the LSPB are the motion data (Algorithm 1).

Algorithm 1: Log Scale Power Bandwidth

Input `Accelerometry Signal`

Output `Log Scale Power Bandwidth coefficients`

This algorithm concerned five stages:

1. The first stage was the preemphasis of the signal in time domain. This stage filters a data sequence (the input segment signal) through a digital filter that increases the magnitude of the signal at high frequencies, thereby improving the signal-to-noise ratio and using with a pre-emphasis factor of 0.97;

2. The second step referred to the framing which divides the input data into a set of 3 (M) frames, each of these with 256 (N) samples;

3. Next, the conversion of the signal segment into the frequency domain was carried out through the Fast Fourier Transform application. However, whenever a finite Fourier Transform is applied and if the start and end of the finite data do not match,

there will be a discontinuity in the signal. In this case, undesirable high-frequencies will show up in the Fourier transform. Therefore, a windowing stage must be computed using a Hamming window to ensure both ends match up;

4. 133 *Hz* and 3128 *Hz* were chosen as the minimum and maximum frequencies for the filter bank creation. The triangular filter bank is a set of triangular overlapping windows in the range 133-3128 *Hz*. This set of triangular filters, shown in figure 4.1, was spaced linearly at lower frequency, below 199 *Hz*, and logarithmic spaced above 199 *Hz*;

5. The triangular filter bank was applied to the data resulting from the third step. Finally, the algorithm took the log (in base 10) of the powers at each frequency and returned the Log Scale Power Bandwidth coefficients as the amplitudes of the resulting spectrum.

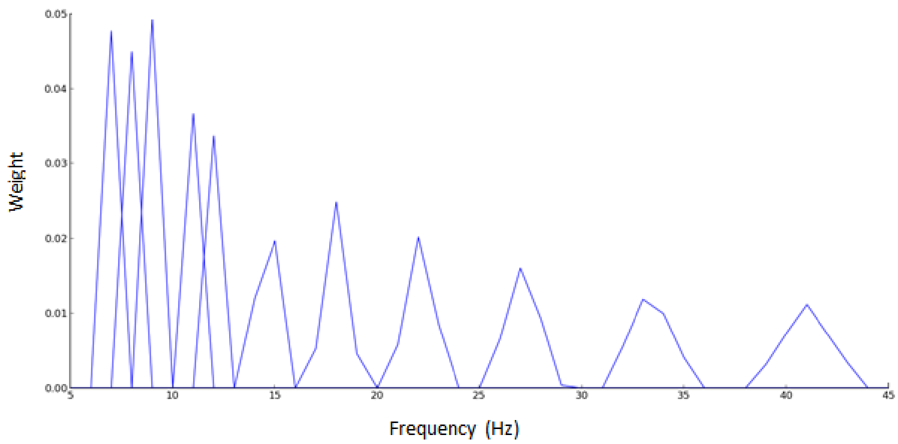


Figure 4.1: Representation of 11 triangular filters belonging to the filter bank used in the Log Scale Power Bandwidth algorithm.

These new features applied to accelerometry, instead of audio signal alone, might contribute to the discovery of important movement characteristics never detected before, such as the Log Scale Power Bandwidth. This new feature is created and applied for the first time to ACC signals in this thesis and revealed promising results (vide Figure 5.7 from chapter 5).

4.1.3 Filtering and Window-Length Influence

After signal acquisition, most of the literature applies a High-Pass Butterworth filter to separate the low frequencies from the high frequencies and to remove gravity acceleration [28], [35], [48]. In the present study and similarly to [35], filtering was applied to the three components accelerometry, x, y, z. The fourth acceleration, the total acceleration, is computed with all three filtered accelerations and expressed by:

$$Acc_{tot} = \sqrt{x^2 + y^2 + z^2} \text{ (eq. 12)}$$

Table 4.1: List of all features used in the present work and respective domains and number of output coefficients for each acceleration component: x, y, z, and total acceleration ($\sqrt[2]{x^2+y^2+z^2}$); [1] Refers to all traditional features already applied in accelerometry [35]; [2] Refers to the CUIDADO features used in audio recognition; [3] Refers to new features created and implemented in this work.

Feature Type	Number of Output Coefficients (for each acceleration component)
Statistical	
Skewness [1]	1
Kurtosis [1]	1
Histogram [1]	10
Mean[1]	1
Standard Deviation[1]	1
Interquartile Range[1]	1
Time	
Root Mean Square[1]	1
Median Absolute Deviation[1]	1
Zero Crossing Rate[1]	1
Pairwise Correlation[1]	3 (in total)
Autocorrelation[1]	1
Temporal Centroid[2]	1
Variance[2]	1
Frequency	
Maximum Frequency[1]	1
Median Frequency[1]	1
Power Spectrum[1]	2
Fundamental Frequency[1]	1
Power Bandwidth[1]	10
Log Scale Power Bandwidth[3]	40
Total Energy[2]	1
Spectral Centroid[2]	1
Spectral Spread[2]	1
Spectral Skewness[2]	1
Spectral Kurtosis[2]	1
Spectral Slope[2]	2
Spectral Decrease[2]	1
Spectral Roll-off [2]	1
Time-Frequency	
Wavelets [2]	20

For a better understanding of the filtering influence on the clustering performance, a study related to its application was carried out and presented in chapter 5, section 5.4.
The next step involves signal segmentation for the subsequent feature extraction. The ACC signal is divided sequentially into a finite number of segments with a predetermined length, called windows. Each segment is used as an input for the clustering algorithms, i.e., each window will present its own features.
The choice of the segment length is based on the resolution and information required in each window. The smaller the segment length, the higher the resolution, however a larger window covers a greater amount of information regarding a particular physical activity or position.

A signal segment should contain only information concerning one physical activity and this must be taken into account for a proper window length selection to avoid ambiguous labeling. For this study the FCHA database was used and several window lengths in multiples of its sampling frequency (800 *Hz*), from 800 to 4000 samples were tested in order to find the most suitable length for the developed framework. The results are presented in chapter 5, section 5.1.

4.1.4 Free Parameters

A smaller set of features were chosen from those available for feature selection (listed in Table 4.1). Similarly to [35], all features used were gathered in a JavaScript Object Notation (JSON) file, which gives all pertinent information such as: feature description, imported libraries, input parameters and free input parameters. Table 4.1 shows all features extracted from each component of the accelerometry signal with the total of 423 features.
The free input parameters are pre-defined values used in some feature algorithms and will influence the respective feature calculation, such as the desirable number of bins for the histogram computation. The free input parameters used in some feature algorithms such as the histogram, cepstral coefficients and power bandwidth, are those with the highest ARI accuracies in [35].
Wavelet decomposition is part of the development and exploration of the present work, which also requires free parameters. The developed algorithm computes the wavelet decomposition of the input signal using the level decomposition value as a free input parameter. Therefore, 15 decomposition levels were studied in order to find the best decomposition level for activity recognition (vide chapter 5, section 5.3). The level with the highest ARI accuracy will be adopted as a free parameter for the wavelets algorithm.

In addition to all the aforementioned information, the JSON dictionary also outputs the number of extracted features from each feature type, the name of the function, the source code and references.

4.1.5 Normalization and Forward Feature Selection

After free parameter identification, it is necessary to select the best features for the HAR system. Since the feature computation is a time consuming and computationally heavy task, it is recommended to identify the most suitable features for activity recognition.
Feature's normalization to zero mean and unit variance is adopted before the generation of any selection of features. This stage avoids unreliable results originated from the influence of different magnitudes on all the features.
The selected features are directly related with the information extracted from the ACC signals which allows data organization inside each cluster by the clustering algorithms.

The protocol for feature selection is based on the "Forward Feature Selection" protocol and aimed to select 10 features at most (vide chapter 5, section 5.6). This study is presented in figure 4.2 and is described by the following steps:

1. Elimination of the redundant information: Correlation between all features and removal of the redundant features. The second feature was removed when two features showed correlation values greater than or equal to 0.98. The resulting set from this correlation and elimination stage was called A;

2. Selection of the best fitting features: 20 features with the highest ARI value were chosen among the set A. This new set of 20 feature types was named set B. The feature type from B with the highest ARI accuracy was collected to a new group, called set C. This feature type was removed from set B and added to the set C. Next, the feature selection algorithm combined the set C with the existing feature types from set B and the combination with the highest ARI value is identified. The new feature type belonging to B and to the identified set was collected by C. In each iteration, a new feature was removed from B and collected by C. This iterative procedure was repeated until C shows the best 10 feature types;

3. Saving the final results: The algorithm finished the procedure and saved the name and the ARI accuracies of the 10 best feature set.

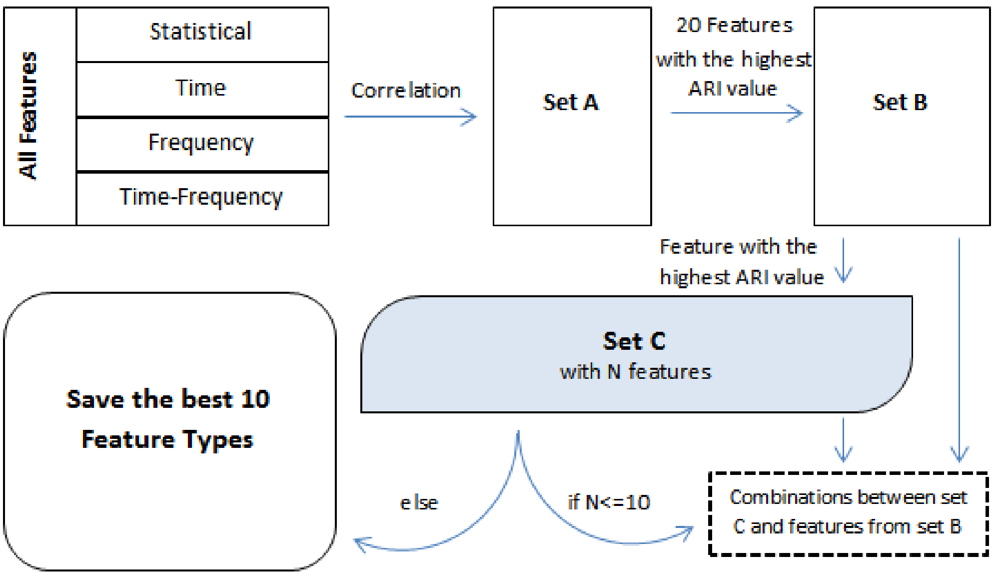


Figure 4.2: Overall structure of the Forward Feature Selection algorithm. All features from all four domains (statistical, time, frequency and time-frequency) are used in this algorithm. The set A is formed by all features with a correlation value lower than 0.98. The set B is formed by 20 features with the highest ARI values. The set C is formed by N features resulting from the combination process between the feature with the highest ARI value and other features from set B. In each iteration, the best combination will be collected by C until N=10.

4.2 Meta Clustering

After feature extraction several unsupervised learning methods are used for activity recognition such as K-means, Mini Batch K-means, Affinity Propagation, Mean Shift, DBSCAN, Spectral Clustering and Ward. All these clustering algorithms explained in chapter 2 structure data such that the inter-cluster similarity is low and intra-cluster similarity is high according to the respective metric.
In the present work, the best clustering methods are identified for each existing clustering class in order to minimize the number of clustering techniques to use and to reduce the computational demands (vide chapter 5, section 5.2).

After the clustering stage, some studies were carried out such as the HMM application in order to improve as much as possible the performance evaluation. All the Meta clustering procedures are supported by an annotation document for the ground truth availability.

4.2.1 Signal Annotation

The annotation document concerns all the labels and time interval logs that delimit each activity performed by a given volunteer. This annotation is a crucial task to select only the activities segments properly. Therefore, segmentation is not applied to the entire signal. Only physical activities are taken into account for the feature extraction and not the transitions between physical tasks.
The initial and final time of each activity was annotated in a JSON file created for each acquired signal. In addition to the annotation task, the present work added an extra stage where motion series of all volunteers were videotaped for label and time interval validation (figure 4.3).

Figure 4.3: Frames of the subject08's videotape, performing four tasks from the protocol: running, lying down, climbing stairs and cycling.

4.2.2 Hidden Markov Model

HMM is a special case of Markov model, as referred in chapter 2, where the observation is a probabilistic function of the state. In the HMM case, the Markov process shows unobserved (i.e. hidden) states and there is no knowledge regarding the existing states and transition probabilities between them [15], [21], [55].
HMM was applied to the clustering results to achieve better results in activity recognition.
The implemented algorithm uses the ground truth (true labels) by collecting the frequency of transitions between different activities/states, in addition it also defines the initial state of a given sequence of activities from the most frequent initial state. Recorded frequencies are then converted to the probabilities of existing symbols and state sequences. Finally, the testing stage estimates the most probable sequence of hidden states based on the trained model and on the Viterbi algorithm.
In the present work, HMM uses the labels from the clustering stage as a test set and the annotated data (ground truth) as training set. The influence of the HMM application was studied and the ARI results and discussion are shown in chapter 5, section 5.5.

4.2.3 Hungarian Accuracy

The ground truth availability is also crucial for the performance evaluation in an unsupervised approach considering that the predicted labels are verified accordingly with this information. For clustering performance evaluation, this work presents the ARI and the HA methods. The HA method outputs a value between 0 and 1, where higher values (close to 1) correspond to data organization closer to the ground truth. For values close to 0 the clustering method does not identify the clusters correctly and assigns multiple clusters to the same activity.
It is possible that the labels provided from the clustering stages present a similar organization to the ground truth but with different values, i.e., the label 1 from the ground truth match value 2 from the predicted labels and vice versa. Therefore, to compute the Hungarian accuracy some steps, detailed in pseudo-code in Algorithm 2, are required which can be explained with some basic concepts from decision analysis.

The process was initiated by computing the confusion matrix with the ground truth and the resulting labels from clustering and HMM algorithms. This confusion matrix represents the profit matrix named matrix A. The higher the value present in each line of the matrix A, the higher is the compatibility between those labels and the correspondent column labels.
The next step involved the construction of the cost matrix B, from A in order to apply the Lower Cost method. This method aimed to maximize the compatibility between labels from columns and rows of B in order to assign each label from the ground truth to the

Algorithm 2 Hungarian Assignment

```
confusion_matrix ← confusion_matrix(true_labels, predicted_labels)    ▷ Confusion matrix computation (matrix A)
cost_matrix = cost_matrix(confusion_matrix)                            ▷ Cost matrix computation (matrix B)
new_labels ← compute(matrix)                                           ▷ Lower Cost method application
final_matrix ← normalize (confusion_matrix (true_labels, new_labels))  ▷ Final matrix Normalization
return mean(diagonal(final_matrix))                                    ▷ Hungarian Accuracy calculation
```

Figure 4.4: Overall structure of the Hungarian accuracy algorithm. The first and second stages referred to the confusion matrix (A) and the cost matrix computation (B), respectively. Then, the lower cost method and normalization were applied to matrix B. The Hungarian Accuracy value was achieved through the average calculation of the resultant matrix diagonal.

most suitable label from the predicted labels set.
Matrix C is the result of the Lower Cost method application to matrix B. Figure 4.5 represents the matrix B and the matrix C resulting from the stages detailed below.

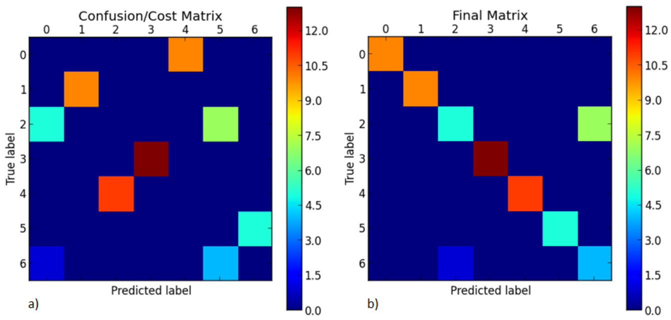


Figure 4.5: Matrixes constructed over the HA method: a) Cost matrix B computed from the profit matrix A which is obtained from the ground truth and predicted labels from the clustering and the HMM stages; b) Matrix C is the result of the Lower Cost method application to the matrix B.

Finally, the matrix C was normalized and the Hungarian Accuracy is computed as:

$$HungarianAccuracy = average(\text{diagonal of the normalized matrix C}) \text{ (eq. 13)}$$

The normalization process converted each value of the diagonal matrix into the ratio between its own value and the sum of all values in the corresponding row. The HA corresponds to the average value of the diagonal on the normalized matrix C.

5

Evaluation of the Performance

Different parameters were studied in order to create the most suitable system for activity recognition and to achieve the best performances. Only the waist-worn accelerometer was used for the ACC signals analysis [35].
The window-length in the segmentation process, the filtering influence, the wavelet level decomposition and the Hidden Markov Model application were tested and the respective results were compiled in the present chapter. Features and clustering methods were also selected according to the resulting ARI values in order to improve the clustering accuracy. In parallel to the unsupervised learning data analysis was also performed with several classification methods, such as Nearest Neighbors, Random Forest and Linear Discriminant Analysis (LDA).

5.1 Window length Influence

The length of the segmentation window is an important decision for feature extraction due to its great influence on the accuracy of the clustering results. Different window length values studied change the signal resolution and the amount of information contained in each signal segment. The most suitable value for segmentation must be defined according to the information chosen to be extracted from each window and the period of all activities present in the data.
Five values in the range 800-4000 *Hz* were tested to identify the most suitable segmentation length, i.e. with the highest ARI accuracy, for HAR systems. This study was carried out with all features implemented from the four domains and with the signals from the FCHA database acquired in a controlled environment. The performance evaluation was

computed from the K-means output. All values used were multiples of the frequency sampling and correspond to time intervals from 1 to 5 seconds.

Table 5.1: K-means performance with ARI algorithm according to each chosen window length. Each result corresponds to the average and standard deviation of the results of 9 individuals, shown in the table in the annex B.1.

Window (samples)	ARI(%)
800	81.1 ± 10.8
1600	81.6 ± 11.3
2400	84.5 ± 10.8
3200	85.3 ± 8.3
4000	87.7 ± 5.6

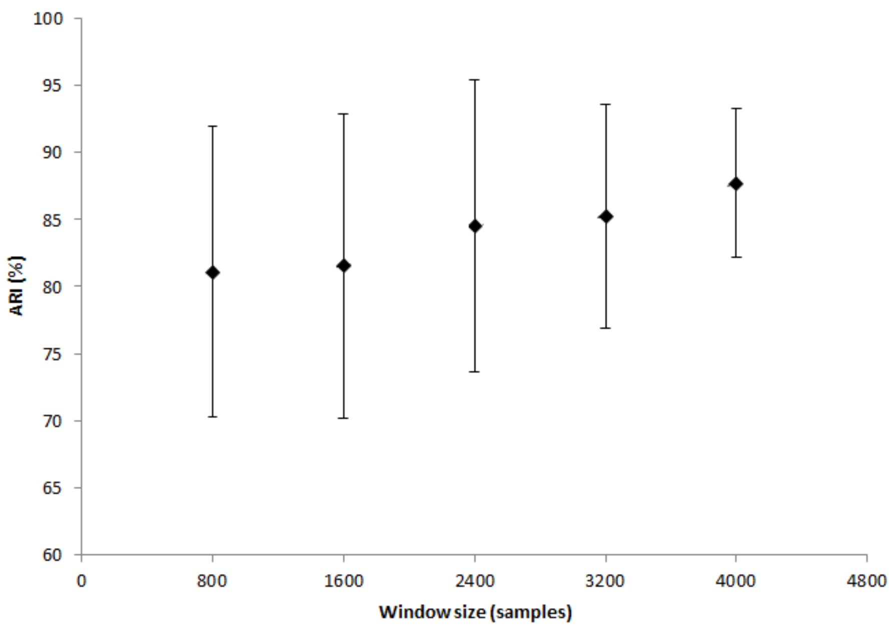


Figure 5.1: ARI accuracy with K-means application according to five segmentation lengths from 800 to 4000 samples. Each point corresponds to the results from the table shown in table 5.1.

The K-means algorithm showed a performance of $81.1\% \pm 10.8\%$ and $87.7\% \pm 5.6\%$ for windows of 800 and 4000 samples, respectively. The window corresponding to the highest ARI value was the 4000 sample. Therefore this value was used in the segmentation process of all studies as the most suitable window length.

It is possible to conclude from the obtained results (Table 5.1) that a smaller window corresponds to a lower clustering performance due to a smaller amount of information contained in each segment.

The studied activities have associated positions and movements similar to each other when analyzed for a short period of time, such as walking, standing and ascending stairs tasks. In a shorter period of time, periodic movements, such as between ascending two

steps of ladders or walking two steps, the volunteer displays a standing position which could easily be confused with the stand still position. Thus, shorter windows belonging to different activities may present a higher similarity that might lead to lower performances.
Furthermore, all windows contain data from only one physical activity in this study and longer time intervals will present more information regarding the respective task. All the referred circumstances may be the reason for higher ARI values with longer segmentation windows.

5.2 Selection of Unsupervised Learning Methods

Most part of the present work involve the study of several clustering techniques such as K-means, Mini Batch K-means, Affinity Propagation, Mean Shift, DBSCAN, Spectral Clustering and the Ward method. It is important to select a smaller group of clustering methods from the original set in order to simplify the ACC studies. The obtained results are shown in the following sections of chapter 5.
The present work aimed to identify at least one method from each existing clustering class. This analysis was carried out with: a window segmentation length of 4000 samples (highest ARI value- vide Table 5.1); all signals from the controlled environment of the FCHA database (vide chapter 3, section 3.1) and with all implemented features (vide chapter 4, Table 4.1). The selected methods were chosen according to different requirements, such as ARI value and time response.
The table 5.2 presents the ARI accuracy according to each unsupervised learning method applied and the respective time response (computational time), in seconds.

Table 5.2: ARI accuracy (%) and time response (in seconds) as a function of different clustering techniques used. Each ARI result shown is the average and standard deviation of the results of 9 individuals, presented in the annex B.2.

Class	Clustering Methods	ARI(%)	Time Response (s)
Partitioning	K-means	87.7 ± 5.6	9.3
	Mini Batch K-means	$84.5 \pm 8,5$	8.7
	Affinity Propagation	$78,4 \pm 6,9$	1.5
Density	Mean Shift	$31,2 \pm 10.7$	2.0
	DBSCAN	78.4 ± 7.0	2.1
Hierarchical	Spectral Clustering	89.1 ± 8.8	4.8
	Ward	86.3 ± 8.7	1.2

From the analysis of the table 5.2 it can be concluded that the Mean Shift presented $31,2\% \pm 10,7\%$ as ARI accuracy and does not show ARI values as high as the other methods. Therefore the Mean Shift algorithm may be less accurate for ACC signals analysis. The DBSCAN was therefore chosen as clustering method from the density class.

Moreover, Mini Batch K-means was considered a familiar K-means with similar clustering results although with an improvement in the time response and slightly worse results in the cluster division. K-means and Mini Batch K-means achieved performances of $87,7\% \pm 5,6\%$ and $84,5\% \pm 8,5\%$ respectively. Thus, K-means is preferred over the Mini Batch K-means.
The Affinity Propagation is another clustering method from the Partitioning class with an ARI accuracy of $78,4\% \pm 6,86\%$. This method although presented a lower ARI accuracy compared with the K-means showed a much lower time response (Affinity Propagation showed a time response of 1.5 seconds and K-means of 9.3 seconds) and it did not require the k-value initially, i.e. number of clusters, which made this algorithm a good choice as well. Overall, the Spectral Clustering, with a performance of $89,1\% \pm 8,8\%$ was better than the Ward method with $86,3\% \pm 8,7\%$. Nevertheless, Ward was chosen from the Hierarchical class due to its shorter time response.

5.3 Wavelet Level Decomposition Influence

Some features used in the present work required input parameters which influence the final results. These input values are called free parameters and are important for the performance of the created framework as can be observed in the final results of the present study.
Wavelets decomposition process is part of an algorithm which computes features from an input signal whose output is formed by the decomposition coefficients, as referred in chapter 4, subsection 4.1.1.
The wavelets process required a decomposition level as input which must be chosen according to the achieved ARI values. Fifteen decomposition levels were analyzed, from 1 to 15, and the respective ARI values are shown in figure 5.3. Only the wavelets decomposition coefficients were used as features on the 4000 sample window length (highest ARI value - vide Table 5.1.). This analysis was performed on the FCHA controlled environment database (vide chapter 3, section 3.1).

Figure 5.2 shows the colormap of the predicted labels from each level decomposition. The decomposition levels from the chosen range show different colors between different activities labels and small color variation within the same label. Moreover, the initial and final time instances of each label must match the same transition moments of the ground truth set. The set with the best defined labels compared to the ground truth was composed by the 7^{th}, 8^{th} and 9^{th} decomposition levels, suggesting that the most suitable level should be within this range.
Analysis of Figure 5.3 confirmed the visual indication and level 8 was the one that best identified the activities present in the ACC data and showed the highest ARI accuracy, with $76.8\% \pm 7,7\%$. All other decomposition levels presented ARI values lower than

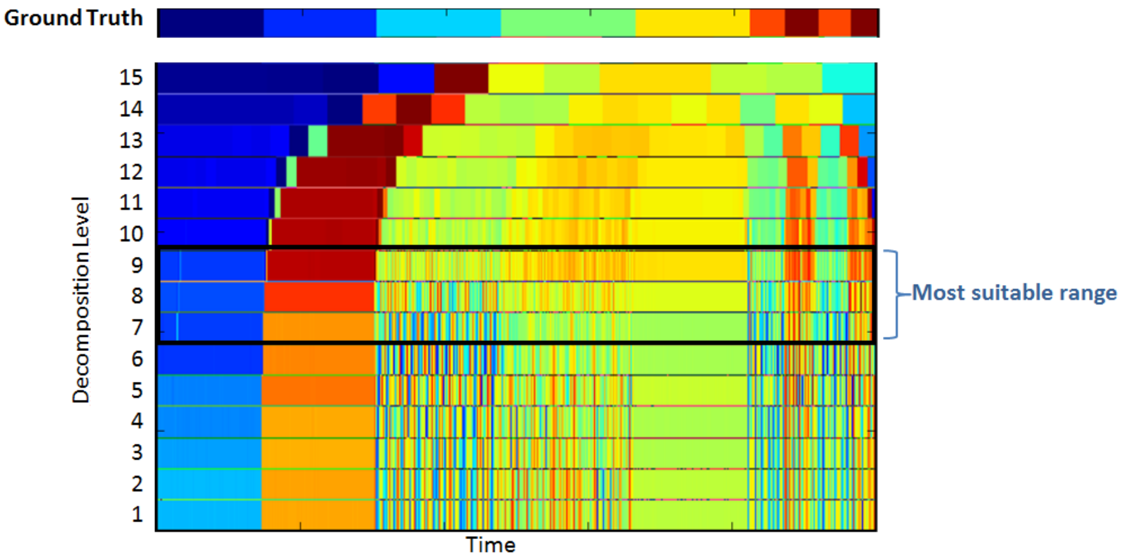


Figure 5.2: Colormap of the ground truth and the predicted labels from each wavelet decomposition level studied, from level 1 to 15. A colormap shows the same color for similar values which means that samples belonging to the same physical activity must show the same color. The most suitable range of decomposition levels for activities recognition is formed by the 7th, 8th and 9th levels.

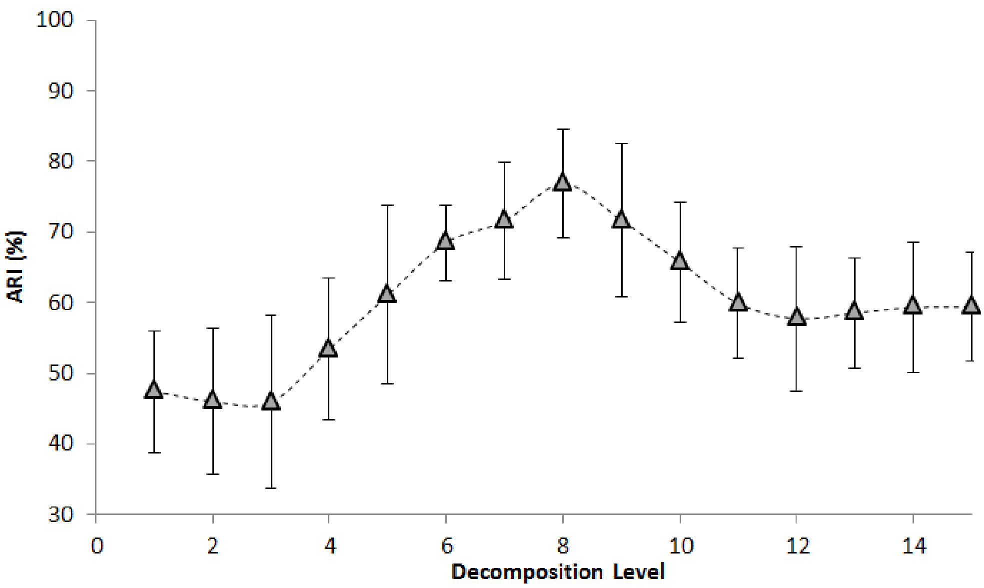


Figure 5.3: Representation of the clustering performances with ARI accuracy according to each wavelet level decomposition applied from 1 to 15. Each point corresponds to the average of the results lodged in the annex B.3.

72%, as shown in Figure 5.3. The chosen decomposition level will be used as a free input parameter in the wavelets algorithm.

5.4 Filtering Influence

The filtering stage, which regards the signal processing step in the revised literature, influences the quantity and quality of the information available to the clustering methods [28], [48].
A High-Pass Butterworth filter was applied with a cut-frequency of 0.25 *Hz* [35] in order to separate the GA and the BA components and analyze its influence in the final performance. The study was performed with the signals from the FCHA database acquired in a controlled environment, with all the implemented features and window length of 4000 samples (vide section 3.1 and Tables 4.1 and 5.1, respectively). All ARI accuracies with and without filtering and its improvement in percentage (%) are shown in table 5.3 and in figure 5.4. The improvement values shown in the present section were achieved by:

$$Improvement(\%) = \frac{100(x_2 - x_1)}{100 - x_1}$$

where:
x_1 - ARI accuracy value for BA component
x_2 - ARI accuracy value for BA and GA components

Table 5.3: Clustering performance with the ARI algorithm with and without filtering stage. Each clustering method result corresponds to the average and standard deviation of the results of 9 subjects, presented in appendix B.4.

Clustering Methods	ARI(%) for BA component	ARI(%) for BA and GA components	Improvement (%)
K-means	64.9 ± 10.0	$87,7 \pm 5,6$	64.9
Affinity Propagation	58.1 ± 10.5	$78,4 \pm 6,9$	48.5
DBSCAN	54.4 ± 9.9	$78,4 \pm 7.0$	52.6
Ward	68.1 ± 10.8	86.3 ± 8.7	57.1

From table 5.3, it can be concluded there was a significant improvement on the ARI accuracy with the addition of the GA components. Improvement results such as 64.9%, 48.5%, $52,6\%$ and 57.1% were obtained for the K-means algorithm, Affinity Propagation, DBSCAN and Ward, respectively.
The improvement with both BA and GA acceleration components, was due to the availability of more information regarding the performed movements at low frequencies. By filter application, the information from the BA component located at low frequencies that may be important for the identification of certain activities is lost. The GA component provides posture information which allows further discrimination on some activities. Furthermore, GA is the acceleration component that allows the distinction of some tasks without movement, i.e., constant over time such as the standing, sitting and the lying down positions. None of the latter positions requires body movement and therefore

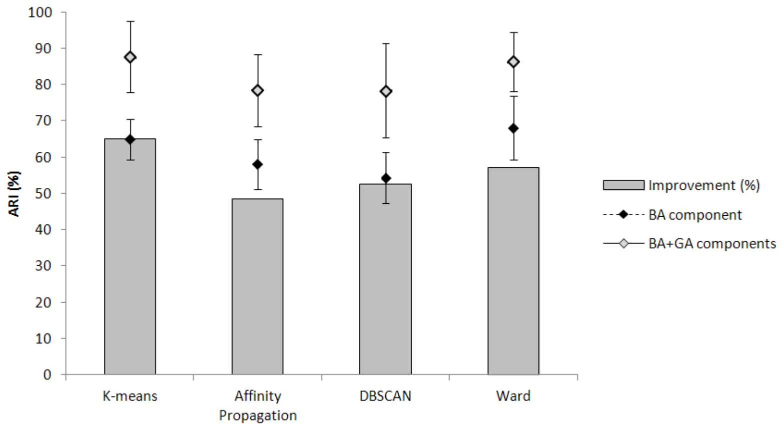


Figure 5.4: Clustering accuracy with ARI algorithm with and without the filtering stage (%) and its improvement (%) in bars shown in table 5.3

the gravitational acceleration is crucial for their identification.
During the protocol, the accelerometer axes vary according to the individual's position while the direction and magnitude of the gravitational acceleration is constant. Figure 3.1 (section 3.1), it is possible to infer that the gravity acceleration and the X-axis of the waist sensor will have the same direction and opposite orientations in the lying position. This layout means that the X component in a lying position will present a negative and higher magnitude from the standing up data.
Therefore, the GA component combined to the BA allowed a significant improvement in the activity recognition (between 49% to 65%), as shown in Table 5.3 and Figure 5.4.

5.5 Hidden Markov Model Application

As previously referred (chapter 4, section 4.2.2), the Hidden Markov Model application may change the resulting labels from clustering according to the existing states, i.e., activities, and its transition probabilities.
All the existing transitions in the test set (predicted labels from the clustering algorithms) with lower probability of occurrence may be a consequence of cluster miscalculation. These transition probabilities were gathered from the ground truth (train data) and all transitions with low probability of occurrence were avoided and replaced by a more likely transition from the analysis process of the test set.
The influence of the Hidden Markov Model application and its improvement (in %) is presented in table 5.4. All implemented features were used in this study (vide Figure 4.1), with 4000 samples segmentation window (higher ARI accuracy from Table 5.1) and with both acceleration components, BA and GA (higher ARI accuracies from Table 5.3). No filters were used in this study. As previously referred only the ACC signals acquired in controlled environments from the FCHA database were used (vide chapter 3, section 3.1).

The improvement values shown in the present section is represented by:

$$Improvement(\%) = \frac{100(x_2 - x_1)}{100 - x_1}$$

where:
x_1 - ARI accuracy value without HMM application
x_2 - ARI accuracy value with HMM application

Table 5.4: Clustering performances through ARI without and with the HMM application and its improvement. All results from the first and second columns correspond to the average and standard deviation of the results from 9 subjects presented in appendix B.5.

Clustering Methods	ARI(%) without HMM application	ARI(%) with HMM application	Improvement (%)
K-means	81.4 ± 9.8	87.7 ± 5.6	33.9
Affinity Propagation	78.0 ± 9.9	78.4 ± 6.9	1.8
DBSCAN	73.9 ± 13.0	78.4 ± 7.0	17.1
Ward	84.8 ± 8.2	86.3 ± 8.7	9.8

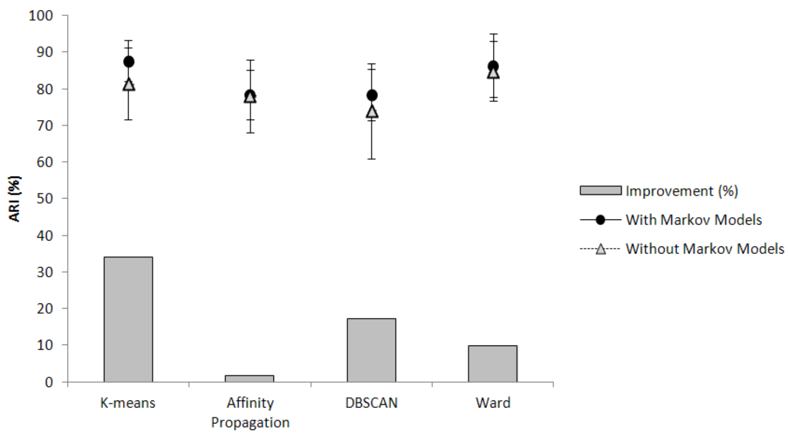


Figure 5.5: Clustering performances through ARI with and without the HMM application (%) and its performance's improvement (%) in bars shown in table 5.4.

The improvement on the obtained performances was 33.9%, 1.8%, 17.1% and 9.8%, for K-means, Affinity Propagation, DBSCAN and Ward method, respectively. These results deserve a special attention because there is a significant improvement for certain clustering methods such as the K-means. Thus, it is important to take into account the HMM application together with the adopted unsupervised learning method. All improvement values were positive and the HMM algorithm does not provide heavy computational demands hence it is applicable to all demonstrated situations.

5.6 Feature Selection

The final studied parameter of the framework for HAR systems was the identification of the best set of features. Two analyses were required for this study: the first involves the search for the best number of feature types for activity recognition and the second test is the identification of the best feature types to be used.
Forward Feature Selection (vide chapter 4 , section 4.1.5) was implemented in this study in order to avoid a process of combinations of features and to minimize the computational demands. As a consequence, parallel computation was not required.
This study used all implemented features from four domains (statistical, time, frequency and time-frequency), segmentation windows of 4000 samples (section 5.1) and K-means, Affinity Propagation, DBSCAN and Ward as the chosen clustering methods (section 5.2). No filters were used in this study and all signals from the FCHA database and acquired in a controlled environment were used in this study.
For each subject, the best 10 features were computed from the Forward Feature Selection Algorithm. Figure 5.6 showed the ARI accuracy (%) for each number of features and for each clustering method in order to select the best number of features to use in HAR recognition.

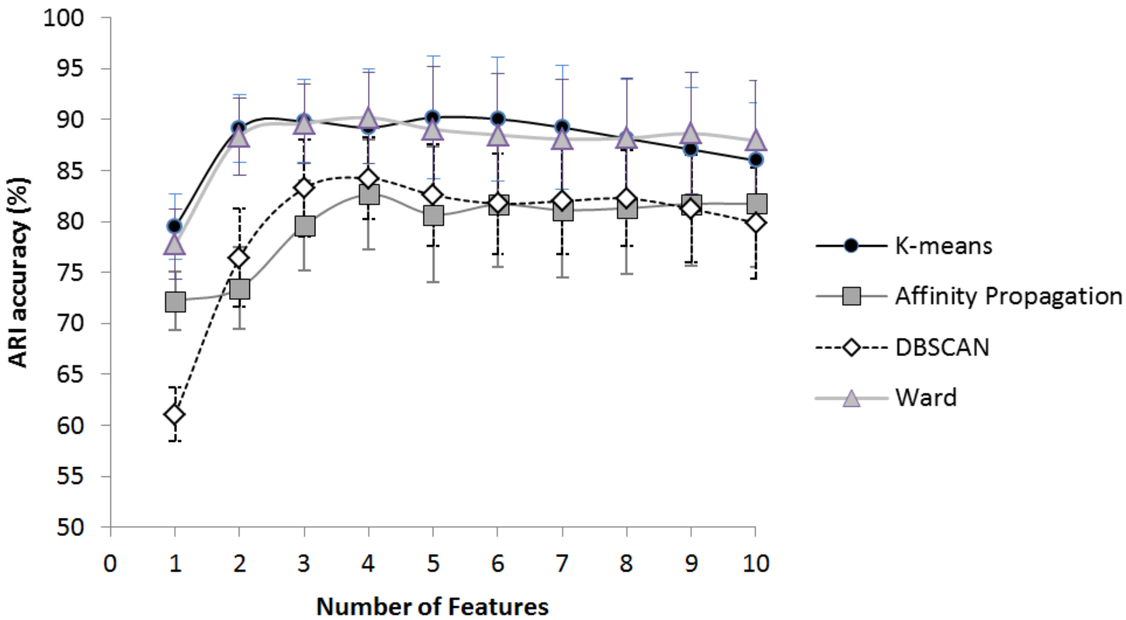


Figure 5.6: ARI accuracy (%) as a function of the number of features used. Each symbol corresponds to a single clustering technique (K-means, Affinity Propagation, DBSCAN and Ward). Each point corresponds to the average and the respective standard deviation of the results from 9 subjects shown in appendix B.6.

The figure 5.6 suggests higher ARI percentages from 4 to 7 features. For a more detailed analysis, Table 5.5 shows the ARI values obtained with different clustering methods and when the framework uses the best 4 to 7 features (sets A, B, C and D respectively) and all features.

Table 5.5: ARI accuracy for all features (first column) and for the best set of features (set A-second, set B-third, set C-fourth and set D-fifth columns). The last row refers to the time interval used to compute the clustering for each set of features. Each ARI result corresponds to the average and standard deviation of the values from appendix B.6.

	ARI (%)				
Clustering Methods	All features	Set A	Set B	Set C	Set D
K-means	87.7 ± 5.6	87.4 ± 12.8	90.0 ± 10.0	84.5 ± 7.5	84.5 ± 9.2
Affinity Propagation	78.4 ± 6.9	76.9 ± 10.3	76.9 ± 10.3	78.3 ± 6.5	81.2 ± 6.0
DBSCAN	78.4 ± 7.0	74.2 ± 9.7	78.1 ± 11.0	80.4 ± 6.3	79.8 ± 10.8
Ward	86.3 ± 8.7	86.0 ± 13.9	88.6 ± 11.7	84.7 ± 9.0	84.7 ± 9.0
Time Response (s)	**347.2**	**107.5**	**132.1**	**142.3**	**145.1**

From table 5.5, it can be concluded that set B is the best set for K-means and Ward performances with $87.4\% \pm 12.8\%$ and $86.0\% \pm 13.9\%$, respectively. On the other hand, the set C showed higher performance for the DBSCAN method with $80.4\% \pm 6.3\%$ while the set D showed best performance for the Affinity Propagation with $81.2\% \pm 6.0\%$. For this reason, any choice from set B, C and D is acceptable. For K-means, Affinity Propagation, DBSCAN and Ward, the clustering performance values are $84.5\% \pm 9.2\%, 81.2\% \pm 6.0\%, 79.8\% \pm 10.8\%$ and $84.7\% \pm 9.0\%$ for set D, $90.0\% \pm 10.0\%$, $76.9\% \pm 10.3\%$, $78.1\% \pm 11.0\%$ and $88.6\% \pm 11.7\%$ for set B. Overall, sets B and D showed similar computing times (with difference of approximately 13 seconds) and the best accuracy values. In the present study set D is the set of the best 7 features thus it was chosen for the proposed framework.

Next, the most suitable set of features that best characterize the overall behavior of ACC signals was tested. After selecting the best number as 7 features, the best group of feature types was found through a histogram, where the occurrences of each feature type was represented. A set of 10 features was computed from each ACC file when the Forward Feature Selection was computed. For all resulting sets, the percentage of occurrence (in %) for each feature type was collected and is presented in figure 5.7. This algorithm was carried out with all implemented features (vide chapter 4, Table 4.1), with segmentation windows of 4000 samples and both acceleration components, BA and GA.

The histogram shown in Figure 5.7 suggests that the Forward Feature Selection algorithm used with Log Scale Power Bandwidth with an higher frequency, Root Mean Square, Total Energy, Autocorrelation, Variance, Wavelet Coefficients and the Mean for HAR systems. Therefore, these feature types are the most used and promising features for the developed framework. These 7 feature types will be used as input for clustering methods in the following studies.

Furthermore, the figure 5.7 suggests that the Log Scale Power Bandwidth occurs more

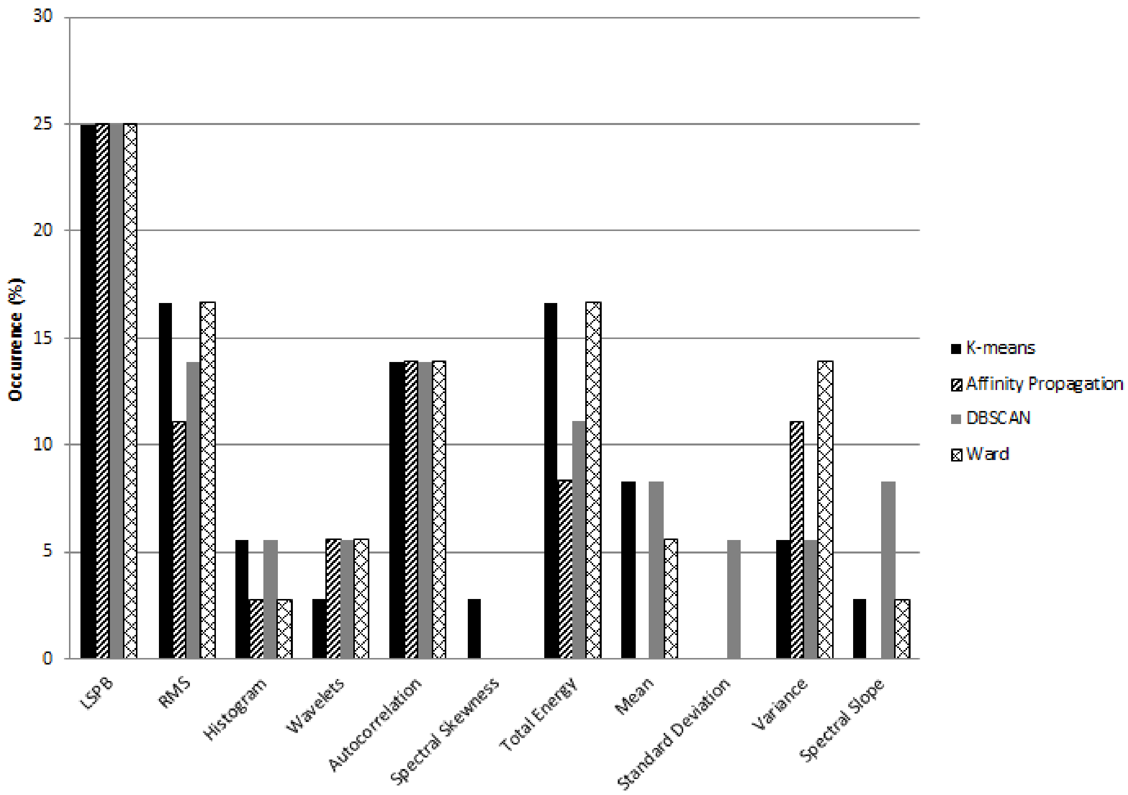


Figure 5.7: Representation of the Forward Feature Selection results. The algorithm outputted the set of the best 10 features for each clustering method. Each column corresponds to all occurrences of each feature type in all resulting sets for each clustering method. The LSPB initials refer to Log Scale Power Bandwidth and RMS to Root Mean Square.

frequently (over 20%) than all the other types of features (with less than 20% in all case occurrences). The Log Scale Power Bandwidth algorithm involves complex stages and offers a wide number of coefficients as output. The resulting data from those 40 output coefficients are complementary. One particular coefficient tended to be more sensible in distinction of two activities due to the variation in behavior over time while the other coefficient seems to identify better other two different tasks.

Moreover, the Log Scale Power Bandwidth feature uses data from the lower frequencies (vide section 4.1.2 from the chapter 4). A detailed analysis in this frequency range suggested that there was important information for activity recognition in accelerometry and the elimination of the filtering stage allowed the information preservation at low frequencies.

It was possible to observe from figure 5.8, and in opposition to the others features, that each Log Scale Power Bandwidth coefficient showed an overall distinction of all subjects activities carried out by the volunteers. Therefore the choice of this type of feature from the Forward Feature Selection as the best feature is justified for its greater ability for activity discrimination.

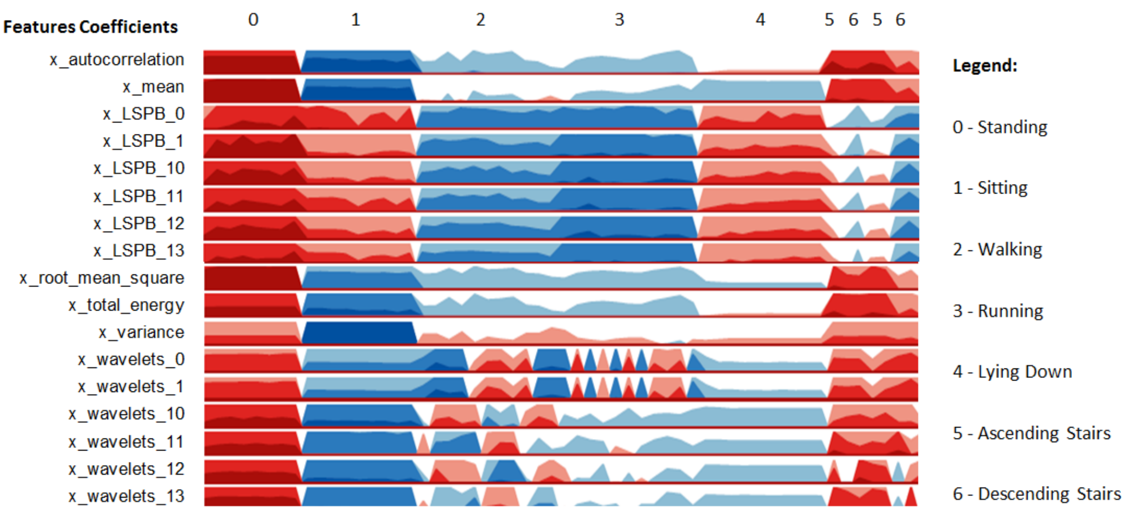


Figure 5.8: Horizon Plot with some features' coefficients from the X-axis acceleration component varying over time according to the type of activity performed. The visualized feature coefficients belong to the FCHA database, acquired in a controlled environment. In the feature coefficients names x refers to the coefficients of the x-axis acceleration component. LSPB refers to Log Scale Power Bandwidth and the final number in each name refers to the feature coefficient's value. The first row of the figure formed by numbers from 0 to 6 refers to different physical activities: Standing (0), sitting (1), walking (2), running (3), lying down (4), ascending stairs (5) and descending stairs (6). The blue and the red colors correspond to positive and negative values, respectively. The color intensity increases with the respective absolute value.

Some difficulties referred in [35] were identified in this work. However some of them were overcome due to the presence of the GA component in the processed data and the use of the Log Scale Power Bandwidth and Wavelet coefficients as features. Some of these difficulties involved the hard discrimination between sitting and standing positions and between walking and running activities. The Horizon Plot in figure 5.8 showed the variation of six Log Scale Power Bandwidth coefficients, six Wavelet coefficients and one coefficient from the Autocorrelation, Mean, Root Mean Square, Total Energy and Variance over time. It is possible to observe that feature types such as Log Scale Power Bandwidth and Wavelets are important for the standing and sitting positions distinction as well as in many other tasks.

5.7 Clustering Evaluation for Activity Recognition

5.7.1 Human Activity Recognition with a controlled approach

After choosing all required parameters for the framework construction, the ACC signals from the FCHA database acquired in a controlled environment and those from the PAMAP database were applied to the developed algorithms set. In this study, unsupervised and supervised learning approaches were applied. Tables 5.6 and 5.7 show all results achieved with the ARI and the Hungarian methodologies application after the clustering and the classification stage, respectively.

The PAMAP signals were acquired with a frequency sampling of 100 *Hz* while the FCHA database is formed by ACC signals acquired with 800 *Hz*. No filters were used in this study. Thus, this data showed different resolutions which influence the amount of information available for the clustering and classification methods.

Table 5.6: Clustering evaluation with the ARI and the Hungarian methodologies (in %) for K-means, Affinity Propagation, DBSCAN and Ward. Each result corresponds to the average and standard deviation of the results shown in the annex B.7.

Clustering (ARI %)				
Databases	K-means	Affinity Propagation	DBSCAN	Ward
FCHA	84.5 ± 9.2	81.2 ± 6.0	79.8 ± 10.8	84.7 ± 9.0
PAMAP	61.6 ± 13.9	63.0 ± 0.2	74.3 ± 16.1	60.5 ± 13.7
Clustering (HA %)				
Databases	K-means	Affinity Propagation	DBSCAN	Ward
FCHA	89.1 ± 7.4	73.2 ± 8.0	76.6 ± 9.7	87.1 ± 8.9
PAMAP	74.5 ± 8.4	83.9 ± 13.7	70.8 ± 10.1	71.1 ± 10.4

Table 5.7: Classification accuracy (in %) with Nearest Neighbors, Random Forest and LDA methods. Each result corresponds to the average and standard deviation of the results shown in the annex B.7.

Classification (Accuracy %)			
Databases	Nearest Neighbors	Random Forest	LDA
FCHA	97.8 ± 6.7	95.4 ± 12.6	98.6 ± 4.3
PAMAP	99.4 ± 1.2	97.9 ± 3.9	98.0 ± 2.4

The results obtained within the ARI method range from $79.8\% \pm 10.8\%$ to $84.7\% \pm 9.0\%$ and from $60.5\% \pm 13.7\%$ to $74.3\% \pm 16.1\%$ for the FCHA and PAMAP databases, respectively. On the other hand, the Hungarian method results range from $73.2\% \pm 8.0\%$ to $89.1\% \pm 7.4\%$ and from $70.8\% \pm 10.1$ to $83.9\% \pm 13.7\%$ for the same databases. The classification showed even higher results: from $95.4\% \pm 12.6\%$ to $99.4\% \pm 1.2\%$ for both databases.

More than 85% of the presented results showed an accuracy higher than 70% which revealed the framework's robustness and versatility for activity recognition with ACC signals, acquired with different sensors and different resolutions.

5.7.2 Human Activity Recognition with a non-controlled approach

The adopted protocol for ACC signals acquisition, explained in the chapter 3, section 3.1, was divided in two parts, with a controlled and non-controlled environments, respectively. This protocol allowed the analysis of the clustering accuracy with these two

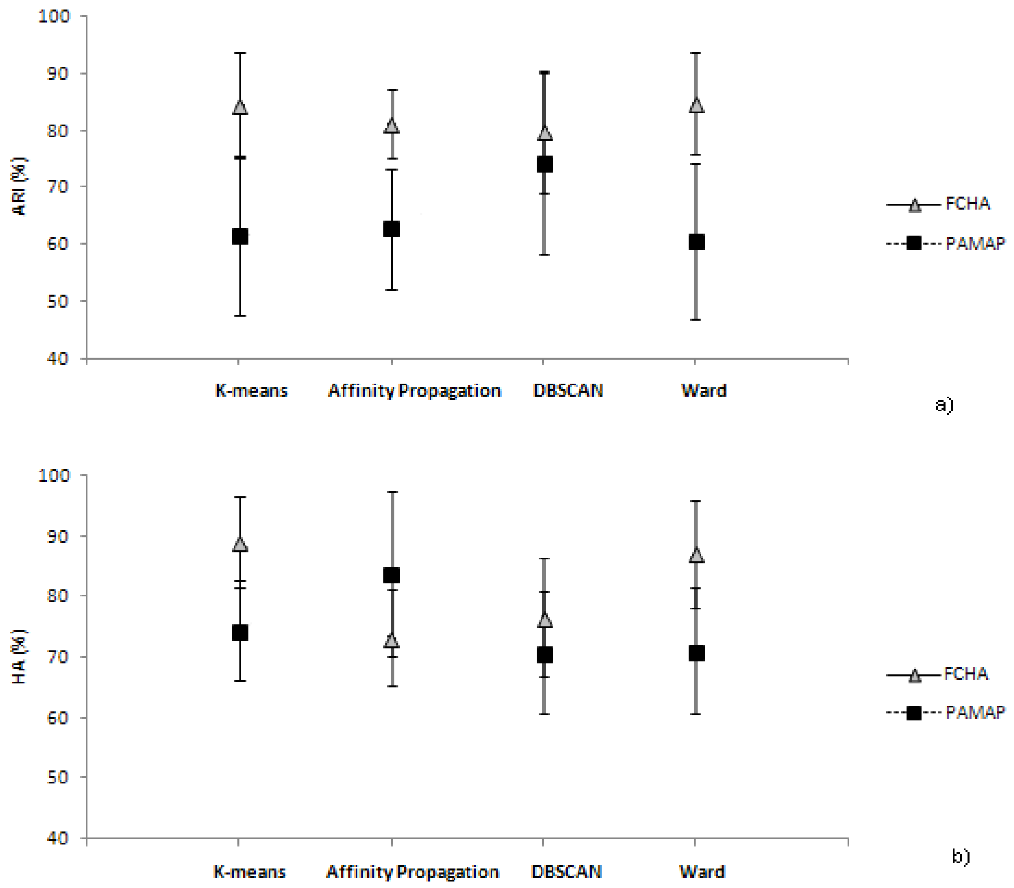


Figure 5.9: Clustering evaluation with a)the ARI and b)the Hungarian methodologies (in %) for K-means, Affinity Propagation, DBSCAN and Ward. Each point corresponds to the average and standard deviation of the results shown in the annex B.7.

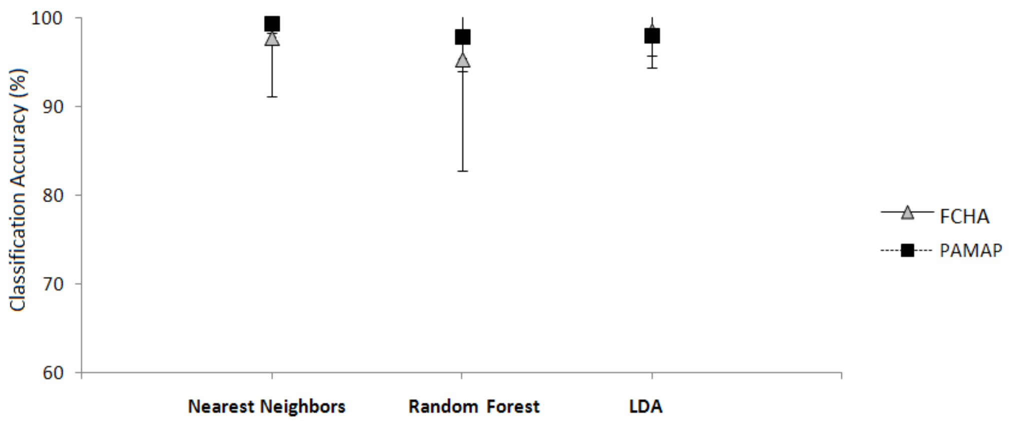


Figure 5.10: Classification methodology (in %) with Nearest Neighbors, Random Forest and LDA methods. Each point corresponds to the average and standard deviation of the results shown in the annex B.7.

environments and to show the versatility of the framework.

This study aimed to show if the developed framework identified satisfactorily activities performed with different approaches and that vary from subject to subject and conducted in an uncontrolled way. This analysis requires a decreased window size decrease due to the short time interval of some of the performed activities. The tasks' duration were not controlled and were performed according to the desire of every subject. However, the results from section 5.1 suggested that the clustering performance increased with the window size. Therefore window duration smaller than 5 seconds and higher than 1 or 2 seconds is advisable in order not to compromise the final performances. Time duration of 3 seconds was chosen for this study and none filtering stage was carried out. All signals available from the FCHA database (acquired with controlled and non-controlled approaches) were used in this study. ARI algorithm was adopted as the clustering performance evaluation methodology.

Table 5.8: ARI accuracy results according to each clustering method and adopted approach (controlled or non-controlled acquisition). The first and the second tables refer to window duration of 3 and 5 seconds, respectively. Each result corresponds to the average and standard deviation of the values presented in appendix B.8.

ARI (%) with windows of 3 seconds

Clustering Methods	Controlled Environment	Non-controlled Environment
K-means	81.9 ± 11.7	71.8 ± 11.7
Affinity Propagation	83.2 ± 11.1	71.2 ± 11.1
DBSCAN	77.6 ± 12.9	70.2 ± 12.9
Ward	83.9 ± 13.6	71.6 ± 13.6

ARI (%) with windows of 5 seconds

Clustering Methods	Controlled Environment	Non-controlled Environment
K-means	84.5 ± 9.2	72.3 ± 11.9
Affinity Propagation	81.2 ± 6.0	64.9 ± 13.6
DBSCAN	79.8 ± 10.8	65.8 ± 12.1
Ward	84.7 ± 9.0	70.9 ± 14.4

The clustering performances from signals acquired in the non-controlled environment range from $70.2\% \pm 12.9$ to $71.8\% \pm 11.7\%$ and from $64.9\% \pm 13.6\%$ to $72.3\% \pm 11.9\%$ for window sizes of 3200 and 4000 samples, respectively. As expected, data acquired in a non-controlled environment showed smaller performances than data acquired in a controlled environment. In the first case variability between subjects' performances was higher because all of them carried out each physical task accordingly to their desire. Different rhythms and velocities difficult the activity recognition and this must be taken into account. The high variability of execution approaches represents a barrier to movement recognition in a non-controlled environment.

However, by adjusting the size of the target window accordingly to the duration of the activities, the accuracy values obtained were higher than 70% which showed that the developed HAR system had a high adaptability to different acquisition environments and to different approaches of the same activity.

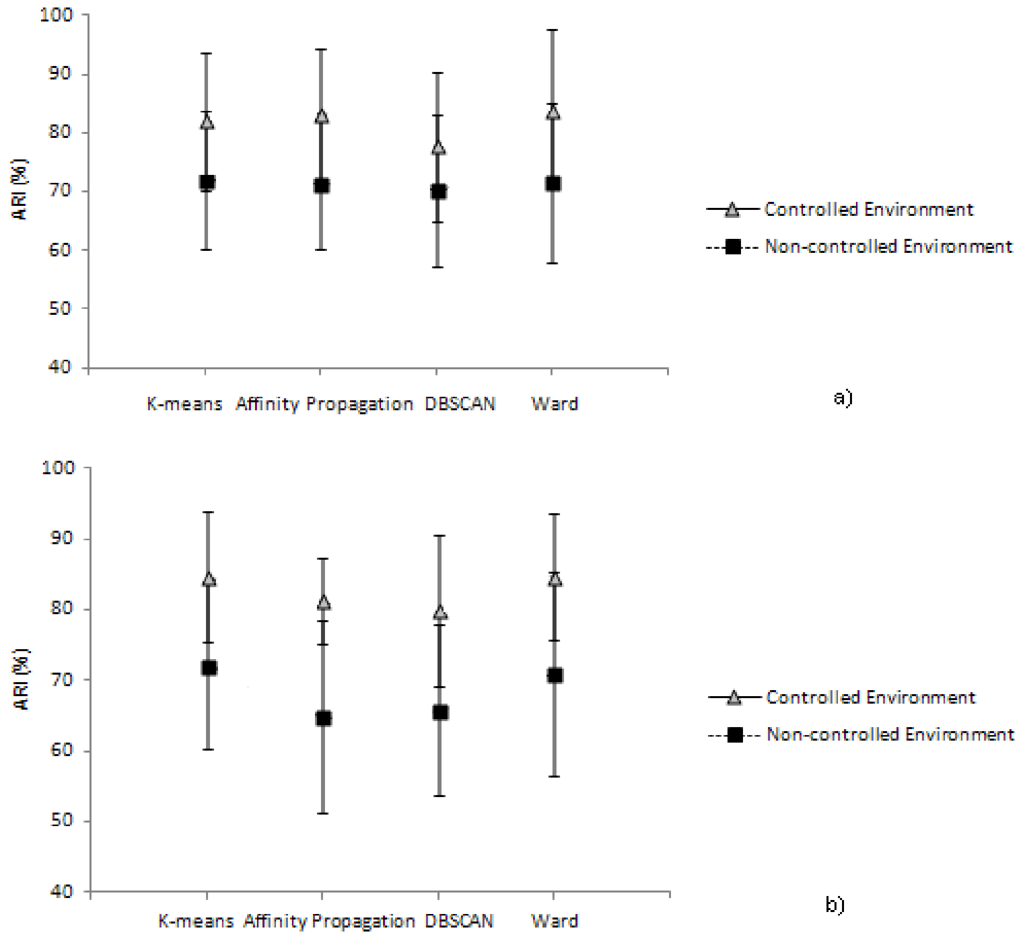


Figure 5.11: ARI results (%) according to each clustering method: K-means, Affinity Propagation, DBSCAN and Ward, and according to the adopted approach: acquisition in a controlled or non-controlled environment for window duration of a) 3 and b) 5 seconds. Each point corresponds to the average and standard deviation of the values presented in appendix B.8.

5.7.3 Clinical State Recognition with Asymptomatic and Symptomatic cases

Clustering techniques were applied for clinical state recognition with asymptomatic and symptomatic cases as an initial approach in the clinical area. This study aimed to recognize the locomotion state of patients that show physical evidences of stroke. The available impairment scores shown in chapter 3, section 3.3 corresponded to the evaluation of several parameters such as locomotion, ease of speech and reasoning and response speed. Therefore, this data cannot be used as ground truth for clustering evaluation.
The clinical database was used in this study and only k-means clustering method was applied: K-means. Windows of 5 seconds were used for the segmentation stage and the best seven features extracted from the clinical data (vide section 5.6). Since there is no ground truth, a k-value of 4 clusters was defined where one cluster corresponds to the normal locomotion and the other three for different abnormal locomotion levels.

Table 5.9: Nine asymptomatic (N0-9) and nine symptomatic (P0-9) volunteers are presented with the respective dysfunction type and impairment score. The clinical data was organized according to the K-means method and its output is shown in the last column. The ND parameter refers to not defined.

Subject	Dysfunction Type	Impairment Score	Clustering Output
N0-9	None	ND	0
P0	CVA	21	0
P1	CVA	22	1
P2	CVA	17	0
P3	CVA	13	1
P4	Diabetic peripheral neuropathy	20	0
P5	CVA + lymphoma	22	2
P6	CVA	15	3
P7	CVA	10	2
P8	CVA	14	3

This study suggests that different subjects with different impairment scores and asymptomatic subjects may belong to the same cluster. These results are only based on the difficulty of locomotion.

Even without the availability of ground truth, it is possible to take some conclusions from the obtained results. Different subjects (P1 with P3 and P5 with P7) showed different impairment scores and were grouped in the same cluster as suggested in Table 5.9. Individuals with different impairment scores may show a similar difficulty in locomotion and different physical evaluations in other parameters.

Furthermore, subjects P0, P2, who suffered strokes, and P4, with diabetic peripheral neuropathy, were clustered with the asymptomatic individuals. These results suggested that are two symptomatic and asymptomatic domains with an overlap which include patients with health complications but with no apparent difficulty in mobility. P6 and P8 were clustered together and present almost the same impairment score (15 and 14, respectively).

This clustering stage is an objective evaluation of the locomotion impairment and its application may represent an important help in the health and therapy fields.

6

Conclusion

All achievements obtained in this thesis are summarised in this chapter. The obtained results and its implications are presented to show the main contributions of this work for scientific knowledge.

6.1 Contributions and Take Home Message

This work aimed to create and develop a novel recognition system based on literature concepts and on the HAR system framework created in [35].
All signals acquired from the MotionPlux sensor in a controlled environment were analyzed to find the best conditions and framework's parameters to cluster daily tasks. This data together with the ACC signals acquired from the same MotionPlux sensor in a non-controlled environment compose the FCHA database.
The main contribution of this work involved the study and choice of the parameters for the framework's construction.
Initially, the window size was studied and the greater window size, with 5 second time interval, was chosen based on the highest ARI value. With an increase of window size more information is available in each segment for clustering. It is important to remind that this study did not take into account the transitions between activities. Similarly to [35] each window contains only information about one physical activity.
Several clustering methods were tested and four of them were chosen according to its performance values: K-means, Affinity Propagation, DBSCAN and Ward algorithms.
In addition to those added in [35] new features were implemented, such as the CUIDADO features, from the audio recognition applications, wavelets and Log Scale Power Bandwidth coefficients, inspired in the Mel Scale. This thesis offers a set of 423 features for

machine learning techniques providing more and new information regarding the performed movement tasks compared to the literature [19], [27], [35], [44], [48], [58].

The Forward Feature Selection aimed to reduce the undesirable redundancy between features and to select the set of features that may lead to the best framework's performance. The chosen features as the most suitable features for HAR systems are the Log Scale Power Bandwidth, Root Mean Square, Total Energy, Autocorrelation, Variance, Wavelet Coefficients and Mean coefficients.

The achieved results also suggested that it is important not to waste any information regarding to the movement. The presence of the information located in the low frequencies, such as the GA component, and the Log Scale Power Bandwidth implementation lead to better static activities distinction.

As expected, the HMM application improves the ARI values according to the clustering method used. Therefore, it is possible to create a database in the future, similar to those used in audio recognition, where all transitions between the existing states (activities) with higher or lower probability are known. The addition of more volunteers and ACC signals to the database may lead to higher performances in the identification of different activities.

In addition to the exposed, the Hungarian methodology was computed as a second performance algorithm evaluation. A second database was analysed (the PAMAP database) where signals showed a different sampling frequency, 100 *Hz*, [45] from those acquired with MotionPlux sensors (FCHA database), of 800 *Hz*. The obtained results show that the developed framework is suitable for activities recognition even for ACC data with very different resolutions. The FCHA signals acquired in a non-controlled environment were also analyzed and their high results showed the framework's versatility and robustness.

All the clustering performance evaluation methods and the HMM algorithm required the availability of a ground truth which is logged along with the annotation. Without the ground truth, both studies could not be applied.

Unlike all studies mentioned, others were performed although without success. The processing of ACC signals through wavelets decomposition and reconstruction before of the clustering stage application was applied. This step involved elimination of noise and signal compression which would be important for higher ARI accuracies. This improvement was not achieved. In this stage of wavelets decomposition and reconstruction important information for activity recognition may be lost which leads to lower ARI values. Moreover, the K-value (number of clusters) prediction also showed no promising results for its application in the final framework.

The activity recognition with Markov Models, similarly to the speech recognition, was also a study with no reliable and acceptable results due to the reduced database's size, providing insufficient information for the Markov Models learning. Moreover, audio data does not show the same properties and behavior as ACC data. Activities carried out by volunteers cannot be studied as Markov states in a similar way as audio data analysis. This supervised learning study requires a deep adaptation for the ACC data application.

The major achievements of the current work referred before allowed the construct of a HAR system ensuring promising clustering performances without heavy computational demands. Therefore, as expected accelerometry is a suitable technique for monitoring movement patterns in free-living subjects over long periods of time. The knowledge acquired over this thesis may be applied into the clinical setting for the diagnosis and physiotherapy fields.

6.2 Future Work

This thesis presents topics that require further investigation and can be exploited in the future, such as:

Increasing data for analysis: the enlargement of the number of volunteers or even patients is an important task for the reliability of the results achieved with this HAR system framework. The increase of the number and duration of the performed activities also contributes for the versatility of the algorithm and leads to a recognition system more adaptable to the complexity of daily activities;

More studies for the continuous development of the HAR systems framework: in a non-controlled data acquisition environment, the number of activities is not known. Therefore, a detailed study of the K-value prediction may be an interesting topic especially for the Ward and K-means application. The transition's tasks between activities must be also taken into account in a signal analysis for a more naturalistic and realistic study;

Adapting the acquired knowledge to the most used technologies: acquiring ACC data from accelerometers present in many technologic devices, such as cellphones, may be an interesting approach by making easier access of more people to the required equipment and their participation in the study. However, this topic may provide additional problems due to the sensor's location (usually far from the CoM) and its variable location over time relatively to the body;

Applying the developed framework into different contexts: testing the ACC algorithm in different environments and situations, mainly for clinical applications. The unceasing need to obtain information in a fast and non-invasive way will lead to data processing algorithms development in the clinical area. Therefore, an activity recognition system based in accelerometry must show high performances and adaptability under several acquisition conditions. This ACC application may be involved in the diagnosis and treatment procedures of some movement disorders, such as CVA consequences, OCD and Parkinson's disease.

Bibliography

[1] *Accelerometer's Position Free Human Activity Recognition Using A Hierarchical Recognition Model*, 12th IEEE International Conference on e-Health Networking Applications and Services (Healthcom), Lyon, 2010, pp. 296 –301. DOI: 10.1109/HEALTH.2010.5556553.

[2] *A Comparison of Unsupervised Learning Algorithms for Gesture Clustering*, 6th international conference on Human-robot interaction, University of Sydney, Australia, 2011, pp. 111–112. DOI: 10.1145/1957656.1957686.

[3] H. J. Aizenstein, J. T. Becker, and O. L. Lopez, "Measuring physical activity using accelerometry in a community sample with dementia", vol. 61, no. 1, 158–159, 2013. DOI: 10.1111/jgs.12050.

[4] A. Alguwaizani, "Degeneracy on k-means clustering", vol. 39, 13–20, 2012. DOI: 10.1016/j.endm.2012.10.003.

[5] S. BM, D. D, W. M, S. Y, B. S, and M. RW, "Physical activity and multiple sclerosis: new insights regarding inactivity", vol. 126, no. 4, 256–262, 2012. DOI: 10.1111/j.1600-0404.2011.01634.x.

[6] S. Boukir, S. Jones, and K. Reinke, "Fast mean-shift based classification of very high resolution images: application to forest cover mapping", vol. 1, no. 7, pp. 111–116, 2012, ISSN: 2194-9042.

[7] M. M.-T. Chiang, "Intelligent choice of the number of clusters in k-means clustering: an experimental study with different cluster spreads", vol. 27, no. 1, pp. 3–40, 2009. DOI: 10.1007/s00357-010-9049-5.

[8] S. J. Chunhao Tua and W. Kohc, "Comparison of clustering algorithms on generalized propensity score in observational studies: a simulation study", *Journal of Statistical Computation and Simulation,*, vol. 83, no. 12, pp. 2206–2218, 2013. DOI: 10.1080/00949655.2012.685169.

[9] (Jan. 2014). Clustering, [Online]. Available: http://scikit-learn.org/stable/modules/clustering.html#clustering.

[10] R. K. Curey, "Gyro and accelerometer panel: 50 years of service to the inertial community", vol. 28, no. 7, pp. 23–29, 2013. DOI: 10.1109/maes.2013.6559378.

[11] P. Esser, H. Dawes, J. Collett, and K. Howells, "Insights into gait disorders: walking variability using phase plot analysis", vol. 38, no. 4, 648–652, 2013. DOI: 10.1016/j.gaitpost.2013.02.016.

[12] I. Eyal, I. Keidar, and R. Rom, "Distributed data clustering in sensor networks", vol. 24, no. 5, 207–222, 2011. DOI: 10.1007/s00446-011-0143-7.

[13] P. Fazio, G. Granieri, I. Casetta, E. Cesnik, S. Mazzacane, P. Caliandro, F. Pedrielli, and E. Granieri, "Gait measures with a triaxial accelerometer among patientswith neurological impairment", vol. 34, no. 4, pp. 435–440, 2013. DOI: 10.1007/s10072-012-1017-x.

[14] F. Foerster and J. Fahrenberg, "Motion pattern and posture: correctly assessed by calibrated accelerometers", vol. 32, no. 3, pp. 450–457, 2000. DOI: 10.3758/BF03200815.

[15] E. Fosler-Lussier, "Markov models and hidden markov models - a brief tutorial", 1998.

[16] J. Gałka and M. Ziółko, "Wavelets in speech segmentation", 2008. DOI: 10.1109/MELCON.2008.4618547.

[17] Z. Ghahramani, *Unsupervised Learning*. Carnegie Mellon University, USA: Springer-Verlag, 2004, pp. 3–4.

[18] W. Günthner, *Enhancing Cognitive Assistance Systems with Inertial Measurement Units*. USA: Studies in Computational Intelligence, Springer-Verlag Berlin Heidelberg, 2008.

[19] Godfrey, R. Conway, D. Meagher, and G. ÓLaighin, "Direct measurement of human movement by accelerometry", *Elsevier*, vol. 30, no. 10, pp. 1364–1386, 2008. DOI: 10.1016/j.medengphy.2008.09.005.

[20] V. Gonzalez-Castro, R. Alaiz-Rodriguez, and E. Alegre, "Class distribution estimation based on the hellinger distance", vol. 218, no. 1, 146–164, 2013. DOI: 10.1016/j.ins.2012.05.028.

[21] P. Guo, Z. Miao, X.-P. Zhang, Y. Shen, and S. Wang, "Coupled observation decomposed hidden markov model for multiperson activity recognition", vol. 22, no. 9, pp. 1306 –1320, 2012. DOI: 10.1109/TCSVT.2012.2199390.

[22] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*. San Francisco, USA: Morgan Kaufmann, Second Edition, 2006, 383–458, ISBN: 10: 1-55860-901-6.

[23] J. He, Z. Liu, Lianwen, J. L.-X. Zhen, and J.-C. Huang, "Weightlessness feature — a novel feature for single tri-axial accelerometer based activity recognition, 19th international conference on pattern recognition", pp. 1–4, 2008. DOI: 10.1109/ICPR.2008.4761688.

[24] I. H.Witten and E. Frank, *Data Mining – Practical Machine Learning Tools and Techniques*. San Francisco, CA, USA: Second Edition, Elsevier, Morgan Kaufmann Publishers, 2005, pp. 1–37, 81, 254–270.

[25] A. K. Jain, "Data clustering: 50 years beyond k-means", vol. 5211, pp. 3–4, 2008. DOI: 10.1007/978-3-540-87479-9_3.

[26] Z. Jiang, Z. Lin, and L. S. Davis, "Class consistent k-means: application to face and action recognition", vol. 116, no. 6, 730–741, 2012. DOI: 10.1016/j.cviu.2012.02.004.

[27] J. J. Kavanagh and H. B. Menz, "Accelerometry: a technique for quantifying movement patterns during walking", vol. 28, no. 1, pp. 1–15, 2008. DOI: 10.1016/j.gaitpost.2007.10.010.

[28] A. M. Khan, Y.-K. Lee, S. Y. Lee, and T.-S. Kim, "A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer", vol. 14, no. 5, pp. 1166–1172, 2010. DOI: 10.1109/TITB.2010.2051955.

[29] H. W. Kuhn, "The hungarian method for the assignment problem", vol. 2, pp. 83–97, 1955. DOI: 10.1007/978-3-540-68279-0_2.

[30] D. Labate, G. Weiss, and E. Wilson, "Wavelets", vol. 60, no. 1, pp. 66–76, 2013. DOI: 10.1090/noti927.

[31] T. Liao, "Clustering of time series data—a survey", vol. 38, no. 11, 1857 – 1874, 2005. DOI: 10.1016/j.patcog.2005.01.025.

[32] R.-S. Lin and L.-H. Chen, "A new approach for audio classification and segmentation using gabor wavelets and fisher linear discriminator", vol. 19, no. 6, pp. 807–822, 2005. DOI: 10.1142/S0218001405004289.

[33] R. Loohach and K. Garg, "Effect of distance functions on k-means clustering algorithm", vol. 49, no. 6, 2012.

[34] M. M., H. FB, Z. C, C.-K. P., N. JG, and C. L., "Trunk accelerometry reveals postural instability in untreated parkinson's disease", vol. 17, no. 7, 557–562, 2011. DOI: 10.1016/j.parkreldis.2011.05.010.

[35] I. P. Machado, "Human activity data discovery based on accelerometry", Master's thesis, Physics Department, Faculty of Sciences, Technology, New University of Lisbon, and Plux, Wireless Biosignals, Portugal, 2013.

[36] *Activity Recognition and Monitoring Using Multiple Sensors on Different Body Positions*, Cambridge, MA, USA: Proceedings of the International Workshop on Wearable and Implantable Body Sensor Networks 0-7695-2547-4/06 - IEEE, 2006, pp. 113–116. DOI: 10.1109/BSN.2006.6.

[37] D. Minnen, T. Starner, J. A. Ward, P. Lukowicz, and G. Troster, "Recognizing and discovering human actions from on-body sensor data", pp. 1545 –1548, 2005. DOI: 10.1109/ICME.2005.1521728.

[38] B. Mirkin, “Choosing the number of clusters”, vol. 1, no. 3, pp. 252–260, 2011. DOI: 10.1002/widm.15.

[39] *Recognizing Human motion with multiple acceleration sensors*, vol. 2, Tucson, AZ, USA: IEEE International Conference on Systems, Man, and Cybernetics, 2001, pp. 747 –752. DOI: 10.1109/ICSMC.2001.973004.

[40] R. Moe-Nilssen, “A new method for evaluating motor control in gait under real-life environmental conditions part 1: the instrument”, vol. 13, no. 4, pp. 320–332, 1998. DOI: 10.1016/S0268-0033(98)00089-8.

[41] *Morphological analysis of acceleration signals in cross-country skiing*, Rome, Italy: BIOSIGNALS 2011 - Proceedings of the International Conference on Bio-inspired Systems and Signal Processing, 26-29 January, 2011.

[42] *Research on k-means Clustering Algorithm - An Improved k-means Clustering Algorithm*, Jinggangshan: Third International Symposium on Intelligent Information Technology and Security Informatics (IITSI), 2010, pp. 63 –67. DOI: 10.1109/IITSI.2010.74.

[43] Y. Nam, S. Rho, and C. Lee, “Physical activity recognition using multiple sensors embedded in a wearable device”, vol. 12, no. 2, 2013. DOI: 10.1145/2423636.2423644.

[44] N. F. M. Nunes, “Algorithms for time series clustering”, Master’s thesis, Physics Department, Faculty of Sciences, Technology, New University of Lisbon, and Plux, Wireless Biosignals, Portugal, 2011.

[45] (Jan. 2014). Pamap, [Online]. Available: http://www.pamap.org/demo.html.

[46] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, “Scikit-learn: machine learning in python”, vol. 12, pp. 2825–2830, 2011.

[47] G. Peeters, “A large set of audio features for sound description (similarity and classification) in the cuidado project”, 2004.

[48] G. Plasqui, A. G. Bonomi, and K. R. Westerterp, “Daily physical activity assessment with accelerometers: new insights and validation studies”, vol. 14, no. 6, pp. 451–462, 2013. DOI: 10.1111/obr.12021.

[49] P. F. Qinpei Zhao Mantao Xu, “Knee point detection on bayesian information criterion”, 2008. DOI: 10.1109/ICTAI.2008.154.

[50] *Activity Recognition from Accelerometer Data*, vol. 3, USA: 17th conference on Innovative applications of artificial intelligence - AAAI Press, 2005, pp. 1541–1546. DOI: 10.1109/MPRV.2002.1037719.

[51] J. S. Ricardo Gomes Neuza Nunes and H. Gamboa, “Long term biosignals visualization and processing”, 2012.

[52] G. W. Robert Tibshirani and T. Hastie, “Estimating the number of clusters in a data set via the gap statistic”, vol. 63, pp. 411–423, 2001.

[53] M. RW, P. L, S. BM, D. D, S. JJ, and P. JH, “Accelerometry as a measure of walking behavior in multiple sclerosis”, vol. 127, no. 6, 384–390, 2013. DOI: 10.1111/ane.12036.

[54] K. Sato, W. A. Sands, and M. H. Stone, “The reliability of accelerometry to measure weightlifting performance”, vol. 11, no. 4, pp. 524–531, 2012. DOI: 10.1080/14763141.2012.724703.

[55] D. Trabelsi, S. Mohammed, F. Chamroukhi, L. Oukhellou, and Y. Amirat, “An unsupervised approach for automatic activity recognition based on hidden markov model regression”, vol. 10, no. 3, pp. 829 –835, 2013. DOI: 10.1109/TASE.2013.2256349.

[56] S. F. Walter, “Clustering by affinity propagation”, Master’s thesis, Department of Physics, ETH Zurich, Switzerland, 2007.

[57] J. Wu, G. Pan, D. Zhang, G. Qi, and S. Li, “Gesture recognition with a 3-d accelerometer”, pp. 25–38, 2008. DOI: 10.1109/APCCAS.2008.4745999.

[58] M. Yanga, H. Zhenga, H. Wanga, S. McClean, and D. Newell, “Igait: an interactive accelerometer based gait analysis system”, *Elsevier*, vol. 108, no. 2, pp. 715–723, 2012. DOI: 10.1016/j.cmpb.2012.04.004.

A

Affinity Propagation Method

The following algorithm refers to the Affinity Propagation method applied in this work and developed and shown in [46]. The present section shows all parameters and code in python language of this clustering method.

A.1 Parameters

$damping$: float, optional, default: 0.5. Damping factor between 0.5 and 1.
$convergence\_iter$: int, optional, default: 15. Number of iterations with no change in the number of estimated clusters that stops the convergence.
max_iter : int, optional, default: 200. Maximum number of iterations.
$copy$: boolean, optional, default: True. Make a copy of input data.
$preference$: array-like, shape ($n\_samples$) or float, optional. Preferences for each point; Points with larger values of preferences are more likely to be chosen as exemplars. The number of exemplars, ie of clusters, is influenced by the input preferences value. If the preferences are not passed as arguments, they will be set to the median of the input similarities.
$affinity$: string, optional, $default =' euclidean'$. Which affinity to use. At the moment precomputed and euclidean are supported. euclideanuses the negative squared euclidean distance between points.
$verbose$: boolean, optional, default: False. Whether to be verbose.
$'cluster\_centers\_indices\_'$: array, shape (n_clusters). Indices of cluster centers.
$'cluster\_centers\_'$: array, shape (n_clusters, n_features). Cluster centers (if affinity != precomputed).
$'labels\_'$: array, shape (n_samples). Labels of each point.

$'affinity\_matrix\_'$: array, shape (n_samples, n_samples). Stores the affinity matrix used in fit.

A.2 Algorithm

```

class AffinityPropagation(BaseEstimator, ClusterMixin):

def __init__(self, damping=.5, max_iter=200, convergence_iter=15,
  copy=True, preference=None, affinity='euclidean', verbose=False):

  self.damping = damping
  self.max_iter = max_iter
  self.convergence_iter = convergence_iter
  self.copy = copy
  self.verbose = verbose
  self.preference = preference
  self.affinity = affinity

@property
def _pairwise(self):
  return self.affinity is "precomputed"

def fit(self, X):
  if X.shape[0] == X.shape[1] and not self._pairwise:
    warnings.warn("The API of AffinityPropagation has changed.")
  if self.affinity is "precomputed":
    self.affinity_matrix_ = X
  elif self.affinity is "euclidean":
    self.affinity_matrix_ = -euclidean_distances(X, squared=True)
  else:
    raise ValueError("Affinity must be "precomputed" or "euclidean
        ". Got \%s instead", \%str(self.affinity))

  self.cluster_centers_indices_, self.labels_ = affinity_propagation(
    self.affinity_matrix_, self.preference, max_iter=self.max_iter,
    convergence_iter=self.convergence_iter, damping=self.damping,
     copy=self.copy, verbose=self.verbose)

  return self
```

B

Evaluation of the Performance - Additional Results

B.1 Window Length Influence

Table B.1: ARI accuracy (%) with K-means method according to different window sizes (in samples) and each ACC file.

Subjects \Window Length (samples)	800	1600	2400	3200	4000
0	70.945	73.550	84.600	72.522	90.292
1	96.888	100.00	100.00	100.00	96.064
2	78.395	76.392	85.118	85.137	80.860
3	63.768	64.924	64.869	83.027	87.530
4	84.055	86.247	82.850	87.384	89.505
5	93.144	82.159	87.834	89.669	84.836
6	86.854	77.481	84.276	92.953	94.866
7	68.190	73.169	71.177	72.750	78.127
8	87.516	100	100	83.803	87.167

B.2 Selection of Clustering Methods

Table B.2: ARI accuracy (%) according to different clustering methods and each ACC file.

Subjects \Clustering Methods	Kmeans	Mini Batch K-means	Affinity Propagation	Mean Shift	DBSCAN	Spectral Clustering	Ward
0	90.292	83.140	74.433	37.164	68.635	86.380	79.900
1	96.064	100.000	90.460	25.055	72.528	100.000	96.064
2	80.859	88.556	73.902	22.052	45.819	95.531	70.259
3	87.530	87.530	56.690	19.991	55.819	78.773	87.530
4	89.505	72.232	77.064	31.710	77.244	72.017	89.505
5	84.836	84.556	78.844	48.915	83.937	88.107	90.503
6	94.866	82.812	86.368	48.779	82.535	92.745	99.067
7	78.127	71.085	76.765	20.793	39.461	88.143	76.765
8	87.167	90.947	82.120	26.279	71.510	100.000	87.167

B.3 Wavelet Level Decomposition Influence

Table B.3: ARI accuracy (%) with the K-means method according to different wavelets levels decomposition (from 1 to 15) and to each ACC file.

Subjects \Decomposition Level	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	53.949	59.345	55.961	67.926	61.633	65.858	65.858	64.338	67.409	63.652	63.652	53.089	53.089	53.089	53.089
1	47.230	47.999	52.078	48.921	67.067	68.782	71.668	67.636	59.039	59.039	43.886	32.699	43.886	40.136	43.886
2	40.707	36.808	37.783	30.448	30.535	58.000	60.088	80.083	71.853	66.387	66.387	66.387	66.387	66.387	66.387
3	50.952	37.187	22.595	54.814	67.005	73.253	72.868	79.513	75.343	57.593	57.593	56.217	55.457	55.457	55.457
4	27.774	27.774	28.016	60.863	64.664	73.803	72.780	71.560	58.587	56.547	54.491	57.938	56.310	59.615	67.491
5	59.261	50.697	53.336	59.103	67.733	75.949	75.773	78.458	83.531	83.250	67.070	67.070	67.176	67.070	67.070
6	53.192	53.429	53.192	52.514	52.514	65.772	90.737	86.368	80.742	74.565	54.427	58.263	58.263	58.263	58.263
7	48.003	41.174	55.033	57.688	60.108	64.213	62.818	89.139	89.139	71.249	70.642	68.552	70.452	74.478	66.980
8	45.851	59.686	55.206	48.785	78.617	70.742	71.602	74.146	59.486	59.486	60.820	59.515	55.975	59.515	55.975

B.4 Filtering Influence

Table B.4: ARI accuracy (%) with K-means, Affinity Propagation, DBSCAN and Ward for signal processing with and without the filtering stage, BA and BA+GA components, respectively.

	BA Component				BA + GA Components			
Subjects \Clustering Methods	Kmeans	Affinity Propagation	DBSCAN	Ward	Kmeans	Affinity Propagation	DBSCAN	Ward
0	69.959	68.1646	44.371	84.680	90.292	74.433	68.635	79.900
1	69.744	68.625	56.127	50.16153	96.064	90.460	72.528	96.064
2	61.8151	53.627	72.941	69.286	80.859	73.902	45.819	70.259
3	64.0872	51.686	70.063	67.281	87.530	56.690	55.819	87.530
4	80.111	76.106	45.930	71.141	89.505	77.064	77.244	89.505
5	69.106	63.284	47.191	63.858	84.836	78.844	83.937	90.503
6	69.963	46.430	48.599	80.816	94.866	86.368	82.535	99.067
7	42.4372	47.127	48.366	72.686	78.127	76.765	39.461	76.765
8	57.148	47.453	55.629	52.651	87.167	82.120	71.510	87.167

B.5 Hidden Markov Models Application

Table B.5: ARI accuracy (%) for K-means, Affinity Propagation, DBSCAN and Ward with and without the HMM application.

	ARI (%) Without Markov Models				ARI (%) With Markov Models			
Subjects Clustering Methods	K-means	Affinity Propagation	DBSCAN	Ward	Kmeans	Affinity Propagation	DBSCAN	Ward
0	68.432	74.855	74.433	79.665	90.292	74.433	68.635	79.900
1	100.000	91.780	89.681	94.012	96.064	90.460	72.528	96.064
2	83.087	73.785	50.779	71.791	80.859	73.902	45.819	70.259
3	64.687	87.530	73.580	82.397	87.530	56.690	55.819	87.530
4	86.278	66.276	77.064	85.451	89.505	77.064	77.244	89.505
5	85.411	59.758	78.844	90.222	84.836	78.844	83.937	90.503
6	80.967	86.368	89.533	97.163	94.866	86.368	82.535	99.067
7	77.640	78.127	78.127	73.973	78.127	76.765	39.461	76.765
8	85.830	83.722	53.014	88.711	87.167	82.120	71.510	87.167

B.6 Features Selection

Table B.6: ARI accuracy (%) with K-means, Affinity Propagation, DBSCAN and Ward using all implemented features.

Subjects Clustering Methods	K-means	Affinity Propagation	DBSCAN	Ward
0	90.292	74.433	68.635	79.900
1	96.064	90.460	72.528	96.064
2	80.859	73.902	45.819	70.259
3	87.530	56.690	55.819	87.530
4	89.505	77.064	77.244	89.505
5	84.836	78.844	83.937	90.503
6	94.866	86.368	82.535	99.067
7	78.127	76.765	39.461	76.765
8	87.167	82.120	71.510	87.167

Table B.7: ARI accuracy (%) with K-means, Affinity Propagation, DBSCAN and Ward using only the best 4 feature types.

Subjects Clustering Methods	K-means	Affinity Propagation	DBSCAN	Ward
0	80.594	74.433	74.433	74.230
1	100.000	90.460	83.867	100.000
2	79.956	73.902	70.259	69.088
3	64.687	51.686	56.948	64.687
4	88.368	77.064	77.064	88.368
5	100.000	78.844	59.758	100.000
6	100.000	86.368	86.368	100.000
7	72.695	76.765	76.765	77.222
8	100.000	82.120	82.120	100.000

Table B.8: ARI accuracy (%) with K-means, Affinity Propagation, DBSCAN and Ward using only the best 5 feature types.

Subjects Clustering Methods	K-means	Affinity Propagation	DBSCAN	Ward
0	80.594	74.433	74.433	74.230
1	100.000	90.460	100.000	100.000
2	79.956	73.902	70.259	69.088
3	88.092	51.686	57.206	88.092
4	88.368	77.064	77.064	88.368
5	100.000	78.844	78.844	100.000
6	100.000	86.368	86.368	100.000
7	72.695	76.765	76.765	77.222
8	100.000	82.120	82.120	100.000

Table B.9: ARI accuracy (%) with K-means, Affinity Propagation, DBSCAN and Ward using only the best 6 feature types.

Subjects Clustering Methods	K-means	Affinity Propagation	DBSCAN	Ward
0	70.173	74.855	74.433	71.493
1	100.000	87.775	90.460	100.000
2	81.321	73.785	70.259	84.595
3	87.530	65.767	87.530	87.530
4	86.278	77.064	77.064	88.221
5	84.056	78.844	78.844	82.293
6	84.131	86.368	86.368	86.368
7	78.834	76.765	76.765	69.539
8	88.328	83.722	82.120	92.506

Table B.10: ARI accuracy (%) with K-means, Affinity Propagation, DBSCAN and Ward using only the best 7 feature types.

Subjects Clustering Methods	K-means	Affinity Propagation	DBSCAN	Ward
0	70.173	74.855	74.433	71.493
1	100.000	91.780	100.000	100.000
2	79.956	73.785	75.700	84.595
3	87.530	87.530	91.883	87.530
4	86.278	77.064	83.097	88.221
5	86.615	78.844	78.844	82.293
6	84.131	86.368	86.368	86.368
7	78.834	76.765	76.765	69.539
8	86.416	83.722	87.167	92.506

B.7 Human Activity Recognition with controlled approach

Table B.11: ARI accuracy (%) with K-means, Affinity Propagation, DBSCAN and Ward for the FCHA database application.

Subjects Clustering Methods	Kmeans	Affinity Propagation	Dscan	Ward
0	70.173	74.855	74.433	71.493
1	100.000	91.780	100.000	100.000
2	79.956	73.785	75.700	84.595
3	87.530	87.530	91.883	87.530
4	86.278	77.064	83.097	88.221
5	86.615	78.844	78.844	82.293
6	84.131	86.368	86.368	86.368
7	78.834	76.765	76.765	69.539
8	86.416	83.722	87.167	92.506

Table B.12: HA accuracy (%) with K-means, Affinity Propagation, DBSCAN and Ward for the FCHA database application.

Subjects Clustering Methods	Kmeans	Affinity Propagation	Dscan	Ward
0	78.592	78.869	71.429	78.869
1	100.000	85.714	100.000	100.000
2	84.734	71.358	84.790	85.643
3	82.857	82.857	91.949	82.857
4	93.591	71.429	83.315	85.656
5	97.339	71.429	71.429	98.641
6	89.833	71.429	71.429	71.429
7	78.219	68.730	68.730	85.367
8	95.716	71.429	84.206	95.476

Table B.13: ARI accuracy (%) with K-means, Affinity Propagation, DBSCAN and Ward for the PAMAP database application.

Subjects Clustering Methods	Kmeans	Affinity Propagation	Dscan	Ward
0	43	89	58	44
1	63	61	59	56
2	58	36	96	60
3	82	67	84	82

Table B.14: HA accuracy (%) with K-means, Affinity Propagation, DBSCAN and Ward for the PAMAP database application.

Subjects Clustering Methods	Kmeans	Affinity Propagation	Dscan	Ward
0	60	100	63	60
1	78	77	59	61
2	79	65	80	80
3	81	94	81	83

Table B.15: Classification with Cross-validation(%) with Nearest Neighbors, Random Forest and LDA for FCHA database application.

Subjects Classification Methods	Nearest Neighbors	Random Forest	LDA
0	83.871	87.097	100.000
1	96.429	96.429	100.000
2	96.667	96.667	100.000
3	96.296	96.296	100.000
4	90.909	93.939	93.939
5	100.000	100.000	100.000
6	96.429	96.429	100.000
7	96.667	100.000	100.000
8	86.207	86.207	96.552

Table B.16: Classification with Cross-validation(%) with Nearest Neighbors, Random Forest and LDA for PAMAP database application.

Subjects Classification Methods	Nearest Neighbors	Random Forest	LDA
0	92.248	93.023	93.023
1	82.540	80.159	87.302
2	92.199	94.326	95.745
3	97.203	96.503	97.902

B.8 Human Activity Recognition with non-controlled approach

Table B.17: ARI accuracy (%) with K-means, Affinity Propagation, DBSCAN and Ward and with window sizes of 3 seconds for the FCHA signals application acquired in a non-controlled environment.

Subjects Classification Methods	Kmeans	Affinity Propagation	DBSCAN	Ward
0	64.833	73.140	57.150	73.795
1	72.144	78.513	75.341	70.890
2	64.776	63.411	63.992	62.607
3	44.198	43.145	41.319	38.780
4	82.118	77.593	79.757	83.975
5	80.526	74.866	80.604	76.857
6	80.961	72.141	74.019	73.378
7	74.976	75.778	76.431	74.911
8	82.034	82.290	83.289	88.995

Table B.18: ARI accuracy (%) with K-means, Affinity Propagation, DBSCAN and Ward and with window sizes of 3 seconds for the FCHA signals application acquired in a controlled environment.

Subjects Classification Methods	Kmeans	Affinity Propagation	DBSCAN	Ward
0	71.088	69.192	74.039	74.793
1	100.000	100.000	100.000	100.000
2	82.570	74.465	62.639	87.387
3	64.869	85.908	78.100	64.869
4	84.012	89.548	75.790	85.689
5	86.709	92.900	76.038	87.322
6	80.569	82.143	80.523	79.674
7	76.834	77.011	72.779	76.353
8	90.617	77.676	78.723	98.973

Table B.19: ARI accuracy (%) with K-means, Affinity Propagation, DBSCAN and Ward and with window sizes of 5 seconds for the FCHA signals application acquired in a non-controlled environment.

Subjects Classification Methods	Kmeans	Affinity Propagation	DBSCAN	Ward
0	62.911	64.166	64.626	65.778
1	76.408	60.849	73.843	76.000
2	67.902	61.386	60.598	72.638
3	43.703	37.264	40.296	32.356
4	82.284	79.622	77.556	83.974
5	77.023	54.779	77.888	80.511
6	77.293	63.097	77.553	74.960
7	79.117	79.960	53.407	74.865
8	83.685	82.949	66.276	77.155

Table B.20: ARI accuracy (%) with K-means, Affinity Propagation, DBSCAN and Ward and with window sizes of 5 seconds for the FCHA signals application acquired in a controlled environment.

Subjects Classification Methods	Kmeans	Affinity Propagation	DBSCAN	Ward
0	70.173	74.855	74.433	71.493
1	100.000	91.780	100.000	100.000
2	79.956	73.785	75.700	84.595
3	87.530	87.530	91.883	87.530
4	86.278	77.064	83.097	88.221
5	86.615	78.844	78.844	82.293
6	84.131	86.368	86.368	86.368
7	78.834	76.765	76.765	69.539
8	86.416	83.722	87.167	92.506

C

8th International Joint Conference on Biomedical Engineering Systems and Technologies - BIOSTEC 2015 and Procedia Computer Science - Elsevier

Human Activity Recognition based on Novel Accelerometry Features and Hidden Markov Models application

Ana Luísa Gomes[1], Vítor Paixão[3] and Hugo Gamboa[1,2]
[1]*Universidade Nova de Lisboa Faculdade de Ciências e Tecnologias, FCT-UNL, Lisbon, Portugal*
[2]*PLUX - Wireless Biosignals, Lisbon, Portugal*
[3]*Champalimaud Neuroscience Programme, Champalimaud Centre for the Unknown, Lisbon, Portugal*
alg.gomes@campus.fct.unl.com, hgamboa@plux.info,vitor.paixao@neuro.fchampalimaud.org

Keywords: Human Activity Recognition, Forward Feature Selection, Hidden Markov Models, Clustering.

Abstract: The Human Activity Recognition (HAR) systems require objective and reliable methods that can be used in the daily routine and must offer consistent results according to the performed activities.
In this work, a framework for human activity recognition in accelerometry (ACC) based on our previous work and with new features and techniques was developed. The new features set covered wavelets, the CUIDADO features implementation and the Log Scale Power Bandwidth creation. The Hidden Markov Models were also applied to the clustering output. The Forward Feature Selection chose the most suitable set from a 423^{th} dimensional feature vector in order to improve the clustering performances and limit the computational demands. K-means, Affinity Propagation, DBSCAN and Ward were applied to ACC databases and showed promising results in activity recognition: from $73.20\% \pm 7.98\%$ to $89.05\% \pm 7.43\%$ and from $70.75\% \pm 10.09\%$ to $83.89\% \pm 13.65\%$ with the Hungarian accuracy (HA) for the FCHA and PAMAP databases, respectively. The Adjust Rand Index (ARI) was also applied as clustering evaluation method. The developed algorithm constitutes a contribution for the development of reliable evaluation methods of movement disorders for diagnosis and treatment applications.

1 INTRODUCTION

Over time, the increasingly demand for objectivity in clinical diagnosis and the continuous pursuit for human wellbeing led to the development of enginery for healthcare. The combined efforts of medicine and engineering created and developed techniques that provide large amounts of information and simultaneously allow to interpret the generated data. According to different studies and our previous work, accelerometry is a reliable system for monitoring and evaluate daily physical activities over time(Inês Prata Machado, 2014), (Nishkam Ravi and Littman, 2005), (A. M. Khan and Kim, 2010). In this study, a framework for HAR systems was developed and tested with different accelerometry databases acquired with a triaxial accelerometer.
Biosignal processing requires an acquisition stage and a transformation with conversion, filtering and extraction of the useful features, which will depend on the aim of the investigation. The feature extraction step becomes very important for activity recognition because it defines what information we will cluster with. The selected features are directly related with the information extracted from the ACC signals which allows data organization inside each cluster by clustering algorithms. The clustering organization must show a lower variation between similar activities than between different activities (Lin and Chen, 2005), (Nishkam Ravi and Littman, 2005).

1.1 UNSUPERVISED LEARNING METHODS

Several techniques for data acquisition and processing have been developed to improve the early diagnosis and to aid clinical treatment of various diseases. ACC signals processing shows the importance of objective monitoring human locomotion through movement quantification when the medical diagnosis of pathologies is subjective and hard to trace such as Parkinson's disease and Cerebrovascular Accident (CVA).
U. Maurer and coworkers (Uwe Maurer and Deisher, 2006) and our group (Inês Prata Machado, 2014) con-

cluded that four features from time and frequency domains can achieve high activity recognition accuracy (about 99%). The work developed in (Inês Prata Machado, 2014) also contributed for the physical activity (PA) recognition in accelerometry with the K-means technique and also inferred that one waist-worn accelerometer can identify the physical activities in an adequate manner. Furthermore, (Adrian Ball and Velonaki, 2011) states that not only k-means clustering has a good performance but also the spectral clustering and the affinity propagation approaches show high accuracy results.
In the K-means method, the k value (number of clusters) is defined and k points are chosen randomly as cluster centers (Ittay Eyal and Rom, 2011),(Ghahramani, 2004), (Liao, 2005). Besides the K-means method, other clustering methods were applied in the present work such as Affinity Propagation (Walter, 2007), DBSCAN and Ward (Liao, 2005). The Affinity Propagation clusters dataset sending messages between pairs of samples until they converge. These messages represent the suitability for one data point to be the exemplar (similar) of the other, which is updated in response to the values from other pairs (Walter, 2007). In the DBSCAN method, clusters are areas of high density, separated by regions of low density, while the Ward method is based on the minimal variance criterion where two clusters will be agglomerated into one when a defined value is not achieved. Otherwise, both clusters will be apart (Liao, 2005).
Clustering techniques applied to biosignals morphology knowledge allows the detection and classification of different physical positions and everyday movements. The clustering stage is crucial for pathology detection and abnormal behavior evaluation (Nunes, 2011) due to changes that can be detected in the morphology of the accelerometry signal. Therefore, it is mandatory to acquire enough knowledge and data in order to be able to distinguish between normal movement patterns and those of certain pathologies.

1.2 HIDDEN MARKOV MODELS

The Hidden Markov Models (HMM) are statistical models where the observation is a probabilistic function of the state. In this case, the observation task is made by inference and the training set will determine the transition probabilities between the existing states (Ping Guo and Wang, 2012),(Fosler-Lussier, 1998).
In (Trabelsi et al., 2013) the HMM were used to identify the sequence corresponding to 12 physical activities and the final results lead to 91.4 % as a mean correct classification rate averaged over all observations. They also concluded that the HMM application leads to a better classification rate (84 %) and with k-means algorithm (60 %). This fact highlights the potential benefit of automatic identification of human activity with the HMM approach.

1.3 CLUSTERING PERFORMANCE EVALUATION

After the clustering and HMM application, it is possible to assess if the separation of the data is similar to the available ground truth set. In an unsupervised learning context, it is important to create a data annotation as a ground truth (A. M. Khan and Kim, 2010). In the present work, two clustering performance evaluation methods were used:

1. Hungarian Accuracy - With two solution sets, the predicted labels and the ground truth set, it is possible to measure the distance between both sets. The labeling of the predicted clusters must correspond to the ground truth available. However, if two partitions of the dataset are equivalent but its labelings are represented with different labels, there will be an ambiguity. To overcome this ambiguity the labelled indices in one predicted solution are permuted in order to increase the agreement between the two solution sets under comparison. This ambiguity can be minimized through the Hungarian method with a matrix construction based on the predicted labels and the ground truth similarity. This performance evaluation method measures the fraction of disagreement between both labels sets through the diagonal of the resulting matrix (Kuhn, 2009).
2. Adjust Rand Index - No conjecture is performed on the cluster arrangement and this technique measures the similarity between the predicted labels and the ground truth set, ignoring permutations. The ARI accuracy ranges from 0.0 (0%) to 1.0 (100%), for a perfect score (Clu, 2014).

It is proposed in the present work a framework implementation for activity recognition through new features and techniques application, presented in figure 1.
New features applied to accelerometry, instead of audio signal alone, might contribute for the discovery of important movement characteristics never detected before, such as the CUIDADO features and wavelets. A new feature inspired in the Mel scale is created and implemented, called Log Scale Power Bandwidth.
The HMM are also applied to the clustering output to improve the final results of the developed framework. Section 2 describes the materials and methods

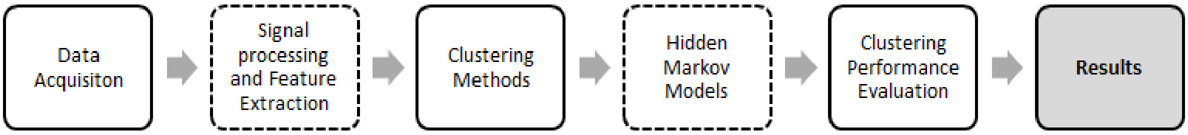


Figure 1: Overall structure of the framework developed for HAR systems. After the ACC data acquisition, the signal processing and the feature extraction stages are carried out and its results are used in clustering methods and in the HMM application. Finally, the clustering performance evaluation is applied. The hatched blocks show novel approaches for activity recognition.

adopted in this work to extract the ACC data for motion analysis. Section 3 describes the implementation of some algorithms that form the developed framework, including the feature selection method and the Log Scale Power Bandwidth implementation. Section 4 shows the results and respective discussion. Section 5 presents the main conclusions and the take home message obtained with this work.

2 METHODOLOGY AND MATERIALS

Two databases were analyzed in the present work: the online available PAMAP database (PAM, 2014) and the Foundation Champalimaud Human Activity (FCHA) database, described in the following subsections.

2.1 FCHA Database

Seven tasks were carried out by 9 volunteers with an age range from 23 to 44 years old: standing, sitting, lying down, walking, running, and ascending and descending stairs. All activities were performed with a predefined order and time, excluding the ascending and descending stairs tasks. The walking and running activities were carried out in a exercise treadmill with predefined velocities (4 km/h and 10 km/h respectively).
A triaxial accelerometry sensor was located on the waist with an acquisition frequency sampling of 800 Hz and a resolution of 16 bits. The ACC data acquired in this protocol formed the FCHA database. Data acquisition was carried out in the Champalimaud Centre for the Unknown with the OpenSignals software (Ricardo Gomes and Gamboa, 2012).

2.2 PAMAP Database

The PAMAP database is available at (PAM, 2014) and was also analyzed alongside the FCHA database in order to verify that the framework created in this work may suggest encouraging performances even from acceleration data with different resolutions. The PAMAP signals were acquired with a sampling frequency of 100 Hz and a resolution of 13 bits. The PAMAP signals show several physical activities and nine were selected: standing, sitting, lying down, walking, running, ascending and descending stairs, jumping and cycling. The data was acquired from 8 volunteers within an age range 25-31 years and the 3D-accelerometer sensor used was placed in the chest. All movement tasks were performed at a variable rhythm, according with each subject in order to acquire data in the most realistic conditions as possible.

2.3 Annotation Stage

The annotation document concerns all the labels and times of each activity performed by a given volunteer. The initial and final time of each activity was annotated in a JSON file created for each acquired signal. In addition to the annotation task adopted, the present work added an extra stage where motion series of all volunteers were videotaped in order to avoid erroneous times or labels and for ground truth validation (Figure 2).

Figure 2: Frames of the subject08's videotape, performing four tasks from the protocol: running, lying down, climbing stairs and cycling.

3 PROPOSED FRAMEWORK

Several algorithms were implemented in this work, including new previously unused features and the feature selection method. Algorithms such as the segmentation process and the feature design stage, used in the present work, were developed previously by our group (Inês Prata Machado, 2014).

3.1 Feature Implementation

There are several features already suggested in other studies and applied to accelerometry, such as the Mean and Standard Deviation (Adrian Ball and Velonaki, 2011),(Ittay Eyal and Rom, 2011), (Godfrey and ÓLaighin, 2008), listed in Table 1 with [1].

It is possible to group features according to different parameters, such as time, statistical and frequency domains. However, in the frequency spectrum analysis, the FFT does not provide information about the time at which these frequency components occur, which leads to the need for a tool that allows us to analyze the signal on both domains. A wavelet is a specific technique for the time-frequency domain and allows the visualization of the frequency content over time and consequently a better transient event description of an accelerometry signal (Godfrey and ÓLaighin, 2008), (Jani Mantyjarvi and Seppanen., 2001). The approximation coefficients from wavelets decomposition reflect the main characteristics of the signal and these values are used as feature coefficients in this work (Demetrio Labate and Wilson, 2013), (Galka and Ziólko, 2008).
There are other features, called the "CUIDADO features", applied for the first time for audio signals by G. Peeters in (Peeters, 2004) and can be applied and useful for accelerometry studies. Some from the CUIDADO features, shown in the Table 1 with [2], were used in this work. For more information about these, see (Peeters, 2004).
Table 1 shows all features from the four domains analyzed in this work.

3.1.1 Log Scale Power Bandwidth

In the present work, the lower frequencies were studied in more detail than higher frequencies through logarithmic scales in order to analyze meticulously the frequency domain. This study was inspired by the audio spectrum and the Mel scale which ultimately lead to the feature Log Scale Power Bandwidth creation.
The Log Scale Power Bandwidth coefficients are computed and its input is the motion data. This algorithm concerned five stages:

1. The first stage was the pre-emphasizing of the signal in the time domain. This stage filters a data sequence (the input segment signal) using a digital filter which emphasizes the energy of the signal at high frequencies with a pre-emphasis factor of 0.97;

Table 1: List of all features used in the present work and respective domains and number of output coefficients for each acceleration component: x, y, z, and total acceleration. [1] Refers to all traditional features already applied in accelerometry (Inês Prata Machado, 2014); [2] Refers to the CUIDADO features used in audio recognition; [3] Refers to the new feature type created and implemented in this work.

Feature Type	Number of Output Coefficients (for each acceleration component)
Statistical	
Skewness [1]	1
Kurtosis [1]	1
Histogram [1]	10
Mean[1]	1
Standard Deviation[1]	1
Interquartile Range[1]	1
Time	
Root Mean Square[1]	1
Median Absolute Deviation[1]	1
Zero Crossing Rate[1]	1
Pairwise Correlation[1]	3 (in total)
Autocorrelation[1]	1
Temporal Centroid[2]	1
Variance[2]	1
Frequency	
Maximum Frequency[1]	1
Median Frequency[1]	1
Power Spectrum[1]	2
Fundamental Frequency[1]	1
Power Bandwidth[1]	10
Log Scale Power Bandwidth[3]	40
Total Energy[2]	1
Spectral Centroid[2]	1
Spectral Spread[2]	1
Spectral Skewness[2]	1
Spectral Kurtosis[2]	1
Spectral Slope[2]	2
Spectral Decrease[2]	1
Spectral Roll-off [2]	1
Time-Frequency	
Wavelets [2]	20

2. The second step refers to the framing which divides the input data into a set of 3 (M) frames, each of these with 256 (N) samples;

3. Next, the conversion of the signal segment into the frequency domain is carried out through the Fast Fourier Transform application. However, whenever a finite Fourier Transform is applied and if the start and end of the finite data do not match, there will be a discontinuity in the signal. In this case, there will show up nonsense and undesirable high-frequencies in the Fourier transform. There-

fore, a windowing stage was computed to the data sample with a Hamming window to make sure that the ends match up;

4. A set of triangular overlapping windows in the range 133-3128 Hz was created. This set of triangular filters was spaced linearly at lower frequency, below 199 Hz, and logarithmic spaced above 199 Hz;

5. The triangular filter bank was applied to the resulting data from the step 3 which gives the powers at each frequency. Finally, the algorithm computes the log (in base 10) of the powers at each frequency and returned the Log Scale Power Bandwidth coefficients as the amplitudes of the resulting spectrum.

3.2 Feature Selection

Feature normalization to zero mean and unit variance is adopted before creating any feature selection. Next, the most suitable features for activity recognition were identified once the feature computation is a time consuming and computationally heavy task.
The protocol for feature selection is based on the Forward Feature Selection protocol and aims to select 10 features at most for each clustering method used. This study may be described by the following steps:

1. Elimination of the redundant information: Correlation between all features and removal of the redundant features. The second feature is removed when two features show correlation values greater than or equal to 0.98. The resulting set from this correlation and elimination stage is called A;

2. Selection of the best fitting features: 20 features with the highest ARI value are chosen among the set A, named set B. From the new set formed by 20 features types, B, the feature type with the highest ARI performance is collected to the set C, which leaves the original set B with only 19 features. Next, the algorithm combines the set C with the existing feature types from the set B. The combination with the highest ARI value and the corresponding set are identified. The new feature belonging to B and to the identified set is collected by C. In each iteration, a new feature is deleted from B and is collected to C. This iterative procedure repeats until C shows the best combination of 10 features;

3. Saving the final results: The algorithm finishes the procedure and saves the name and ARI performance of the 10 features set.

3.3 Hidden Markov Models Application

HMM were applied to the clustering results to achieve higher ARI accuracies in activity recognition. Its application was formed by the training and testing stages.
The implemented algorithm uses the ground truth (true labels) by collecting frequencies of the transitions between all different activities/states and also defines the initial state of a given sequence of activities through the most frequent initial state. The frequencies recorded are then converted to the probabilities of the existing symbols and state sequences. Finally, the testing stage estimates the most probable sequence of hidden states based in the trained model and with the Viterbi algorithm.
In the present work, the HMM topology is a completely connected structure of an ergodic model and it uses the labels from the clustering methods as test set and the annotation data (ground truth) as training set. Finally, the HMM output is used in the clustering performance evaluation stage.

4 Results and Discussion

All studies here presented used defined parameters according to the highest ARI accuracy for each of these, such as: window segmentation, filtering stage and wavelet level decomposition. Segments with time duration of 5 seconds were used as window segmentation. The 8^{th} level was selected as the best level decomposition for the wavelets algorithm and the filtering stage was eliminated from the signal processing stage. The referred parameters were used in all studies presented in the following sections, 4.1, 4.2 and 4.3.
Clustering methods were also selected according to the ARI performances in order to improve the clustering accuracy. The K-means, Affinity Propagation, DBSCAN and Ward showed the highest ARI for ACC data. In parallel to the unsupervised learning techniques, three classification methods were used: K-Nearest Neighbors, Random Forest and Linear Discriminant Analysis (LDA).

4.1 Best set of Features

Feature selection is formed by two stages. The first part aims to find the best number of feature types for activity recognition and second stage aims to identify which feature types must be used.

The best 10 features were computed from the Forward Feature Selection Algorithm, for each subject and clustering method. Figure 3 shows the ARI performance (%) for each number of features and for each clustering method in order to select the best number of features to use in ACC recognition. This study used as ACC data the FCHA database and all implemented features from four domains: statistical, time, frequency and time-frequency, presented in Table 1.

Figure 3 suggests higher ARI percentages in 4 to 7 features. For a more detailed analysis, table 2 shows the ARI performances achieved with different clustering methods when the framework uses the best 4 to 7 feature types (sets A, B, C and D respectively) and all features.

From table 2, it can be concluded that the set B is the best set for K-means and Ward performances with 89.97% $\pm$ 9.97% and 88.56% $\pm$ 11.72%, respectively. On the other hand, the set C showed higher performance for the DBSCAN method with 80.43% $\pm$ 6.29% while the set D showed best performance for the Affinity Propagation with 81.19% $\pm$ 5.99%. For this reason, any choice from set B, C and D is acceptable. For K-means, Affinity Propagation, DBSCAN and Ward, the clustering performance values are 84.54% $\pm$ 9.23% ,81.19% $\pm$ 5.99%, 79.84% $\pm$ 10.75% and 84.73% $\pm$ 9.00% for set D, 89.97% $\pm$ 9.97%, 76.85% $\pm$ 10.30%, 78.12% $\pm$ 10.95% and 88.56% $\pm$ 11.72% for set B. Overall, sets B and D showed similar computing times (with difference of approximately 13 seconds) and the best accuracy values. In the present study set D is the set of the best 7 features and it was chosen for the formed framework.
After selecting the best number as 7 features, the best group of features was found through a histogram, where the occurrences of each feature type were represented. The 10 most used features in each clustering method were pooled, some of them belonging to the same feature type, presented in the figure 4.

The histogram shown in figure 4 suggested that the Forward Feature Selection algorithm used with a higher frequency the Log Scale Power Bandwidth, Root Mean Square, Total Energy, Autocorrelation, Variance, Wavelet Coefficients and the Mean for HAR systems. Therefore, these feature types are the most used and promising features for the developed framework.

Furthermore, figure 4 suggested that the Log Scale Power Bandwidth occurs more frequently (over 20%) than all the other types of features (with less than 20% in all occurrences). The Log Scale Power Bandwidth algorithm involved complex stages and offered a wide number of coefficients as output. The resulting data from those 40 output coefficients are complementary. One particular coefficient tended to be more sensible in activity distinction due to the variation in behavior over time while the other coefficient may identify better other different tasks. Thus, by using this type of feature, all these coefficients are used together and there will be more information related to the activity distinction compared with other type of features with fewer information and lower number of coefficients.
Moreover, the Log Scale Power Bandwidth feature considered data from the lower frequencies. A detailed analysis in this frequency range suggested that there was important information for activity recognition in accelerometry.The information located at low frequencies was preserved due to the elimination of the filtering step in the signal processing stage. Therefore, no information was lost and the GA component was maintained in the ACC data.

It was possible to observe from figure 5, and in opposition to others features, that each Log Scale Power Bandwidth coefficient showed an overall distinction for all activities carried out by the volunteers. Therefore the choice of this type of feature from the Forward Feature Selection as the best feature is justified for its greater ability for activity recognition.

Some difficulties referred in (Inês Prata Machado, 2014) such as the hard discrimination between sitting and standing positions and between walking and running activities were also identified in this work. These difficulties were subdued due to the presence of the GA component in the processed data and the use of the Log Scale Power Bandwidth and Wavelet coefficients as features. The Horizon Plot in figure 5 showed the variation of six Log Scale Power Bandwidth coefficients, six Wavelet coefficients and one coefficient of the Autocorrelation, Mean, Root Mean Square, Total Energy and the Variance from the x-axis component over time. It is possible to observe that feature types such as Log Scale Power Bandwidth and Wavelets are important for the standing and sitting positions distinction as well in many other tasks.

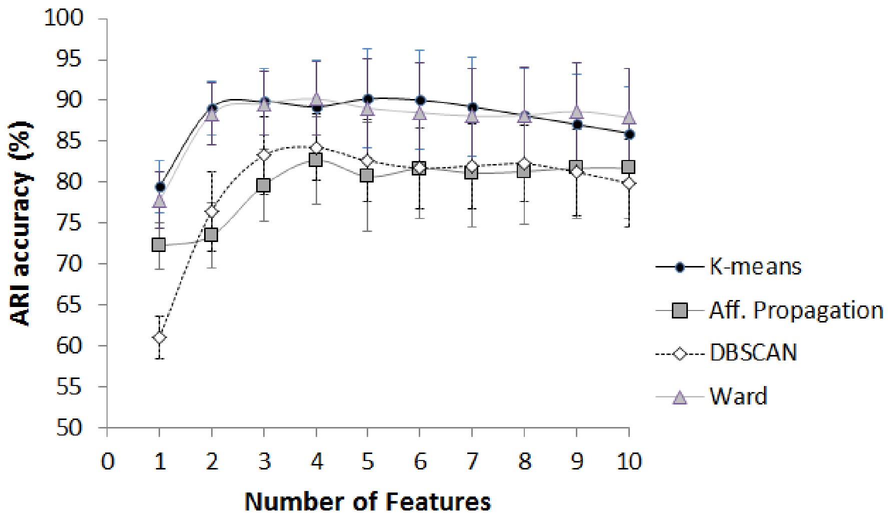


Figure 3: ARI performances (%) as a function of different numbers of features (from 1 to 10 features) and according to different clustering techniques: K-means, Affinity Propagation, DBSCAN and Ward.

Table 2: ARI performances for all features (first column) and for the best 4 to 7 features (set A-second, set B-third, set C-fourth and set D-fifth columns). The last row refers to the time interval used to compute each set of features.

	ARI (%)				
Clustering Methods	All Features	Set A	Set B	Set C	Set D
K-means	87.69 ± 5.56	87.37 ± 12.78	89.97 ± 9.97	84.52 ± 7.56	84.54 ± 9.23
Affinity Propagation	$78.41 \pm 6,86$	76.85 ± 10.30	76.85 ± 10.30	78.33 ± 6.48	81.19 ± 5.99
DBSCAN	78.36 ± 6.96	74.18 ± 9.67	78.12 ± 10.95	80.43 ± 6.29	79.84 ± 10.75
Ward	86.31 ± 8.68	85.96 ± 13.93	88.56 ± 11.72	84.73 ± 9.00	84.73 ± 9.00
Time Response	**347.16**	**107.46**	**132.13**	**142.25**	**145.05**

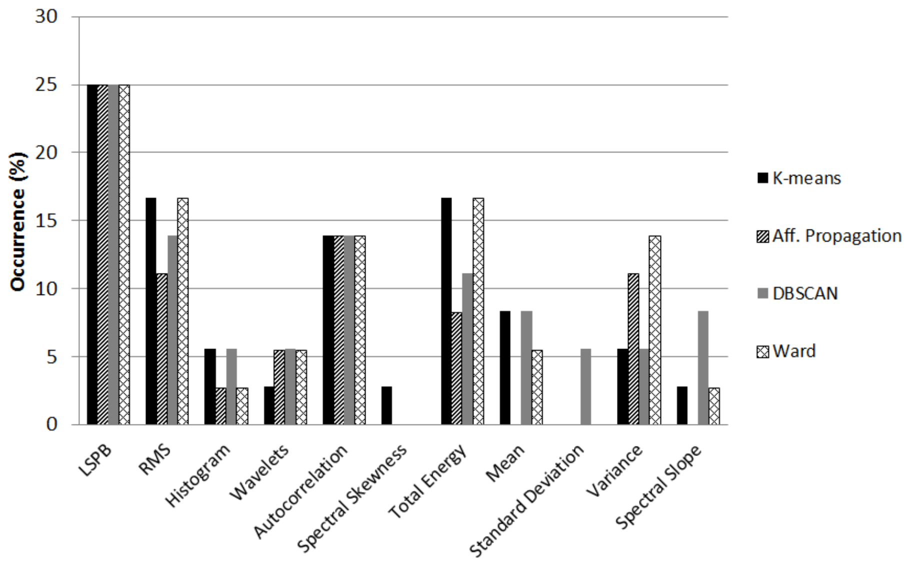


Figure 4: Representation of the Forward Feature Selection results. The algorithm outputted the set of the best 10 features for each clustering method. Each column corresponds to all occurrences of each feature type in all resulting sets for each clustering method.

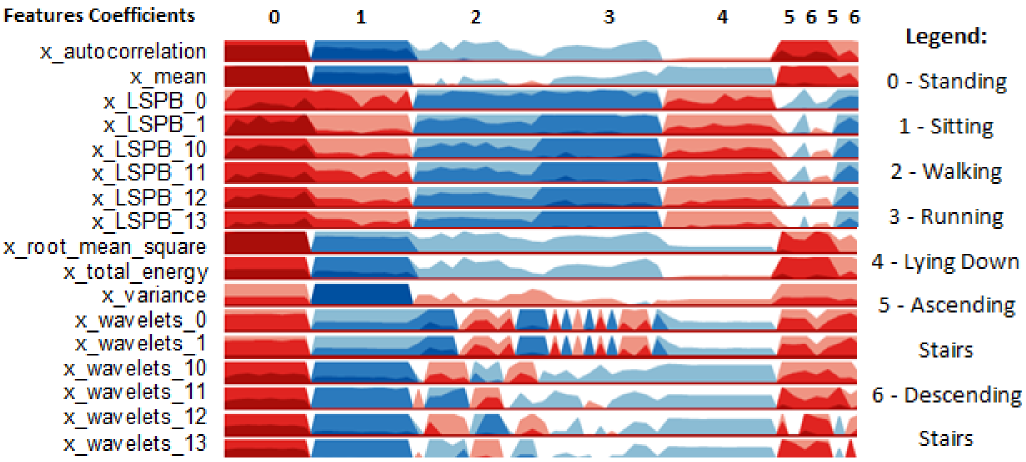


Figure 5: Horizon Plot with some feature coefficients from the X-axis acceleration component which vary over time according to the activity type performed.

4.2 Hidden Markov Models Application

All the existing transitions in the test set (predicted labels from the clustering algorithms) with lower probability of occurrence may be a consequence of cluster miscalculation. These transition probabilities were gathered from the ground truth (train data) and all transitions with low probability of occurrence are avoided and replaced by a more likely transition.
The influence of the Hidden Markov Model application and its improvement (in %) is presented in figure 6 and in table 3. All implemented features were used in this study and only the FCHA database was analyzed. The improvement values shown in the present section is represented by:

$$Improvement(\%) = \frac{100(x_2 - x_1)}{100 - x_1} \qquad (1)$$

where:
x_1 - ARI accuracy value without HMM application
x_2 - ARI accuracy value with HMM application

The obtained improvement values were 33.92%, 1.76%, 17.10% and 9.81%, for K-means, Affinity Propagation, DBSCAN and Ward method, respectively. These results deserve a special attention because there is a significant improvement for certain clustering methods such as the K-means. Thus, it is important to take into account the HMM application together with the adopted unsupervised learning method. All improvement values were positive and in this case the HMM algorithm does not provide heavy computational demands hence it may be applicable to all demonstrated situations.

4.3 Evaluation of the Performance of the Framework

After the framework construction, the FCHA database and the PAMAP database were applied to the developed algorithms set. In this study, unsupervised and supervised learning approaches were applied. The tables 4 and 5 showed all results achieved with the ARI and the HA application and with classification methods: Random forest, LDA and K-Nearest Neighbors, which K value equals the number of activities carried out. The accuracy score method was used in classification techniques in which the ground truth was used as training set and the clustering output as test set.
The PAMAP signals were acquired with a frequency sampling of 100 Hz while the FCHA database is formed by ACC signals acquired with 800 Hz. Thus, this data showed different resolutions which influence the amount of information available for the clustering and classification methods.

Table 4: Clustering evaluation with the ARI and the HA (in %) for K-means, Affinity Propagation, DBSCAN and Ward.

Clustering (ARI %)		
Databases	FCHA	PAMAP
K-means	84.54 ± 9.23	61.56 ± 13.93
Affinity Propagation	81.19 ± 5.99	63.00 ± 0.19
DBSCAN	79.84 ± 10.75	74.26 ± 16.06
Ward	84.73 ± 9.00	60.53 ± 13.70
Clustering (HA %)		
Databases	FCHA	PAMAP
K-means	89.05 ± 7.43	74.47 ± 8.35
Affinity Propagation	73.20 ± 7.98	83.89 ± 13.65
DBSCAN	76.62 ± 9.68	70.75 ± 10.09
Ward	87.10 ± 8.87	71.13 ± 10.37

Table 3: Clustering performances through ARI without and with the HMM application and its improvement.

Clustering Methods	ARI(%) without HMM application	ARI(%) with HMM application	Improvement (%)
K-means	81.37 ± 9.83	87.69 ± 5.56	33.92
Affinity Propagation	78.02 ± 9.90	78.41 ± 6.86	1.76
DBSCAN	73.89 ± 12.98	78.36 ± 6.96	17.10
Ward	84.82 ± 8.17	86.31 ± 8.68	9.81

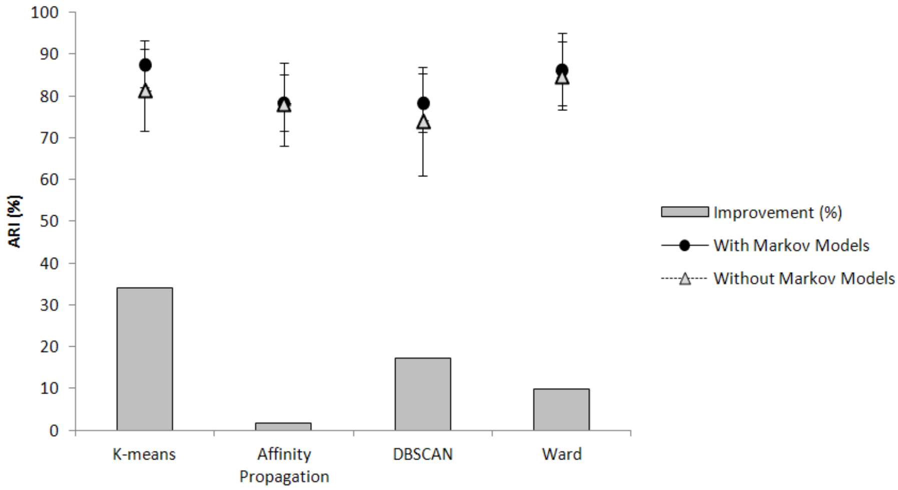


Figure 6: Clustering performances through ARI with and without the HMM application (%) and its performances improvement (%) in bars shown in table 3.

Table 5: Classification accuracy (in %) with K-Nearest Neighbors, Random Forest and LDA methods.

Classification (Accuracy %)		
Databases	FCHA	PAMAP
K-Nearest Neighbors	97.78 ± 6.67	99.40 ± 1.19
Random Forest	95.39 ± 12.64	97.89 ± 3.89
LDA	98.57 ± 4.30	98.03 ± 2.37

The results obtained within the ARI accuracy ranged from $79.84\% \pm 10.75\%$ to $84.73\% \pm 9.00\%$ and from $60.53\% \pm 13.70\%$ to $74.26\% \pm 16.06\%$ for the FCHA and PAMAP databases, respectively. On the other hand, the Hungarian accuracy results ranged from $73.20\% \pm 7.98\%$ to $89.05\% \pm 7.43\%$ and from $70.75\% \pm 10.09\%$ to $83.89\% \pm 13.65\%$ for the same databases. The main cause of the difference between the two databases may be related to the large difference in resolution, since the sampling frequencies for the FCHA base and the PAMAP database are 800 and 100 Hz, respectively. The FCHA data shows eight times more information than PAMAP data, which leads to a higher accuracy values. Unlike clustering, classification uses the ground truth for training and also showed high results: from $95.39\% \pm 12.64\%$ to $99.40\% \pm 1.19\%$ for both databases.
More than 85% of the presented results showed an accuracy higher than 70% which revealed the frameworks robustness and versatility for activity recognition with ACC signals, acquired with different sensors and different resolutions.

4.4 Conclusions

This work aimed to create and develop a novel gesture recognition system based on the consulted literatures concepts and presented in (Inês Prata Machado, 2014).
In the present work and in addition to those used in our previous work new features were implemented, such as the Log Scale Power Bandwidth coefficients. Other features previously used in audio recognition were also used in ACC data such as the CUIDADO features and wavelets coefficients. This work offered a set of 423 feature types for machine learning techniques which provide more and new information regarding the performed movement tasks compared to the literature (Ghahramani, 2004),(Jani Mantyjarvi and Seppanen., 2001), (Nishkam Ravi and Littman, 2005).
The Forward Feature Selection aimed to reduce the undesirable redundancy between features and to select the set of features that may lead to the best frameworks performance. The chosen features selected as the most suitable feature types for HAR systems are the Log Scale Power Bandwidth, Root

Mean Square, Total Energy, Autocorrelation, Variance, Wavelet Coefficients and Mean coefficients.
The achieved results also suggested that it is important not to waste any information regarding to movement. The presence of the information located in lower frequencies, such as the GA component, and the Log Scale Power Bandwidth implementation may lead to better static activity distinction.

The obtained results with the FCHA database and PAMAP database showed that the developed framework is suitable for activity recognition even for ACC data with a large difference in resolution.
The major achievements of the current work allowed to construct a novel HAR system with HMM which may lead to better performances in activity recognition. The created framework with a small number of feature types also ensure high machine learning results without heavy computational demands. Therefore, as expected accelerometry is a suitable technique for monitoring movement patterns in free-living subjects over long periods of time. The knowledge acquired over this thesis may be applied into the clinical setting for the diagnosis and physiotherapy fields.

REFERENCES

(January 2014). Pamap, http://www.pamap.org/demo.html.

(September 2014). Clustering, http://scikit-learn.org/stable/modules/clustering.html.

A. M. Khan, Y.-K. Lee, S. Y. L. and Kim, T.-S. (2010). a triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer. 14(4):11661172.

Adrian Ball, David Rye, F. R. and Velonaki, M. (2011). A comparison of unsupervised learning algorithms for gesture clustering. pages 111–112.

Demetrio Labate, G. W. and Wilson, E. (2013). Wavelets. 60(1):66–76.

Fosler-Lussier, E. (1998). Markov models and hidden markov models - a brief tutorial.

Galka, J. and Ziólko, M. (2008). Wavelets in speech segmentation.

Ghahramani, Z. (2004). Unsupervised learning. pages 3–4.

Godfrey, R. Conway, D. M. and ÓLaighin, G. (2008). Direct measurement of human movement by accelerometry. *Elsevier*, 30(10):1364–1386.

Inês Prata Machado, Ricardo Gomes, H. G. V. P. a. (2014). Human activity recognition from triaxial accelerometer data feature extraction and selection methods for clustering of physical activities.

Ittay Eyal, I. K. and Rom, R. (2011). Distributed data clustering in sensor networks. 24(5):207222.

Jani Mantyjarvi, J. H. and Seppanen., T. (2001). Recognizing human motion with multiple acceleration sensors. 2:747 – 752.

Kuhn, H. W. (2009). the hungarian method for the assignment problem. 2:83–97.

Liao, T. (2005). Clustering of time series data - a survey. 38(11):1857 – 1874.

Lin, R.-S. and Chen, L.-H. (2005). A new approach for audio classification and segmentation using gabor wavelets and fisher linear discriminator. 19(6):807–822.

Nishkam Ravi, Nikhil Dandekar, P. M. and Littman, M. L. (2005). Activity recognition from accelerometer data. 3:1541–1546.

Nunes, N. F. M. (2011). Algorithms for time series clustering. Master's thesis, Physics Department, Faculty of Sciences and Technology, New University of Lisbon and Plux, Wireless Biosignals, Portugal.

Peeters, G. (2004). A large set of audio features for sound description (similarity and classification) in the cuidado project.

Ping Guo, Zhenjiang Miao, X.-P. Z. Y. S. and Wang, S. (2012). Coupled observation decomposed hidden markov model for multiperson activity recognition. 22(9):1306 – 1320.

Ricardo Gomes, Neuza Nunes, J. S. and Gamboa, H. (2012). Long term biosignals visualization and processing.

Trabelsi, D., Samer Mohammed, Faicel Chamroukhi, L. O., and Amirat, Y. (2013). An unsupervised approach for automatic activity recognition based on hidden markov model regression. 10(3):829 – 835.

Uwe Maurer, Asim Smailagic, D. P. S. and Deisher, M. (2006). Activity recognition and monitoring using multiple sensors on different body positions. pages 113–116.

Walter, S. F. (2007). Clustering by affinity propagation. Master's thesis, Department of Physics, ETH Zurich, Switzerland.

Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Procedia Computer Science 00 (2014) 1–9
Procedia Computer Science

Human Activity Data Discovery from Triaxial Accelerometer Sensor: non-supervised learning sensitivity to feature extraction parametrization

Inês P. Machado[a,c,*], A. Luísa Gomes[a], Hugo Gamboa[a,b], Vítor Paixão[c], Rui M. Costa[c]

[a] *Faculty of Sciences and Technology, New University of Lisbon, 2829-516 Caparica, Portugal*
[b] *PLUX, Wireless Biosignals, Avenida 5 de Outubro, 70, 1050-059, Lisbon, Portugal*
[c] *Champalimaud Neuroscience Programme, Champalimaud Institute for the Unknown, Avenida de Braslia, 1400-038, Lisbon, Portugal*

Abstract

Background Our methodology describes a human activity recognition framework based on feature extraction and feature selection techniques where a set of time, statistical and frequency domain features taken from 3-dimensional accelerometer sensors are extracted. This framework specifically focuses on activity recognition using on-body accelerometer sensors. We present a novel interactive knowledge discovery tool for accelerometry in human activity recognition and study the sensitivity to the feature extraction parametrization.

Results The implemented framework achieved encouraging results in human activity recognition. We have implemented a new set of features extracted from wearable sensors that are ambitious from a computational point of view and able to ensure high classification results comparable with the state of the art wearable systems [1]. A feature selection framework is developed in order to improve the clustering accuracy and reduce computational complexity. Several clustering methods such as K-Means, Affinity Propagation, Mean Shift and Spectral Clustering were applied. The K-means methodology presented promising accuracy results for person-dependent and independent cases, with 99.29 % and 88.57 %, respectively.

Conclusions The presented study performs two different tests in intra and inter subject context and a set of 180 features is implemented which are easily selected to classify different activities. The implemented algorithm does not stipulate, *a priori*, any value for time window or its overlap percentage of the signal but performs a search to find the best parameters that define the specific data. A clustering metric based on the construction of the data confusion matrix is also proposed. The main contribution of this work is the design of a novel gesture recognition system based solely on data from a single 3-dimensional accelerometer.

Keywords:
Human Activity Recognition, Interactive Knowledge Discovery, Feature Extraction, Dimensionality Reduction, Clustering Algorithms

*Corresponding author
Email addresses: ip.machado@campus.fct.unl.p (Inês P. Machado), alg.gomes@campus.fct.unl.pt (A. Luísa Gomes), hgamboa@plux.info (Hugo Gamboa), vitor.paixao@neuro.fachampalimaud.org (Vítor Paixão), ruicosta@fchampalimaud.orgt (Rui M. Costa)

[1]The software OpenSignals [2] was used for signal acquisition and signal processing algorithms were developed in Python Programming Language [3] and Orange Software [4].

1. Background

The past two decades have seen an increase in research activity in the field of human activity recognition. With activity recognition having considerably growing so did the number of challenges in developing, implementing and evaluating solutions that may help the medical diagnosis of chronic motor diseases. A major challenge in our networked world is the increasing amount of data, which require efficient and user-friendly solutions with intuitive approaches of data visualization and extraction of relevant information.
The constant concern with the human physical and psychological well being has been the drive for research studies that have led to promising advances in medicine and engineering. In Biomedical Engineering, the demand for objectivity in clinical diagnosis of pathologies so subjective and hard to trace such as Obsessive Compulsive Disorder [5] or Autism Spectrum Disorder [6], has been one of the greatest challenges. The physical activity disturbance is one of the essential signs of psychiatric disorders and many of these, including Depression [7] and Parkinson's Disease [8], exhibit diagnostic criteria that require an assessment of the motor activity changes of the patient. The behavioural classification usually relies on observation and therefore a highly experienced analyst is always needed. Analysing human action is particularly challenging due to its complexity in the non rigid and self occluding nature of articulated human motion. Human body movement shows up to 244 degrees of freedom [9], which leads the modelling of structural and dynamic features for human activity recognition to a complex task.

The current research arises the hypothesis of accelerometry (ACC) as suitable technique for monitoring movement patterns in freely moving subjects during long periods of time. This technique can be used to measure quantitative parameters which may provide clinical insight into the subject health.

This study aims to understand the signals produced by a triaxial accelerometer through the human movement interpretation and clinical relevant parameters identification from the data. Signal processing techniques are implemented with the purpose of examining accelerometer data and finding new information that is difficult to identify from the raw data. We present a method for convenient monitoring of ambulatory movements in daily life, using a portable measurement device employing a single triaxial accelerometer.

This tool is based on an architecture of signal sensor processing, feature extraction, feature selection and clustering algorithms. To achieve the proposed goals, the following steps were implemented:

- Acquire and analyse the data produced by a triaxial accelerometer, placed on different parts of the body, during human movement, creating a database and proposing an annotation structure for motion data;
- Develop a framework for the interpretation of the data provided by an accelerometry monitoring system: design a set of time, statistical and frequency domain features from several research areas such as speech recognition and activity recognition;
- Develop algorithms to extract relevant information from the data;
- Apply clustering techniques based on feature representation of accelerometer data;
- Explore the choice of features and signals window size on the performance of different clustering algorithms;
- Develop a framework to differentiate human activities, such as sitting, walking, standing, running and lying;
- Evaluate the use of the system in daily life settings and the data interpretation framework: intra and inter subject context discovery;
- Implement a new metric based on classification performance evaluation: classification-based evaluation.

We present a method for convenient monitoring of detailed ambulatory movements in dailylife through a portable measurement device which employs a single triaxial accelerometer and machine learning techniques.

2. Related Work

Most approaches to activity recognition, using body-worn accelerometers, involve a multi-stage process [10]. Firstly, the sensor signal is divided into a number of small time segments, referred to as windows, each of which is considered sequentially. For each window, one or more features are derived to characterize the signal. These features are then used as input to a clustering algorithm which associates each window with a cluster.

2.1. Composition of Triaxial Accelerometer Signal

Before using the accelerometer system in any monitoring context, and before the algorithms development to interpret data recorded by the system, it is mandatory to understand the nature of the signals produced by the triaxial accelerometer unit. The signal measured by each fixed-body accelerometer is a linear sum of, approximately, three components [11]:

- Acceleration resulting from body movement - Body Acceleration Component;
- Acceleration resulting from gravity - Gravitational Acceleration Component;
- Noise intrinsic to the measurement system.

The first two components provide different information about the wearer of the device: the gravitational acceleration (GA) provides information about the space orientation of the device, and the body acceleration (BA) provides information about the movement of the device. The separation of the information regarding the movement of the device is important, however these two components have overlapping frequency spectra. According to previous studies, the BA component ranges from 0 Hz to an estimation 20 Hz, but it is mostly contained in the range above 0 and below 3 Hz [11]. This range overlaps the area covered by the GA component, which goes from 0 to 0.3 Hz. It is possible to approximately separate the BA and the GA components by filtering techniques. Figure 1 represents the motion data processing of a typical recording from the accelerometer, showing several minutes of recorded data during a supervised test where the subject performs specific tasks. Mathie and co-workers (2003) have tested a wide range of different filters types with different characteristics and different windowing percentages in order to determine their ability to differentiate the components of the acceleration signal. In the present study, in order to isolate the body acceleration component, a second-order Butterworth High-Pass filter with cut-off frequency of 0.25 Hz was used, as in [11].

3. Methodology

Triaxial accelerometers produce three separated accelerometer data time series, one time series for acceleration on each axis ACC_x, ACC_y and ACC_z. Complementary to the three axes data, an additional time series, ACC_{tot}, have been obtained by computing the magnitude of the acceleration:

$$ACC_{tot} = \sqrt{ACC_x^2 + ACC_y^2 + ACC_z^2} \tag{1}$$

Each time series ACC_i, with $i = x, y, z$ was filtered with a High-Pass Butterworth filter in order to separate the low frequency component and the high frequency component, as suggested in [11] and [12]. This way, for each time series, three more time series BA_i are obtained, with $i = x, y, z$, representing the time series for body acceleration component. Finally, the features for each time series are extracted.

3.1. Feature Extraction

Features can be defined as the abstractions of raw data since they are reduced sets of original raw data which represent main characteristics and behaviours of the signal [10]. The reduced subset of large input data can be called feature vector, it contains important hints for the activity to be recognized and it is the main input for clustering algorithms. Tests are performed in order to assess the following parameters:

- The influence of the signal window size on the clustering performance.

- The influence of the free parameters in that same performance.
- The best feature combination that leads to a better performance of the implemented algorithm.

The present study investigates a new method of feature extraction for clustering problems. A dictionary of features extracted from the motion data was created in JavaScript Object Notation (JSON) format [13]. For each feature, the following information is collected: Description, Imports, Use, Free Parameters, Parameters, Number of Features, Function, Source and Reference, as shown in Table 1.

3.2. Relevant Features for Activity Recognition Systems

Recognizing human activities depends directly on the features extracted for motion analysis. People tend to perform the same movement in many different ways, which can lead to substantial variability in the features derived from body-fixed sensor data. Therefore, to achieve effective clustering, the identification of features with high discriminative ability is important. A good feature set should show little variation between repetitions of the same movements performed by the same individual but should vary between different subjects performing the same task. This set must also show considerable variations between different activities. Table 2 shows the list of features considered in this study.

The developed dictionary divides the features into the following categories: statistical, temporal and spectral. Some of them have been intensively investigated in previous studies and proved to be useful for activity recognition. Others, like an adapted Mel-frequency cepstral coefficients (MFCC) have appeared for the first time in activity recognition and have revealed to be quite promising due to the achieved performances. By manipulating this dictionary, it is possible to reproduce the clustering tests with different feature combinations. For each signal, three new vectors were created: one with the feature information per window, another with the names of the features that were extracted in the respective clustering test and another with the label of the activity that corresponds to each window (the available groundtruth). Therefore, a 180th - dimensional feature vector is obtained: from each one of the four signal vectors, we compute fifteen features with only one output and three features (histogram, cepstral coefficients and mel-frequency cepstral coefficients with ten outputs each.This feature vector has the ability to describe the characteristics of any instance. Features are then fed to a module which implements a specific clustering algorithm.

3.3. Feature Normalization

The obtained features all have different magnitudes. This can cause problems for some machine learning algorithms, where features with higher magnitude will, *a priori*, be given a higher emphasis. Applying a normalizing step before clustering can counter-act this unwanted effect. Because the scale factors and units of the described features are different, before proceeding to the feature selection stage, all features are normalized to zero mean and unit variance.

3.4. Graphical Perception of Feature Visualizations: The Horizon Plot

Previous studies investigated techniques for the visualization of time series data and evaluate their effect in value comparison tasks. Line charts were compared with horizon graphs - an efficient time series visualization technique across a range of chart sizes [14]. To study the behaviour of each feature throughout different activities, horizon graphs were used. This procedure ensures a visual perception of the features that better separate certain activities, those who do not change their value between activities and those who only add redundant information.

In a 180^{th}- dimensional feature vector, it becomes difficult to have a relative perception of the importance of each feature to cluster different activities. Figure 2 represents an example of a horizon graph generated for a feature vector, resulting from an ACC signal composed by seven distinct activities: standing, sitting, walking, running, lying down (belly up), lying down (right side down) and lying down (left side down). It is possible to analyse the behaviour of each feature in each activity. First, the area between data curve and zero y-axis is filled in so that dark reds are very negative and dark blues are very positive. Then, negative values are flipped and coloured red, cutting the chart height by half [14]. Finally, the chart is divided into bands and overlaid, again halving the height.

3.5. Feature Selection Techniques and Free Parameters

To systematically assess the usefulness and identify the most important features to discriminate different activities, a simple feature selection technique was used. Not all features are equally important for a specific task and some of the variables may be redundant or even irrelevant.

The feature selection stage chooses a smaller subset of the original features, which is useful to identify the informative features, and to limit computational demands when executing the recognition system on new observations. A matrix with all possible combinations of the free parameters was created and for each signal, similarly to Holzinger et al. [15], tests were made in order to infer the impact of the window size and to obtain a better performance of the implemented algorithm. The free parameters are those used as input to the features algorithms such as the number of bins in the histogram.

The choice of features acquired from a data set and the window length over which these features are computed is of high importance. For activity recognition, where accelerometer data from PA is windowed, the choice of the number of frames is guided by a trade-off between two aspects: information and resolution [16].Instead of a window size preselection, a combination of different values distribuited in a logarithmic scale were used. According to Table 3, tests were performed with window size ranging from 1000 to 4000 samples, in a logarithmic scale due to its detailed study on the lower frequencies than in the higher frequencies. For each window size, different clustering performances were obtained that allow the choice of an appropriate window size to feature extraction.

3.6. Signal Annotation Method

In the present study, an aspect of activity recognition that was explored was the method applied to annotate sample data, which can be used to compute the performance of the activity model. Most of the researchers have published results of experiments in which the participants are required to manually annotate each activity performed in a given moment [17, 18]. In other cases, the experimenters told the participants in which order the specified activities should be performed, so the correct activity labels were identified before the sensor data was even collected. Still in other studies, the raw sensor data was manually inspected in order to annotate it with a corresponding activity label [19]. If true class labels (ground truth) are known, the validity of a clustering can be verified by comparing the predicted labels and the true labels. For each signal, an annotation, in JSON format [13] was created, as shown in the Figure 1.

$$
\begin{array}{lcl}
"Labels & = & [l_1, l_2, l_3, \ldots, l_i] \\
"Initial\_Times" & = & [init_1, init_2, init_3, \ldots, init_i] \\
"End\_Times" & = & [end_1, end_2, end_3, \ldots, end_i]
\end{array}
$$

Figure 1. Ground Truth: Annotation Structure.

In the presented study, participants were continuously observed during experiments and an observer was stating starting/ending time of each activity. The subjects know in which order the specified activities should be performed and latter, raw sensor data was manually inspected in order to annotate it with a corresponding activity label. The annotation dictionary has information about the number and label of the movements that took place and the time intervals that delimit them. Each label corresponds to one, and only one, activity, regardless of the subject. To compute this annotation format, the implemented function receives an array with the initial and the final times of each activity. It also receives as input the window size of the signal whose features will be extracted and the overlap percentage of the signal to be considered.

4. Experimental Setup

4.1. Unsupervised Learning

Machine learning algorithms based on the feature representation of accelerometer data have become the most widely used approaches in activity prediction [20]. In this study, unsupervised learning is used to distinguish different activities. Clustering mechanisms separate and organize unlabeled data into different groups whose members are similar to each other in some metric. A good clustering process returns a set of clusters in which the intra-class

similarity is high and the inter-class similarity is low. In this study several clustering methods were tested: the K-Means, the Affinity Propagation, the Mean Shift and the Spectral Clustering techniques.
The K-Means clustering algorithm [21] gives a single set of clusters, with no particular organization or structure within them. The clustering tests were performed, individually, for each subject and with the respectively concatenated data. The affinity propagation method separates the dataset into clusters through messages between data samples which characterize the suitability for one data point to fit in the cluster of the other sample. These messages are updated in response to the values from other pairs till convergence [22]. Similarly to the previous clustering method, the Mean Shift also infers the number of clusters and studies the features distribution as a probability density function, where each point with relative high density corresponds to the clusters centers. This method computes the mean between each maximum point and each data sample. The window between these two points is moved to the mean value. The process is iterative and ends when the final clustering is given [23]. Finally, the spectral clustering requires the number of clusters and it uses the similarity matrix between data samples to compute the best eigenvalues for posterior clustering [24].

4.2. Classification-based Evaluation: Proposed Metric

Evaluation of unsupervised approaches is usually difficult due to the lack of ground truth to which one can compare the discovered structure. The presented activity recognition method includes three stages:

- Clustering the data into homogeneous groups.
- Creating rules that connect instances to the correct clusters.
- Recognizing activities inside the clusters.

A confusion matrix contains information about true and predicted labels from a clustering system. The performance of such systems is commonly evaluated using the data in the matrix. Once the clustering algorithm randomly associates the clustering results to non-annotated groups, a novel algorithm, that links these groups to their corresponded activity, is implemented. This function receives the confusion matrix of random assignment and goes through each row of the matrix and stores the index that contains the maximum value of each row. It is checked whether the index is unique throughout the matrix. If the index is unique, it makes the direct correspondence between the vector of true and predicted labels. Otherwise, it checks the index with the maximum value, and assigns it. The process is recursively repeated. After obtaining the assignment vector, the matrix with the labels already associated is reconstructed.

In unsupervised learning, unlabelled data is grouped into different clusters that reflect the underlying structure of the data, based on similarity groups within the data. The input data points are allocated to one of the existing clusters according to the square of the Euclidean distance from the clusters, choosing the closest. This method receives the number of clusters to form as well as the number of centroids to generate. In the presented study, the number of clusters was defined, a priori, from the designed protocol of the performed activities. Figures 3 ,4, 5 and 6 show the structure of clusters according to some of the best set of features: mean, autocorrelation, root mean square and Mel-Frequency Cepstral Coefficients (MFCC).

Activities which are static in nature including postures, such as lying and standing, are easier to recognize than the activities which are periodic in nature, such as running and walking. However, postures that are highly similar, such as sitting and standing, are also very hard to discriminate as they overlap significantly in the feature space.

5. System Architecture and Data Acquisition

In the present study study, the experiments have been carried out with a group of 8 volunteers within an age bracket of 16-44 years. All procedures were approved by the Ethics Committee of the Champalimaud Foundation. The first test consisted in the performing of a gym circuit, in a supervised atmosphere.

Each person performs seven activities in sequence lasting about one minute each - standing, sitting, walking, running and lying down (belly up, right side down and left side down) - wearing an accelerometer on the waist and another on the wrist. Table 4 enumerates the recorded movements of this investigation. Using this system, data with 3-axial acceleration at a constant rate of 800 Hz and 12 bits of resolution was acquired (MotionPlux [25]).

Firstly, the clustering performance for each feature was obtained. Then, tests were carried out with the best set of features. Given the knowledge of the ground truth class assignments (labels true) and the clustering algorithm assignments of the same samples (labels predicted), the Adjusted Rand Index (ARI) is a function that measures the similarity of the two assignments, ignoring permutations and with chance normalization [26].

5.1. Influence of Window Length on the K-Means Performance

Different segmentation methods can be applied to time-series data which enhance signal behaviour and enable the gathering of useful information from continuous stream of data. The first step in the feature extraction process is to divide the acceleration stream in to frames. In this framework, each segment consists of 1000, 2000 or 4000 samples. This time interval proved to be sufficient for analysing the proposed activities (walking, sitting, standing, running, and lying down). Table 5 shows the obtained performance for each value of window size, considering the best implemented set of features: mean, autocorrelation, root mean square and MFCC. An average of the performances obtained for the 8 subjects was calculated.

Based on these results, the Human Activity Recognition (HAR) system reaches a clustering accuracy between 89.46% ±0.4% and 99.14% ±0.6%, with 1000 and 4000 samples, respectively, and a classification accuracy between 81.94% ± 0.7% and 99.29% ± 0.5%. If the subject is performing a single activity for a long time interval, a long window will include more information about that activity. In this study, the windowed data contains a single activity during about one minute, which is how laboratory data usually is collected (long bouts of a single activity) so a long time interval allows a better clustering performance.

6. Results and Discussion

The performance of the proposed HAR system was validated in two studies: subject-independent context and subject-dependent context. Different features and free parameter combinations were explored in order to improve the classification accuracy on physical activity measured from waist accelerometer data. The conditions in which data was collected from movements performed in a supervised laboratory setting are described in this section.

6.1. Clustering Evaluation for Subject-independent Context

The subject-independent performance was evaluated using the K-Means Clustering Algorithm. Even in the case of healthy people, the performed movements can be highly individualistic. Nonetheless, there are a set of basic parameters that are similar to all instances of these movements. A person-independent accuracy of 88.57% and standard deviation of 4.0% was obtained, with window size of 4000 samples and the best set of features: mean, autocorrelation, root mean square and MFCC. Compared to the subject-dependent case, the accuracy is much lower which can be explained by the variations in human motion for different subjects. Each experiment was repeated eight times.

In addition to the K-means application, other clustering methods have been used to corroborate the obtained results, such as the Affinity propagation, the Mean Shift and the Spectral Clustering (figure 4). The Affinity Propagation application presents 32.74%, the Mean Shift shows 44.01% and with the Spectral Clustering 45.74%, for inter subject context.

A confusion matrix contains information about true and predicted labels done by a clustering system. Table 6 shows the confusion matrix for all subjects data. The algorithm successfully distinguish all activities but in 28.3%±6.9% of the cases, the algorithm confuses sitting with standing position.

6.2. Clustering Evaluation for Subject-dependent Context

This procedure investigates how the recognizer performs in a subject's dependent context. To evaluate the subject-dependent accuracy of the proposed algorithm, the K-Means Clustering Algorithm [26] was performed for each subject data. An average person-dependent accuracy of 99.29% and standard deviation of 0.5% were obtained, with window size of 4000 samples and the best set of features: mean, autocorrelation, root mean square and MFCC. On the other hand, the Affinity Propagation application shows 89.17%±7.96 %, the Mean Shift presents 53.04%±20.71 % and with the Spectral Clustering 93.96%±4.87 %.

7. Conclusions

The major achievements of this work are: the design of a gesture recognition system based solely on data from a single 3-dimensional accelerometerwhich finds the best free parameters and window sizes to clustering analysis; the new set of features extracted from wearable data able to ensure high classification results; and the clustering method application in intra and inter subject context.

The experimental results presented here suggest that using only one waist-worn accelerometer can adequately identify user's activities. Using only a minimal number of sensors in wearable activity recognition system can be a key for the success in system acceptance. Furthermore, the present study investigated several combinations of features used for activity recognition to find the best features able to distinguish physical activities. A new set of features has been taken into account and compared to the most commonly features used in literature for activity recognition [1]. The techniques that operate on the statistical, time and frequency domains, as well as on data representations that can be used to discriminate between user activities such as Horizon Plot were described. A set of 180 features is implemented and these are easily selected to test different groups of subjects and different activities. The implemented algorithm does not stipulate, *a priori*, any value for window length, or overlap percentage, but searches the best parameters that define the specific data. A clustering metric based on the construction of the data confusion matrix is also proposed. The results of clustering accuracy of human activities recognition with ARI in K-means application were very encouraging. An average person-dependent ARI of 99.29% and a person-independent ARI of 88.57% were reached. Compared to the subject-dependent case, the accuracy of all subjects for the Affinity propagation, the Mean Shift and the Spectral Clustering applications is much lower as it was expected from the previous results with K-means. The obtained results may be explained by the variations in human motion for different subjects.

Competing interests

The authors declare that they have no competing interests.

Author's contributions

IPM carried out the human activity recognition studies, participated in the algorithm implementation, performed the statistical analysis and drafted the manuscript. ALG participated in the algorithm implementation, statistical analysis and in the manuscript composition. HG conceived the study, and participated in its design and coordination and helped to draft the manuscript. VBP conceived the study, and participated in its design and coordination and helped to draft the manuscript. RMC participated in the design of the study and helped to draft the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank Plux- Wireless Biosignals who provided the biosensors for the accelerometry signals acquisition. RMC and VBP are supported by Human Brain Project and BrainFlight grants.

Availability and requirements

The algorithm described in the manuscript is available for testing by reviewers at:
https://www.dropbox.com/sh/3fkil2tws7z101g/J0pTQTmCXw

References

[1] A. Mannini, S. S. Intille, M. Rosenberger, A. M. Sabatini, W. Haskell, Activity recognition using a single accelerometer placed at the wrist or ankle., Medicine and science in sports and exercise.
[2] R. Gomes, N. Nunes, J. Sousa, H. Gamboa, Long term biosignals visualization and processing, in: BIOSIGNALS, 2012, pp. 402–405.
[3] G. V. Rossum, J. de Boer, Interactively testing remote servers using the python programming language 4 (1991) 283–304.

[4] T. Curk, J. Demsar, Q. Xu, G. Leban, U. Petrovic, I. Bratko, G. Shaulsky, B. Zupan, Microarray data mining with visual programming, Bioinformatics 21 (2005) 396–398.
[5] National institute of mental health. obsessive compulsive disorder (ocd)", Available at www.nimh.nih.gov/health/topics/obsessive-compulsive-disorder-ocd/index.shtml/.
[6] B. J. Simon, J. P. Errico, J. T. Raffle, Non-invasive magnetic or electrical nerve stimulation to treat or prevent autism spectrum disorders and other disorders of psychological development, uS Patent App. 13/783,319 (2013).
[7] M. R. Song, Y.-S. Lee, J.-D. Baek, M. Miller, Physical activity status in adults with depression in the national health and nutrition examination survey, 2005–2006, Public Health Nursing 29 (3) (2012) 208–217.
[8] L. Palmerini, S. Mellone, G. Avanzolini, F. Valzania, L. Chiari, Classification of early-mild subjects with parkinson's disease by using sensor-based measures of posture, gait, and transitions, in: AIME, 2013, pp. 176–180.
[9] C. H., Human activity representation, analysis and recognition, Department of Electrical Engineering , Indian Institute of Technology (May).
[10] Z. Ghahramani, Unsupervised learning, Gatsby Computational Neuroscience Unit, University College London.
[11] M. Mathie, Monitoring and interpreting human movement patterns using a triaxial accelerometer.
[12] A. Mannini, A. Sabatini, Machine learning methods for classifying human physical activities from on-body sensors. (2010).
[13] D. Crockford, The application/json media type for javascript object notation (json), RFC 4627, IETF (7 2006).
[14] J. Heer, N. Kong, M. Agrawala, Sizing the horizon: The effects of chart size and layering on the graphical perception of time series visualizations, in: ACM Human Factors in Computing Systems (CHI), 2009.
URL `http://vis.stanford.edu/papers/horizon`
[15] S. C. B. M. A. A. S. H. G. H. . F. A. Holzinger, A., On applying approximate entropy to ecg signals for knowledge discovery on the example of big sensor data, Berlin Heidelberg: Springer (2012) 646–657.
[16] K. Schindler, L. Van Gool, Action snippets: How many frames does human action recognition require?, in: Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, IEEE, 2008, pp. 1–8.
[17] L. Liao, D. Fox, H. Kautz, Location-based activity recognition using relational markov networks, in: Proceedings of the 19th international joint conference on Artificial intelligence, IJCAI'05, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2005, pp. 773–778.
URL `http://dl.acm.org/citation.cfm?id=1642293.1642417`
[18] E. M. Tapia, S. S. Intille, K. Larson, Activity recognition in the home using simple and ubiquitous sensors, in: In Pervasive, 2004, pp. 158–175.
[19] C. R. Wren, E. M. Tapia, Toward scalable activity recognition for sensor networks, in: In Lecture Notes in Computer Science, Springer, 2006, pp. 168–185.
[20] E. F. I. H. Witten, M. A. Hall, Data Mining: Practical Machine Learning Tools and Techniques, 2011.
[21] S. P. Lloyd, Least squares quantization in pcm, IEEE Transactions on Information Theory 28 (1982) 129–137.
[22] S. F. Walter, Clustering by affinity propagation, masters thesis, department of physics, eth zurich, switzerland.
[23] S. J. S. Boukir, K. Reinke, Fast mean-shift based classification of very high resolution images: application to forest cover mapping 1 (2012) 111–116.
[24] T. Liao, Clustering of time series dataa survey 38 (2005) 1857 – 1874.
[25] Motionplux, www.biosignalsplux.com/index.php/product.
URL `www.biosignalsplux.com/index.php/product`
[26] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, Journal of Machine Learning Research 12 (2011) 2825–2830.

Table 1. Information collected from each Feature.

Description	Brief description of the information extracted from the feature.
Imports	Necessary imports for the properly work of feature extraction.
Function	Function that allows the extraction of the respective feature.
Free Parameters	Particular inputs of the function.
Parameters	Inputs.
Number of Features	The function's number of outputs.
Use	If, for a given clustering iteration, the feature is used or not.
Reference	Code Reference.
Source	Code Source.

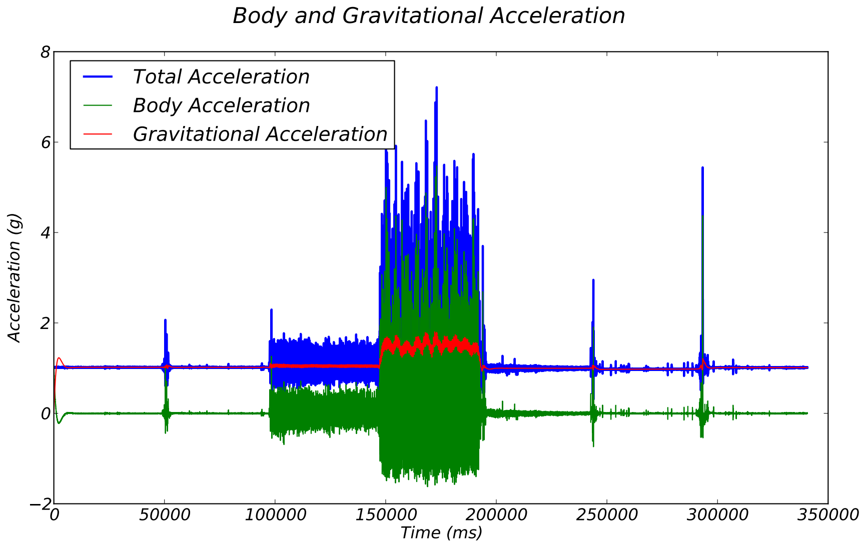


Figure 2. Body and Gravitational Acceleration of Accelerometer Sensor.

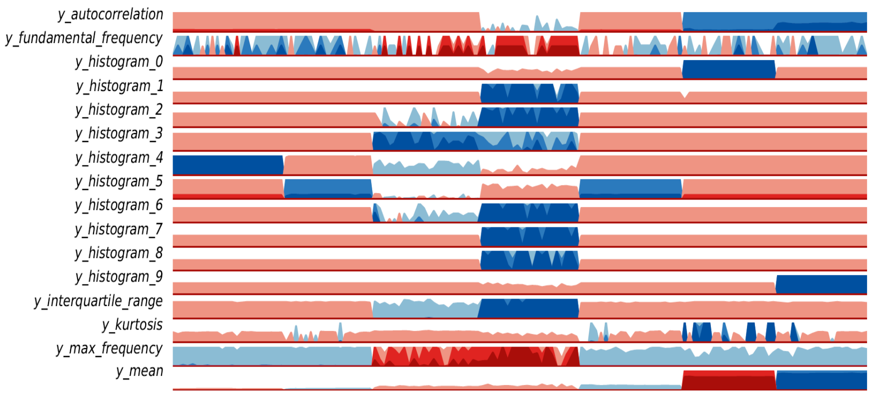
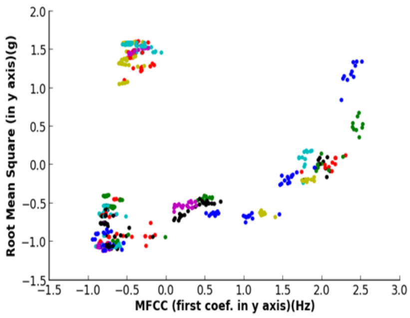
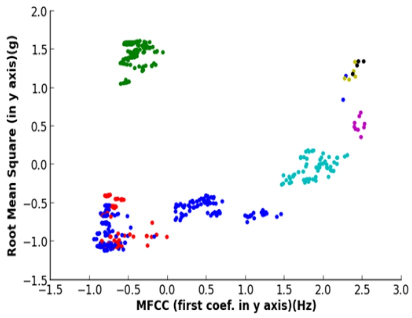


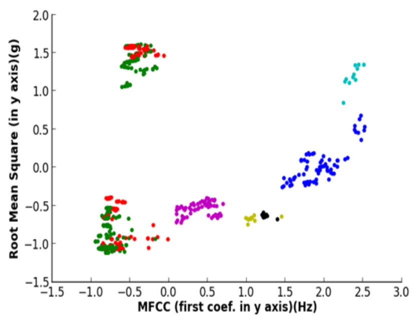
Figure 3. Horizon Graph: Time Series Visualization Technique.



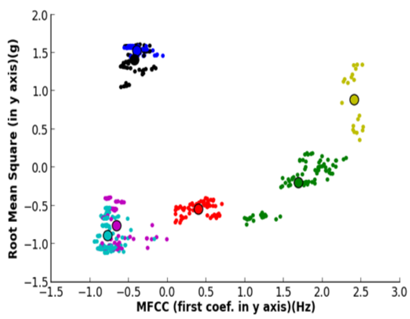
(a) RMS as a function of MFCC for Affinity Propagation.



(b) RMS as a function of MFCC for MeanShift.



(c) RMS as a function of MFCC for Spectral Clustering.



(d) RMS as a function of MFCC for K-means.

Figure 4. Structure of Clusters according to different domain features. Root Mean Square as a function of MFCC with Affinity Propagation (a), Mean Shift (b), Spectral Clustering(c) and K-means(d) applications with concatenate data for the inter-subjects case.

Table 2. Statistical, Temporal and Spectral Domain Features.

Statistical Domain	Kurtosis Skewness Mean Standard Deviation Interquartile Range Histogram Root Mean Square Median Absolute Deviation
Temporal Domain	Zero Crossing Rate Pairwise Correlation Autocorrelation
Spectral Domain	Maximum Frequency Median Frequency Cepstral Coefficients Power Spectrum Mel-Frequency Cepstral Coefficients Fundamental Frequency Power Bandwidth

Table 3. Different combinations of Free Parameters and Window Size of the Signal. First column describes the free parameters, the second and the third onde show the minimum and maximum values used for the studied combinations. The fourth column presents the number of features taken from each free parameters with the defined Minimum and Maximum.

Free Parameter	Minimum	Maximum	Number of Features
Windows Size (samples)	1000	4000	3
Bins of Histogram	10	20	3
Range of Histogram	1	3	2
Number of Cepstral C.	1	11	3
Number of MFCC	10	20	3
Power Bandwidth (samples)	10	20	3

Table 4. Recorded Movements.

Label of Movement	Type of Movement	Time of Movement
Label 0	Standing	about 60 seconds
Label 1	Sitting	about 60 seconds
Label 2	Walking	about 60 seconds
Label 3	Running	about 60 seconds
Label 4	Lying Down (belly up)	about 60 seconds
Label 5	Lying Down (right side down)	about 60 seconds
Label 6	Lying Down (left side down)	about 60 seconds

Table 5. Clustering and classification accuracy (mean value and standard deviation) as a function of different signals window length.

Window Size	Clustering Accuracy (%)	Classification Accuracy (%)
1000 samples	89.46% ± 0.4%	81.94% ± 0.7%
2000 samples	96.21% ± 0.9%	83.61% ± 0.6%
4000 samples	99.14% ± 0.6%	99.29% ± 0.5%