Inês Prata Machado

Bachelor of Science in Biomedical Engineering

Human Activity Data Discovery based on Accelerometry

Dissertation to Obtain Master Degree in
Biomedical Engineering

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

Co-Adviser : Dr. Vítor Paixão, Champalimaud Foundation, Champalimaud Neuroscience Programme

Jury:

Chairperson: Dr. Mário António Basto Forjaz Secca

Supervisor: Dr. Carla Maria Quintão Pereira

Members of the Committeeis: Dr. Hugo Filipe Silveira Gamboa
Dr. Vítor Manuel Bordona de Sousa Paixão

FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

October, 2013

Human Activity Data Discovery based on Accelerometry

To my parents and sisters

Acknowledgements

I would like to express my sincere gratitude to my supervisor Dr. Hugo Gamboa for the opportunity to work in the exciting field of human activity recognition, for the guidance, support and for the knowledge passed on which allowed me to achieve my goals and dreams. I am very thankful to Dr. Vítor Paixão, my co-supervisor from Champalimaud Foundation, whose support, dedication, motivation and enthusiasm helped me with no doubt during this research.

I would like to thank all of *PLUX* team members Neuza Nunes, Filipe Carvalho, Ricardo Gomes, Angela Pimentel, Nídia Batista, Lídia Fortunato and Tiago Araújo for creating a healthy environment where moments of hard work and pure entertainment were reconciled harmoniously. A special thanks to Neuza Nunes for her supervision over my work, for her dedication and patience while helping me during the development of my thesis. I would like to thank all Costa Lab team members from Champalimaud Foundation for all the support and help during this research.

Finally, also a special thanks goes to my colleagues Marília Barandas e Mafalda Câmara for all the support, laughs and lunches that we shared over these last months. To my friends Pedro Garcia, Jorge Ferreira, Pedro Moura, Emanuel Marques, Vanessa Cunha and Margarida Rebelo for making these last five years unforgettable.

Last but not least, I'm very thankful to my family. To my sisters Rita and Ana who have been by my side everyday, and helped me during my research with patience and dedication. A special thank to my parents for always trusting me and for giving me the opportunity to be here today allowing me to grow and evolve personally and professionally.

Abstract

The demand for objectivity in clinical diagnosis has been one of the greatest challenges in Biomedical Engineering. The study, development and implementation of solutions that may serve as ground truth in Physical Activity (PA) recognition and in medical diagnosis of chronic motor diseases is ever more imperative. This thesis describes a Human Activity Recognition (HAR) 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 was designed. In the present study, machine learning algorithms were applied in a non-supervised environment using feature representation of accelerometer data to discover the activities performed by different subjects. A feature selection framework is developed in order to improve the clustering accuracy and reduce computational costs. The features which best distinguish a particular set of activities were selected from a 180^{th} - dimensional feature vector through machine learning algorithms. The implemented framework achieved very encouraging results in human activity recognition: an average person-dependent Adjusted Rand Index (ARI) of $99.29\% \pm 0.5\%$ and a person-independent ARI of $88.57\% \pm 4.0\%$ were reached. Accurate and detailed measurement of an individual's PA is a key requirement for helping researchers understand the relationship between motor activity and health.

Keywords: Human Activity Recognition, Sensor Signal Processing, Feature Extraction, Feature Selection, Clustering Algorithms.

Resumo

A procura pela objectividade no diagnóstico clínico tem sido um dos grandes desafios na área da Engenharia Biomédica. O desenvolvimento e implementação de ferramentas que podem servir de *ground truth* no estudo do comportamento humano e no diagnóstico médico de doenças crónicas é cada vez mais imperativo. Este trabalho descreve uma ferramenta de reconhecimento de actividade humana com base na extracção e selecção de características de sinais de acelerometria. São usadas técnicas de aprendizagem automática baseadas num vector de características do domínio estatístico, temporal e espectral. As características que melhor distinguem um determinado conjunto de actividades foram seleccionadas através da aplicação de algoritmos de aprendizagem automática a um vector de 180 características. Foi desenvolvida uma ferramenta de selecção de características com o objectivo de optimizar os resultados obtidos e diminuir os custos computacionais. O algoritmo implementado apresenta resultados bastante promissores. Foi alcançada uma performance de $99.29\% \pm 0.5\%$ para testes dependentes do sujeito e $88.57\% \pm 4.0\%$ para testes independentes do sujeito. Uma avaliação precisa e detalhada da actividade física é um requisito fundamental para entender a relação entre comportamento e o respectivo estado de saúde do indivíduo.

Palavras-chave: Reconhecimento de Actividade Humana, Processamento de Sinal, Extracção de Características, Selecção de Características, Algoritmos de Clustering.

Contents

List of Figures

List of Tables

Acronyms

ACC Accelerometry

AAL Ambient Assisted Living

ASD Autism Spectrum Disorder

ARI Adjusted Rand Index

BA Body Acceleration

DBSCAN Density-based Spatial Clustering of Applications with Noise

FFT Fast Fourier Transform

GA Gravitational Acceleration

HAR Human Activity Recognition

HRV Heart Rate Variability

JSON JavaScript Object Notation

MEMS Microelectromechanical Systems

MFCC Mel-Frequency Cepstral Coefficients

OCD Obsessive Compulsive Disorder

PA Physical Activity

PD Parkinson's Disease

PSD Power Spectral Density

RI Rand Index

TA Triaxial Accelerometer

1

Introduction

1.1 Motivation

The constant concern with the human physical and psychological well-being has been the drive for research studies that have led to a promising evolution of medicine and engineering. In Biomedical Engineering, the demand for objectivity in clinical diagnosis has been one of the greatest challenges. The study, development and implementation of solutions that may serve as ground truth in the medical diagnosis of pathologies so subjective and hard to trace such as Obsessive Compulsive Disorder (OCD) or Autism Spectrum Disorder (ASD) is ever more imperative.

The change of motor activity is one of the essential signs of psychiatric disorders and many of these, including Depression and OCD, exhibit diagnostic criteria that require an assessment of the motor activity changes of the patient. The behavioural classification usually relies on observation, therefore a highly experienced analyst is always needed, which is the result of a too focused observation on small movements, causing difficulties in long term experiences. Human body movement has up to 244 degrees of freedom [1], making the modelling of structural and dynamic features for activity recognition of such a tough object, a complex task. Analysing human action is particularly challenging because of the complex non rigid and self occluding nature of the articulated human motion. There is a variety of methods to quantify levels of usual Physical Activity (PA) during daily life, including objective measurements, such as Heart Rate, Heart Rate Variability (HRV) and Accelerometry (ACC).

In the medical and therapeutic field, the accelerometer is used in the evaluation of human movement, detection of sleep disorders and fall detection, amongst other applications [2]. The aim of this research is to study and propose novel tools for ACC in human activity recognition. In the present study, the software OpenSignals [3] was used for signal acquisition and signal processing algorithms were developed in Python Programming Language [4] and Orange Software [5]. This dissertation was developed at PLUX - *Wireless Biosignals* in collaboration with the Champalimaud Foundation - Champalimaud Neuroscience Programme.

1.2 Objectives of the Current Work

The main focus of the present work is understanding signals produced by a Triaxial Accelerometer (TA), interpreting them in the context of human movement and identifying clinically relevant parameters from the data. Signal processing techniques are implemented with the purpose of examining accelerometer data and finding new information that would be difficult to identify directly from the raw data.

The hypothesis of the current work is that ACC is a suitable technique for monitoring movement patterns in free-living subjects over long periods of time and that it can be used to measure quantitative parameters that can provide clinical insight into the health status of the subject. A method for convenient monitoring of detailed ambulatory movements in daily life, using a portable measurement device employing a single TA is presented. The goal of this thesis is to extract information on statistical, temporal and spectral domains of ACC signals for human activity recognition.

Tests will be made, based on machine learning methodologies, which will allow the identification of different activities performed by different subjects. 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:

1. Acquire and analyse the data produced by a TA, placed on different parts of the body, during human movement: create a database and propose an annotation structure for PA data;

2. Develop a framework for the interpretation of the data provided by an ACC monitoring system: design a set of time, statistical and frequency domain features from several areas such as speech recognition and PA;

3. Develop algorithms to extract relevant information from the data;

4. Apply clustering algorithms based on feature representation of accelerometer data: visualization of time series features - The Horizon Plot [6];

5. Explore the choice of features and signal window size on the performance of different clustering algorithms;

6. Develop a framework to differentiate human activities, such as sitting, walking, standing, running and lying;

7. Evaluate the use of the system in daily life settings and the data interpretation framework: intra and inter subject context discovery;

8. Implement a new metric for assessing the obtained results from unsupervised techniques: classification-based evaluation.

1.3 State of the Art : Brief Historical Perspective

In recent decades, there has been an increasing interest in the use of ACC to monitor human behaviour. The advance of Microelectromechanical Systems (MEMS) has helped the development of small size and low cost accelerometers, making it a very convenient tool for monitoring free-living subjects. Accelerometers are inexpensive, require relatively low power [7], and are embedded in most of today's cellular phones. Many HAR systems have been developed in the past which incorporate the use of accelerometers [8] and nowadays, TA are, perhaps, the most broadly used sensors to recognize ambulatory activities [9].

Those physical activity recognition systems recognize a variety of everyday postures, physical and household activities and common exercise routines from a small set of wearable accelerometers using machine learning techniques. One of the key point is the diversity of areas where ACC has been used in the past. The most studied have been: physical activity recognition (defining and comparing group of subjects with different activity levels), balance and postural sway, metabolic energy expenditure (which is the standard reference for the measurement of physical activities), gait, detection of falls and sit-to-stand transfers (which is an important indicator for postural instability). The use of accelerometers has helped on diagnose and prevention of a number of diseases such as Parkinson's Disease (PD), OCD , ASD and Depression [10, 11, 12, 13, 14].

Obsessive Compulsive Disorder is one of the most debilitating neuropsychiatric disorders, characterized by unreasonable thoughts (obsessions) that lead to repetitive behaviours (compulsions) [10]. The diagnosis of OCD is mostly based on descriptions of the experiences lived by the patient, in the behaviours described by relatives and by evaluating the patient's mental state. The change of motor activity is one of the essential signs of psychiatric disorders and many of these, including Depression and OCD, exhibit diagnostic criteria that require an assessment of the motor activity changes of the patient.

Parkinson's Disease was initially characterized by James Parkinson in 1817. PD is classified as a chronic movement disorder. Predominant attributes of PD are rigidity from an increase in muscle tone, shuffling gait, impaired balance and tremor during resting status [11]. Accelerometers have been tested and evaluated for the characterization of PD status, and the integration of accelerometer technology has the potential to further advance treatment of PD [15].

Autism Spectrum Disorder is a disorder affecting a child's social development characterized by deficits in social skills, communication and repetitive or restricted interests. Children with autism will often exhibit behaviours, such as vocal stutters and brief bouts of vigorous activity (e.g., violently striking the back of the hands) to cope with everyday life. Over the past few years, research in both physical activity and autism with ACC has been rigorous, leading to more advanced methods of assessment and sampling [12]. Recognizing this behaviours via wearable sensors can provide valuable information regarding an individual's degree of functional ability.

Depression is a major public health problem and is characterized by one or two important depressive episodes, such as lowered mood, increased sense of worthlessness, fatigue, and preoccupation with death and suicide [13]. By using accelerometers, the effects of psychomotor retardation on the level of gross body movement can be measured. Accelerometers can also be used to identify regions of physical activity, sleep, and sedentary behavioural states throughout the day [14].

One type of system designed for elderly people aims to detect potentially dangerous situations in a person's life in order to call for external help automatically. Such systems can be seen as a complement to traditional emergency systems such as smoke or fire alarms, by detecting, for instance, when a person has fallen [16] or when vital body signs indicate imminent health threats [17]. Other type of health-related system aims to use context-information to promote a more active and thus healthy lifestyle, or to actively support elderly or disabled people in performing everyday activities.

Physical-activity recognition via wearable sensors can provide valuable information regarding an individual's degree of functional ability and lifestyle. Even in the case of healthy people, these movements can be highly individualistic. Nonetheless, there are a set of basic parameters that are common to all instances of these movements. This allows, through ACC, the identification and classification of certain behaviours of daily life. To study daily life activities, accelerometers need to be able to detect accelerations between $\pm 12\,g$, in general, and greater than $\pm 6\,g$ if placed at the waist [18].

Regarding the used classification systems, two approaches have been used to pattern recognition. The first approach uses fixed-threshold classification while the second approach uses reference-pattern-based classification. In fixed-threshold classification, activities and postural orientations are discriminated by applying a threshold (derived empirically) to the accelerometer signal. In reference-pattern-based classification, activity patterns are compared to template reference patterns [19]. The traditional approach in HAR for classifying a time series into physical activity, is to use regression-based thresholds called cut-points [20] which allow researchers to estimate the time spent performing physical activities at different intensity levels. Researchers, however, have found cut-points to be inaccurate, and are turning to machine learning methods to identify physical activity types and estimate energy expenditure more accurately [21].

Several papers have reported high recognition accuracy (92.25% [9], 95% [22], 97% [23], and up to 98% [24]) under different evaluation methodologies. This work aims to go beyond the state of the art in HAR, presenting solutions that address some of current limitations: implement a novel set of features, test different signal segmentation methods, selecting different window sizes and overlap percentage and propose a novel metric for unsupervised methodologies.

The main objective of this study is to present a method for convenient monitoring of detailed ambulatory movements in daily life, by the use of a portable measurement device employing a single TA and machine learning techniques.

1.4 Thesis Overview

The structure of this thesis is schematically represented in Figure 1.1.
In the first two chapters the basis that support the present research is reported. Chapter 1 provides a brief historical perspective, where objectives and motivation of this research are clarified. The importance of objective monitoring human movement is discussed. Also in this chapter, an overview on other studies about HAR with wearable sensors is presented. In Chapter 2, different activities, sensors, and machine learning approaches that have been proposed are discussed. Techniques for the assessment of human movement are reviewed and the choice of ACC for unsupervised home monitoring is established. The impact of the sensor specifications has also been analysed.

Chapter 3 presents a set of time, statistical and spectral domain features used in this investigation. The choice of features is a fundamental first step in applying machine learning methodologies to sensor data, and it can have a strong influence on the outcome of any approach.

In Chapter 4, the composition of the TA signal is explained. The signal is made up of several components, and each of these is examined. The proposed methodology and the approach used in the work to extract and select features based on motion data is described.

In Chapter 5, the carried out tests to pick the features which best distinguish a particular set of activities are described. The experimental evaluation of the ACC data collection and interpretation system is discussed.

In Chapter 6, the results obtained from the investigation are presented. This chapter draw the main conclusions from the work and provides recommendations for further research.

The Appendix contains the paper accepted in the context of the presented research work.

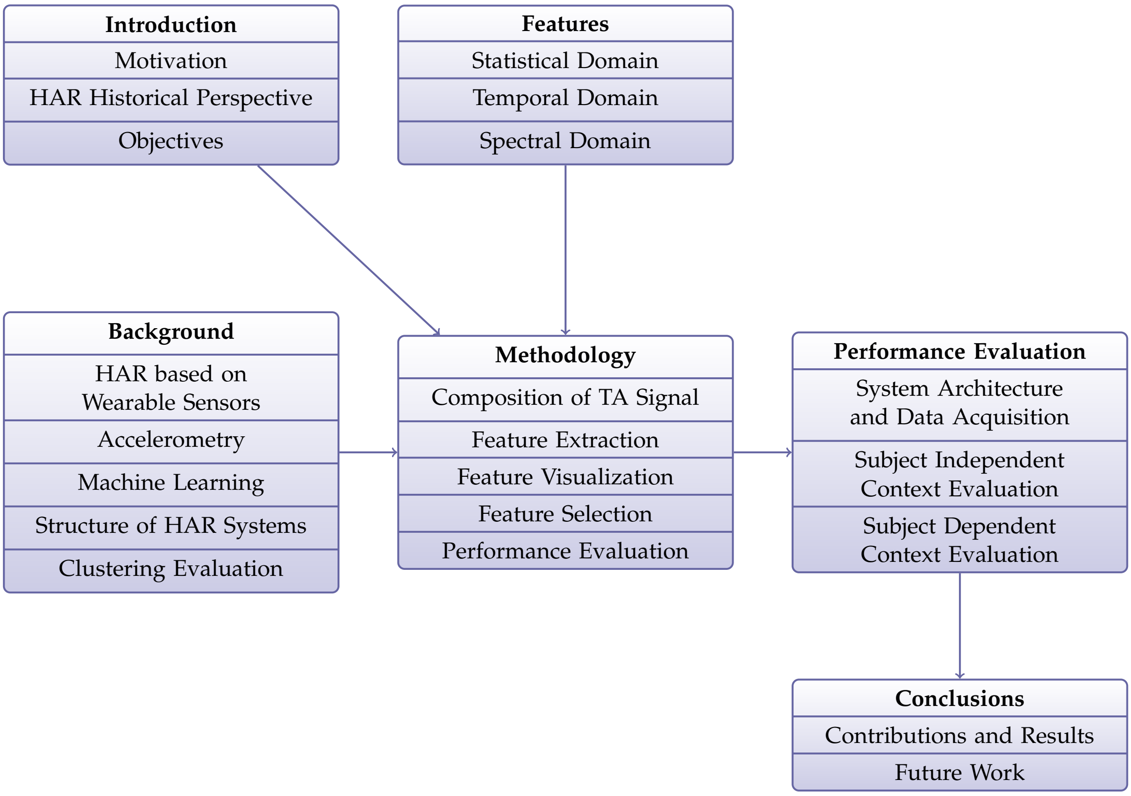


Figure 1.1: Thesis Overview.

2

Theoretical Background

Monitoring human movement can provide valuable information on a patient life style, health status, rate of rehabilitation and other potentially useful clinical data. The present work addresses these challenges in the context of wearable accelerometer-based simple activity recognition. ACC is a method of movement kinematic analysis which allows, through the use of an accelerometer, the quantification of caused or suffered accelerations of the human body [25]. It has been pointed out that ACC, a technique that is increasingly being used for monitoring human movement in laboratories and research studies, is suitable for long term monitoring of human movement [21].

2.1 Activity Recognition based on Wearable Sensor

In signal acquisition, the sensor is the part of the instrument sensible to variations of the physical parameter to be measured and must be specific to the nature of the signal to be acquired [25]. Besides the fact that accelerometers usually lead to good results in recognition of physical activities, they are small and cheap, require relatively little energy, memory and processing power [26]. The basic principle of operation behind the accelerometer based on MEMS is the displacement of a small proof mass etched into the silicon surface of the integrated circuit and suspended by small beams [27]. Consistent with Newton's second law of motion, when a force is applied to the device, a developed acceleration displaces the mass. Acceleration can be defined as the rate of change of direction or magnitude in the velocity of an object [25]. Therefore, its units are ms^{-2} or g units, where $1g = 9.81ms^{-2}$. When choosing an accelerometer, it should be considered:

- **Dynamic Range:** the $\pm$ maximum amplitude that the accelerometer can measure before distorting or clipping the output signal. Dynamic range is typically specified units in g's.
- **Frequency Response:** the frequency range to which the sensor will detect motion and report a true output. Frequency response is typically specified as a range measured in Hertz.
- **Sensitive Axis:** accelerometers are designed to detect inputs in reference to an axis. Single-axis accelerometers can only detect inputs along one plane. TA can detect inputs in three orthogonal plans and are suitable for most applications.
- **Size and Mass of an Accelerometer:** these parameters can change the characteristics of the tested object. The mass of the accelerometers should be significantly smaller than the mass of the system to be monitored [25].

The most common accelerometers used in human activity research measure accelerations either in the vertical plane (uni-axial), or in three planes (triaxial) and respond both to frequency and intensity of the movement [28]. As the majority of human motion occurs in more than one movement axis, TA are used to measure the acceleration in each orthogonal axis. Table 2.1 presents some accelerometer specifications used in the present investigation.

Table 2.1: Target Accelerometer System Parameters.

Parameter	**Target Value**		
Number of axes	3		
Sampling Frequency	1.25 to 800 Hz		
		12 bits	8 bits
Maximum Acceleration Amplitude and Acceleration Resolution (in bits and g)	$\pm 2\ g$	2/2048	2/128
	$\pm 4\ g$	4/2048	4/128
	$\pm 8\ g$	8/2048	8/128
Maximum Acceleration without Damage	5000 g		

The reference point is usually chosen so that 0 g corresponds to a free-fall condition and that the maximum output number corresponds to the maximum amount of g that the device can register. The output of an accelerometer worn on the body is dependent on four factors: the position at which it is placed, its orientation at this location, the posture of the subject and the activity being performed by the subject [29]. If the subject is at rest, the output of the accelerometer is determined by its inclination relative to the gravitational vector. If the orientation of the accelerometer relative to the person is known, then the resulting accelerometer recordings can be used to determine the posture of the subject relative to each direction [29].

2.2 Number and Body Position of the Accelerometer

The placement of the accelerometer is a relevant point of discussion. A device that is to be worn over extended periods must be designed to be as simple to put on and comfortable to wear in order to encourage patient compliance [30]. A system with multiple sensors placed across the body can provide superior data to a system that has only a single sensor location. However, such system will be more time-consuming and thus, more inconvenient to put on and wear, which will lead to reduced compliance rates.

One of the key points given is that in order to measure human acceleration, it is important to understand the motion of the human body and realize which physical property one wishes to measure [31]. This is necessary in order to choose the right combination of measurement range and accelerometer placement.

Generally, body motion can be measured with a single accelerometer placed close to the body's center of mass, which is located within the pelvis [32]. The advantage of this placement is that attachment at the waist allows monitoring of accelerations near the center of mass. Any movement of the body will cause the center of mass to shift [33].

There are other common placement locations such as chest or thigh [19]. Normally, accelerometers are attached to the part of the body whose movement is being studied. Ultimately, the optimal position to place the accelerometer depends on the application and the type of activities to be recognized. The present study aims to develop a HAR framework, for a waist mounted accelerometer based system.

2.3 Human Activities Acceleration

The magnitude of the acceleration tends to increase from the head to the ankle, and is generally greater in the vertical direction [18]. Just as magnitude, frequency tends to decrease from the ankle to the head, and is greater in the vertical direction than in the transverse plane [18].

The behaviour of the recognition accuracy as a function of the accelerometer sampling rate was studied by Maurer *et al.* (2006) [34]. It was shown that no significant gain in accuracy is achieved above 20 Hz for ambulation activities. Table 2.2, adapted from [18], shows some amplitude intervals for some typical activities, all units are in g.

Table 2.2: Physical activities: Amplitude of the Movement with different Sensor Locations.

Motion	Vertical (g)			Horizontal (g)		
	Head	**Body**	**Ankle**	**Head**	**Body**	**Ankle**
Walking	-	-0.3; 0.8	-1.7; 3.3	-0.2; 0.2	-0.3; 0.4	-2.1; 0.4
Running	0.8; 4.0	0.9; 5.0	3.0; 12.0	-	-	-

Since the measurement ranges have an impact on the precision and cost of the accelerometer, and the frequency determines the sampling rate of the device, this knowledge becomes important when designing a sensor for clinical assessment.

2.4 Machine Learning Techniques for Activity Recognition

An efficient approach based on machine learning methods has been recently proposed in several research projects with focus on activity recognition. Machine learning algorithms based on the feature representation of accelerometer data have become the most widely used approaches in PA prediction [35]. There are two main types of machine learning algorithms: unsupervised and supervised learning. In this thesis, unsupervised learning was chosen because of it collection of methods for grouping unlabelled data into subsets (called clusters) that are believed to reflect the underlying structure of the data, based on similarity groups within the data.

2.4.1 The K-Means Clustering Algorithm

K-means is a commonly used partitioning based clustering technique that searches for a specified number of clusters, which are represented by their centroids, by minimizing an error function [36]. This algorithm takes as input the number of clusters to generate, C, and a set of observation vectors to cluster and returns a set of centroids, one for each of the C clusters. 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. The mean (centroid) of each cluster is then computed so as to update the cluster center. This update occurs as a result of the change in the membership of each cluster. The processes of re-assigning the input vectors and the update of the cluster centers is repeated until no more change in the value of any of the cluster centers. An observation vector is classified with the centroid index of the centroid closest to it.

The k-means algorithm tries to minimize distortion, which is defined as the sum of the squared distances between each observation vector and its dominating centroid [37]. Each step of the k-means algorithm refines the choices of centroids to reduce distortion. The change in distortion is used as a stopping criterion: when the change is lower than a threshold, the k-means algorithm is not making sufficient progress and terminates. One can also define a maximum number of iterations. If there is a set of observations vector x_j with $j = 1, \ldots, N$, to be organized into C_i partitions with $i = 1, \ldots, k$, then the squared error criterion is defined by Equation 2.1:

$$J(M) = \sum_{i=1}^{k} \sum_{j=1}^{N} \|x_j - m_i\|^2 \tag{2.1}$$

where m_i is an element of the cluster prototype or centroid matrix M [38].

2.4.2 Clustering Distance Metric

A good clustering test will produce clusters in which the intra-class similarity is high and the inter-class similarity is low. The K-Means Clustering Algorithm [39] gives a single set of clusters, with no particular organization or structure within them. An important component of a clustering algorithm is the distance measured between data points. If the components of the data, for instance vectors, are all in the same physical units then it is possible that the simple Euclidean distance metric is enough to successfully group similar data instances. Another type of distance measurement that can be used is Hamming, Manhattan or Pearson Correlation [39].

2.5 General Structure for Human Activity Recognition Systems

In the literature, there are many different methods to extract activity information from raw sensor data [40]. The main steps can be categorized as: Preprocessing, Segmentation, Feature Extraction, Dimensionality Reduction and lastly, Clustering or Classification Process. Figure 2.5 represents the general structure for HAR systems.

Preprocessing: Initial samples received from any type of sensor are called raw data. Accelerometers respond to gravitational and body acceleration and the aim of filtering the signal was to approximately separate the Body Acceleration Component and the Gravitational Acceleration Component [41]. 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.

Segmentation: Different segmentation methods can be applied to time-series data which enhance relevant signal properties and enable the gather of useful information from continuous stream of data: in the present study, timing windows and sliding windows were considered as segmentation methods [40]. For activity recognition, where accelerometer data from physical activity is windowed, the choice of the number of frames is guided by a trade-off between two aspects: information and resolution. In the present study, the accelerometer data was collected, cleaned, and preprocessed to extract features that characterize 1000, 2000 and 4000 samples data windows with different overlap percentages.

Feature Extraction: Features can be defined as the abstractions of raw data since they are reduced sets of original raw data which basically represent main characteristics and behaviours of the signal. The reduced subset of large input data can be called as a feature vector, it contains important hints for the activity to be recognized and it is the main input for clustering algorithms [40]. In the present investigation, features are grouped as time, frequency and statistical domains.

Dimensionality Reduction: The aim of dimensionality reduction is to reduce the computational complexity and increase the performance of the activity recognition process. After the previous steps, collected data can be used directly in the clustering step. But some part of data may not even contribute to the results of the clustering process. Therefore, feature selection chooses distinguishing features from a set of candidates and feature extraction uses data transformations to generate useful and novel features.

Clustering Algorithms: Clustering mechanisms separate and organize unlabeled data into different groups whose members are similar to each other in some metric [42]. These groups are called clusters. Being a method of unsupervised learning, the learner only receives unlabeled inputs with no class information. A good clustering test will produce clusters in which the intra-class similarity is high and the inter-class similarity is low. The ultimate goal of clustering is to provide meaningful insights from the original data. In the present work, K-Means Clustering Algorithm [39] was used.

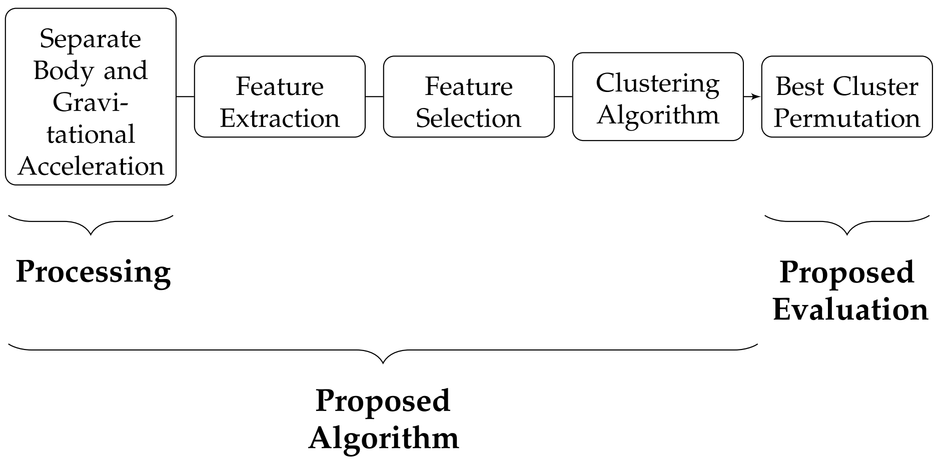


Figure 2.1: General Structure for Human Activity Recognition Systems.

2.6 Clustering Performance Evaluation

Ideally, after a feature space clustering procedure, each cluster should contain samples of only one activity. This would indicate that the data of the given feature was clearly separable and thus, well-suited as an input for classification.

In the worst case, the clustering performance is equal to the probability of the performed activity ($\sim n^{-1}$ where n is the number of activities). This would imply that the feature was not discriminative for the given set of activities and thus unlikely to be suited for recognition [43]. The clustering performance should be higher than random guessing.

Evaluating the performance of a clustering algorithm is not as trivial as counting the number of errors or the precision and recall of a supervised classification algorithm. In particular, any evaluation metric should not take the absolute values of the cluster labels into account, but rather if this clustering define separations of the data similar to some ground truth set of classes or satisfying some assumption such that members belong to the same class are more similar that members of different classes according to some similarity metric.

2.6.1 Adjusted Rand Index as a Metric for Comparing Partitions

Given the knowledge of the ground truth class assignment (labels true) and predicted labels from the clustering algorithm, the ARI is a function that measures the similarity of the two assignments, ignoring permutations and with chance normalization [39]. If C is a ground truth class assignment and k the clustering labels, x and y are defined as:

- x, the number of pairs of elements that are in the same set in C and in the same set in k.
- y, the number of pairs of elements that are in different sets in C and in different sets in k.

The raw (unadjusted) Rand index is then given by Equation 2.2:

$$RI = \frac{x+y}{C_2^{n_{samples}}} \tag{2.2}$$

Where $C_2^{n_{samples}}$ is the total number of possible pairs in the dataset (without ordering). However, the Rand Index (RI) score does not guarantee that random label assignments will get a value close to zero (mainly if the number of clusters is in the same order of magnitude as the number of samples) [39]. To counter this effect the expected of random labels can be discounted by defining the adjusted Rand index shown in Equation 2.3:

$$ARI = \frac{RI - E[RI]}{max(RI) - E[RI]} \tag{2.3}$$

2.7 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 [42]. 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, the **Algorithm 1**, that links these groups to their corresponded activity, was implemented.

Algorithm 1 Best Cluster Permutation.

```
dim ← matrix.shape[0]                          ▷ Compute matrix dimension
p ← matrix.argmax(axis = 1)                    ▷ Find maximum value index for each row
up ← unique(p)                                 ▷ Sort unique index of p
if len(up) == dim then
    return p                                   ▷ Check if there is only a maximum value per row
else
    newP ← zeros(dim) − 1.0
    for i in range(len(up)) do
        ind ← find(up[i] == p)
        if len(ind) == 1 then
            newP[ind] ← up[i]
        else
            bInd ← argmax(amax(m[ind, :], 1), 0)   ▷ Find the maximum index and assign it
            newP[ind[bInd]] ← up[i]
        end if
    end for
    ind ← find(newP == −1)                     ▷ Check labels not assigned
    miss ← range(dim) − set(up)
    tempP ← bcp(matrix[ind, :][:, missing])    ▷ Matrix is recursively built
    newP[ind] ← miss[tempP]                    ▷ Returns vector with true assignments
    newCm ← zeros((dim, dim))
    for i in arange(len(newP)) do
        newCm[i, :] ← matrix[newP[i], :]
    end for
    return newCm                               ▷ Returns Confusion Matrix
end if
```

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.

3

Statistical, Temporal and Spectral Domain Features

The choice of features is a fundamental first step in applying machine learning methodologies to sensor data, and it can have a strong influence on the outcome of any approach. In the present work, each activity is model based on timing and sliding window strategies. Specifically, the continuous sensor streams are divided into fixed length windows. By choosing a proper window length, all the information of each activity can be extracted from each single window. The information is then transformed into a feature vector by computing a 180^{th} - dimensional feature vector over the sensor data within each window. Three sets of features which are incorporated in the recognition investigation framework are now described. Figure 3 summarizes the list of features considered in this work.

3.1 Statistical Domain Features

A fundamental task in many statistical analysis is to characterize the location and variability of the time series. A further characterization of the data includes skewness, kurtosis and histogram. Histogram is an effective graphical technique for showing both the skewness and kurtosis of a data set. The histogram of Body Acceleration (BA) component

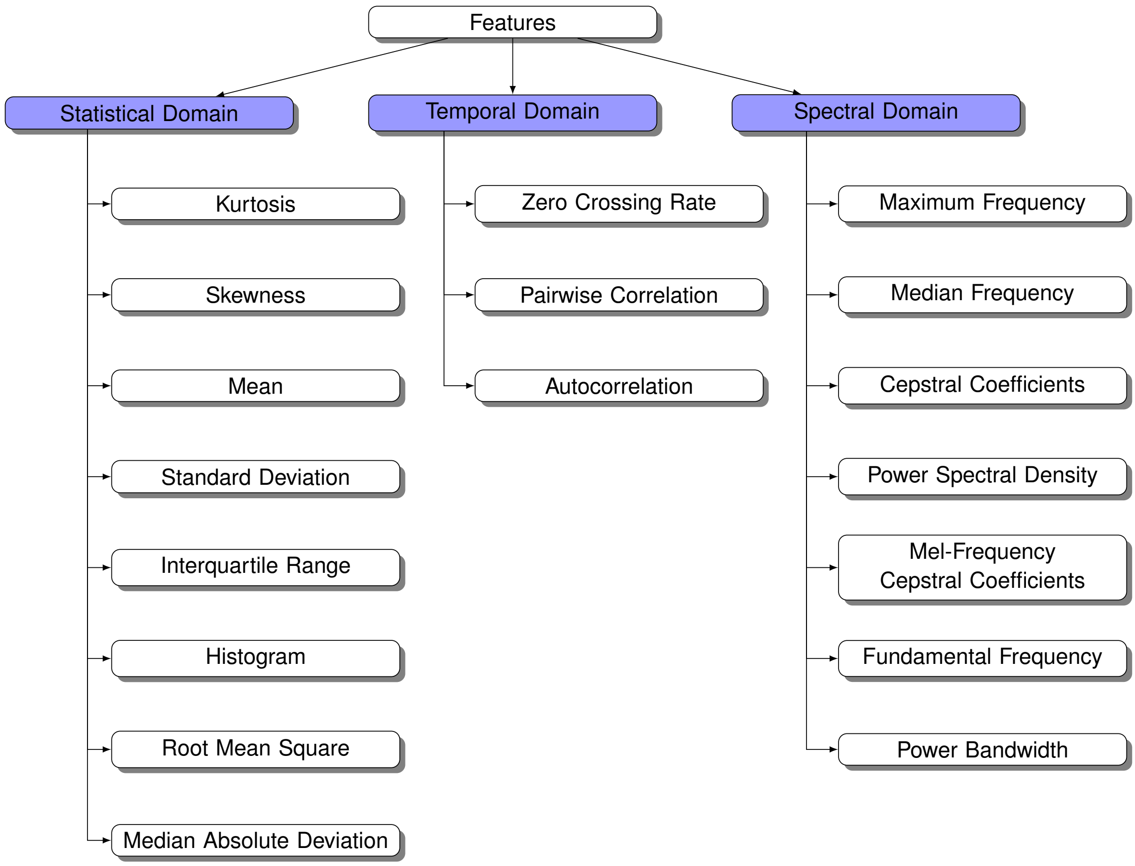


Figure 3.1: Statistical, Temporal and Spectral Domain Features.

can be used to evaluate that variable as a predictor of static and dynamic activities. Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution [44]. Skewness is a measure of the degree of asymmetry of the sensor signal distribution [44]. This function receives as inputs a n-dimensional array with data and the axis along which skewness is calculated. Figure 3.2 represents a typical recording from the accelerometer showing seven minutes of motion data, where the subject is asked to perform specific tasks. Time, in milliseconds, is represented in the x-axis and the magnitude of the acceleration is represented, in g, in the y-axis. Histogram is a graphical representation of the distribution of data, Figure 3.3. This function groups data into bins, plotting the number of members in each bin against the bin number. It receives the number of bins and defines the number of equal-width bins in the given range (the lower and upper range of the bins).

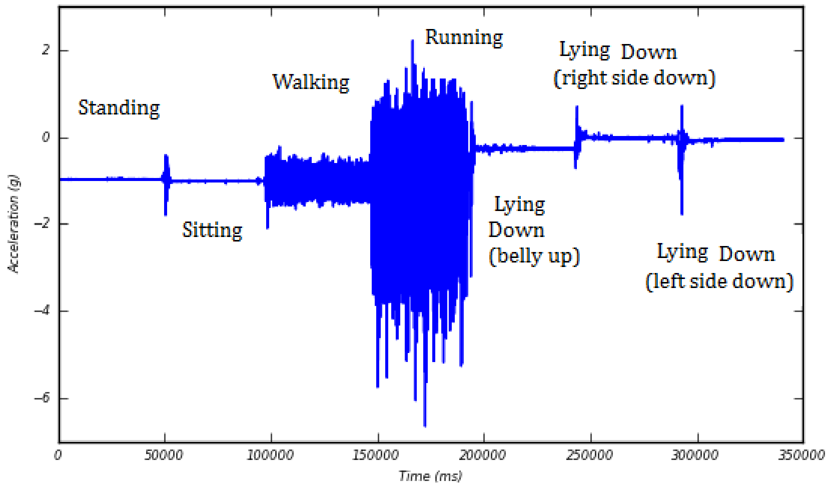


Figure 3.2: Typical recording from the accelerometer showing seven minutes of motion data where the subject is asked to perform specific tasks.

The result of a normalized histogram is the value of the probability density function at the bin, normalized such that the integral over the range is 1 [44]. An array of different number of bins is calculated for each axis, representing the distribution of acceleration values. The amplitude, number of members in each bin, is considered as a new feature to add to the feature vector.

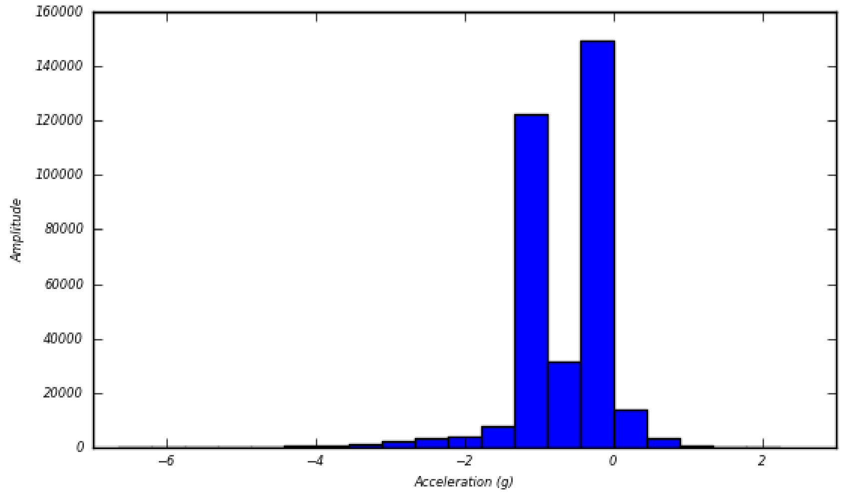


Figure 3.3: Histogram representation of a typical recording from the accelerometer showing seven minutes of recorded data where the subject is asked to perform specific tasks.

Mean is the DC component (average value) of the signal over the window. There have been various early uses of the mean metric in activity recognition [45, 46]. Several researchers have used the mean to identify user posture (sitting, standing or lying) and also to discriminate the type of activity as either dynamic or static [47]. The Standard Deviation can give an indication of the stability of a signal, measuring the variability of the signal over the window. The standard deviation of the acceleration signal is useful in capturing the range of possible acceleration values to separate activities that may look similar in nature but different in their speed and acceleration (e.g. walking vs. running). In the past, Median Absolute Deviation was used in the automated detection of epileptic seizures [48].

The Root Mean Square is the quadratic mean value of the signal over the window. In the past, root mean square has been used to classify wavelet results by distinguishing walking patterns [49] and is present in works of activity recognition [50]. Interquartile Range represents a measure of the statistical dispersion, being equal to the difference between the 75^{th} and the 25^{th} percentiles of the signal over the window. When the mean values of different classes are similar, the interquartile range represents the dispersion of the data and avoids the effect on range caused by extreme values in the data. As Median Absolute Deviation, this feature was used in the automated detection of epileptic seizures and in fall detection studies [48].

3.2 Temporal Domain Features

This section highlights the common uses of temporal-domain analysis for the recognition of user activity from accelerometer data. Zero-crossing can be defined as the total number of times that the signal changes from positive to negative or vice versa, normalized by the window length. A non-zero threshold can be defined, which can be an extracted mean value. Zero Crossing Rate is commonly applied to audio signals to identify the surrounding environment or the type of sound such as music, speech, and noise [51]. In this work, zero crossing rate is used in human activity recognition perspective.

Signal correlation is used to measure the strength and direction of a linear relationship between two signals. In activity recognition, correlation is especially useful in differentiating between activities that involve translation in a single dimension [52]. In order to calculate the degree of correlation it is necessary to calculate the correlation coefficients between the signals for the various axes. The most commonly used is Pearson's coefficient [53], $\rho_{x,y}$, also known as the sample correlation coefficient, calculated as the ratio of the covariance of the signals along the x-axis and the y-axis to the product of their standard deviations, as shown in Equation 3.1:

$$\rho_{x,y} = \frac{cov(x,y)}{\sigma_x \sigma_y} \tag{3.1}$$

The sample correlation coefficient was applied in [52], in order to determine which are the best classifiers (or combination of them) for recognizing activities, and which among several features are the most relevant. The correlation among accelerometer axes is useful in distinguishing activities that may appear similar but are performed in different dimensions. Autocorrelation is a commonly used feature of activity recognition to apply in a sliding window algorithm in order to measure the self-similarity of time series segments [54].

3.3 Frequency Domain Features

Understanding the difference between the resulting vector of a Fourier transform and the vector of accelerometer readings is crucial to understand the Fourier transform itself. In order to derive frequency-domain features, the window of sensor data must first be transformed into the frequency domain, normally using a Fast Fourier Transform (FFT). The vector produced by the Fourier transform represents the distribution of values over a range of frequencies. The output of a FFT gives the basis coefficients which represent the amplitudes and phases of the frequency components of the signal and the distribution of the signal energy. This section highlights the common uses of frequency-domain analysis for the recognition of user activity from accelerometer data. Maximum and Median Frequency are computed. The inputs of maximum frequency are the motion data and the sampling frequency of the signal's acquisition, **Algorithm 2**.

Algorithm 2: `Maximum Frequency`

```
Input:  motion data and sampling frequency.
Output:  maximum frequency point.
```

This algorithm computes the FFT of the input signal and, from the obtained frequency signal distribution, finds the frequency point where the FFT reaches its 95% of distribution. This value is considered the maximum frequency point. The pseudo code to compute maximum frequency is shown below:

Algorithm 2 Maximum Frequency

1: $f, fs \leftarrow fft(signal, samplingfrequency)$ ▷ Compute Fast Fourier Transform
2: $cfs \leftarrow cumsum(fs)$ ▷ Cumulative Sum of the FFT elements
3: $mag\_index \leftarrow find(cfs > cfs[-1] * 0.95)[0]$ ▷ Find index
4: $max\_freq \leftarrow f[mag\_index]$ ▷ Returns maximum frequency

Regarding median frequency, the inputs of this function are the motion data and the sampling frequency. Like the previous algorithm, this one computes the FFT of the input signal and finds the frequency point in which the FFT reaches its 50% of distribution. This value is, by definition, the median frequency point.

Usage of Power Spectral Density (PSD) functions for extraction of features is a standard approach in various areas of pattern recognition, including activity recognition, acoustics and imaging [55].

Algorithm 3: `Power Spectral Density`

`Input:  ACC signal and sampling frequency.`

`Output:  maximum power and respective frequency peak.`

First, this algorithm computes the PSD of the given signal. The PSD describes how the power of a time series is distributed with frequency. This function returns an array of frequencies ranging from 0 Hz to sampling frequency, and the power corresponding to each frequency, **Algorithm 3**. With this information, the algorithm finds the maximum peak power and returns its value and the frequency where it occurs. PSD reduces the redundancy in signals by concentrating energy into smaller areas of the frequency domain. In the present work, a straightforward approach is to use magnitude averages of PSD over a few frequency intervals. The inputs of this function are the motion data and sampling frequency of the signal. The pseudo code to PSD is shown below:

Algorithm 3 Power Spectrum Density

1: $power, freq \leftarrow psd(sig/std(sig), FS)$ ▷ Compute Power Spectrum Density
2: $maxPower \leftarrow max(power)$ ▷ Find the maximum value of PSD
3: $peakFreq \leftarrow freq[argmax(power)]$ ▷ Return the respective frequency

The cepstrum is defined as the inverse Fourier transform on the log-magnitude Fourier spectrum [56]. The coefficients that make up the resulting cepstrum are known as the Cepstral Coefficients. In the past, cepstral coefficients have only been used for the identification of echoes that are present in an acoustic signal [56] but later, cepstral coefficients have been shown to be a feasible set of features for speaker identification and even for musical instrument identification [57]. In the present work, cepstral coefficients are used in human activity recognition perspective.

Stevens et al. (1937) proposed the mel scale, which is a scale of pitches judged to be equal in distance from one another according to human perception [58]. Mel-frequency cepstrum is mapped onto the mel scale before the log and inverse fourier transform is taken. As such, the scaling in mel frequency cepstrum mimics the human perception of distance frequency and its coefficients are know as the Mel-Frequency Cepstral Coefficients (MFCC). MFCC are now widely used in speaker recognition tasks [59] and has been shown to yield excellent results [60], [61]. In [60], it is also shown that MFCC outperforms normal cepstral coefficients for speaker recognition. MFCC are used today in voice recognition areas - audio finger printing [62]. For the first time, these coefficients appear in physical activity recognition and have revealed to be quite promising due to the achieved performances.

The fundamental frequency, f_0, of a periodic or quasi-periodic signal is the inverse of the repeating period pattern length, which is the longer repeating unit of a signal. Considering the signal as a superposition of sinusoids, the f_0 is the lowest frequency harmonic [63]. Regarding activity recognition, for example, as cycling involves a uniform movement of the legs, a frequency-domain analysis of thigh acceleration shows a single dominant frequency. In contrast, running or walking may result in more complex acceleration patterns and often displays many major FFT components.

4

A Framework for Activity Recognition

Most approaches to activity recognition, using body-worn accelerometers, involve a multi-stage process. 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. Before using the accelerometer system in any monitoring context and before the development of algorithms to interpret data recorded by the system, it is necessary to understand the nature of the signals produced by the TA unit. The signal is made up of several components and each component is examined. The difficulties in distinguishing between the different signal components are discussed.

4.1 Composition of Triaxial Accelerometer Signal

This section describes the composition of the TA signal. According to [41], the signal measured by each fixed-body accelerometer is a linear sum of, approximately, three components:

- 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 - BA Component - is important, however these two components have overlapping frequency spectra.

According to [41], the BA component ranges from above 0 Hz to possibly up 20 Hz, but it is mostly contained in the range above 0 and below 3 Hz. This range overlaps the area covered by the GA component, which goes from 0 to several Hertz. This makes the identification of the signal parts which correspond to the BA or to the GA components, difficult. It is possible to approximately separate the BA and the GA components with some filtering. In [41], a wide range of different filters types with different characteristics and different windowing percentages were tested, in order to determine their ability to differentiate the components of the acceleration signal.

4.1.1 Body and Gravitational Acceleration Components

Accelerometers respond to Gravitational and Body Acceleration. The GA is also referred to the static component, while acceleration due to body movement is referred to the dynamic component. According to [41], the BA component depends on three factors:

- The nature of the activity being undertaken;
- The location on the body at which the acceleration is measured;
- The orientation of the accelerometer relative to the body;

Considering that x, y, and z are the outputs of the TA along the x, y and z-axes respectively, and that the GA component along the x-axis is represented by x_{GA}, and similarly for the y and z-axes, then:

$$x = x_{GA} + x_{BA} \tag{4.1}$$

$$y = y_{GA} + y_{BA} \tag{4.2}$$

$$z = z_{GA} + z_{BA} \tag{4.3}$$

The resultant acceleration, measured by an accelerometer, is the vector sum of all of the accelerations acting on the device along the sensitive axis - Equations 4.1, 4.2, 4.3. This is equal to the gravitational acceleration component plus the body acceleration component, neglecting the effects of noise - Equation 4.4.

$$\begin{aligned}\rho &= \sqrt{x^2 + y^2 + z^2} = \sqrt{(x_{GA} + x_{BA})^2 + (y_{GA} + y_{BA})^2 + (z_{GA} + z_{BA})^2} = \\ &= \sqrt{(x_{GA}^2 + y_{GA}^2 + z_{GA}^2) + (x_{BA}^2 + y_{BA}^2 + z_{BA}^2) + 2(x_{GA}x_{BA} + y_{GA}y_{BA} + z_{GA}z_{BA})} = \\ &= \sqrt{\rho_{GA} + \rho_{BA} + 2(x_{GA}x_{BA} + y_{GA}y_{BA} + z_{GA}z_{BA})}\end{aligned} \tag{4.4}$$

where ρ is the acceleration magnitude vector.

All human movement contains some postural reorientation, therefore, when the TA is worn by a person, changes in the acceleration signals are made up of simultaneous changes in the GA and BA components. As there are temporal and frequency overlaps between the two components, it is not possible to perfectly separate them, and approximations must be made. In the present work, a cut-off frequency of 0.25 Hz was chosen, as it is consistent with the frequencies used in other research works. For example, [64] and [65] choose to use 0.5 Hz, while [66] choose 0.1 Hz.

Figure 4.1 illustrates the motion data processing of a typical recording from the accelerometer, showing seven minutes of recorded data during a supervised test where the subject is asked to perform specific tasks. In order to isolate the BA Component, a Butterworth High-Pass Filter was used. The acceleration signal in blue represents the original signal (raw data), in red it is presented the GA component and the green color represents the filtering result.

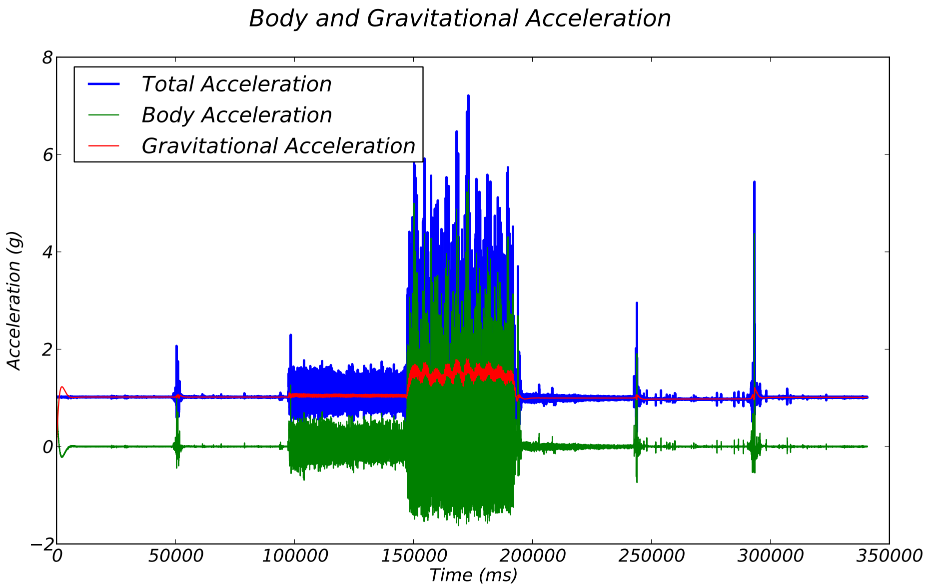


Figure 4.1: Body and Gravitational Acceleration of Signal Accelerometer Sensor.

4.2 Feature Design

TA are made up of 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 - Equation 4.5:

$$ACC_{tot} = \sqrt{ACC_x^2 + ACC_y^2 + ACC_z^2} \tag{4.5}$$

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 [41] and [67]. **Algorithm 4** receives as inputs the original ACC signal, the sampling frequency of data, the cut-off frequency and the order of the filter.

Algorithm 4: `Body Acceleration Component`

`Input: ACC data, sampling frequency, cut-off frequency, filter order.`

`Output: filtered signal.`

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.

4.3 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. 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 [40]. In this section, 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.

4.3.1 Preprocessing Techniques: Domains and Approaches

The present study investigates a new method of feature extraction for clustering techniques. A dictionary of features that were later extracted from the motion data, was created, in a JavaScript Object Notation (JSON) [68]. **Algorithm 5** describes how the feature dictionary was built. The pseudo code to create this dictionary is shown below:

Algorithm 5 Feature Dictionary

$data \leftarrow jsonLoad(open($feature_Dictionary.Json$))$ ▷ Load Feature Dictionary
$domain \leftarrow data.keys()$ ▷ Find different domains
for $feat$ in $domain$ **do** ▷ Goes through all features
    $DomainFeats \leftarrow data[feat].keys()$
    for $params$ in $DomainFeats$ **do** ▷ Collected all information
        $data[feat][params] \leftarrow dict(Default.items() + data[feat][params].items())$
    end for
end for

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 4.1.

Some of these parameters are common for a certain group of features. A default dictionary was created with that information, and afterwards it was compared to the

Table 4.1: Collected Information from each Feature.

Description	Brief description of the information extracted from the feature.
Imports	Necessary imports for the feature extraction properly work.
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.

features dictionary. For example, for the histogram, the extracted parameters are shown in Figure 4.2

```
{"statistical domain":
   {"histogram":
      "description": "The grouping of data into bins.
      Number of members in each bin against the bin number.",
      "imports":"from openSignalsAccel.features import hist",
      "function": "hist",
      "free parameters":"nbins": [10, 20, 3], "r": [1,3,2],
      "number of features": 3,
      "source": "Python 2.7, Numpy Library",
      "reference": "Python 2.7, Numpy Library"
      }
}
```

Figure 4.2: Feature Dictionary: histogram collected information.

4.3.2 Relevant Features for HAR Systems

Recognizing human activities depends directly on the features extracted for motion analysis. People tend to perform the same movement in a variety of 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 of high importance. A good feature set should show little variation between repetitions of the same movements and across different subjects, but should vary considerably between different activities.

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, as described in Chapter 3. Others, like MFCC have appeared for the first time in PA recognition and have revealed to be quite promising due to the achieved performances. By manipulating this dictionary, it is possible the reproduction of clustering tests with different feature combination. Table 4.2 shows the list of features considered in the present work.

Table 4.2: 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*

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. This way, 180^{th} - dimensional feature vector was obtained.

Algorithm 6 creates a matrix with n-samples by m-features. It also outputs an array with feature names. This feature vector has the ability to describe all known characteristics of any instance. Features are then fed to a module which implements a specific clustering algorithm.

Algorithm 6: `Feature Extraction`

```
Input:  motion data, window length, sampling frequency,
feature dictionary, matrix of free parameter combinations,
overlap percentage.
Output:  feature vector and an array with feature names.
```

4.3.3 Feature Normalization

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 must be normalized to zero mean and unit variance, using Equation 4.6:

$$f_{normalized} = \frac{f_{row} - \mu}{\sigma} \tag{4.6}$$

where μ and σ are the empirical mean and standard deviation of a particular feature across all activity classes.

4.4 Graphical Perception of Feature Visualizations: The Horizon Plot

Some studies investigate techniques for the visualization of time series data and evaluate their effect in value comparison tasks [6]. Line charts were compared with horizon graphs - an efficient time series visualization technique across a range of chart sizes [6]. 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.

Figure 4.3 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).

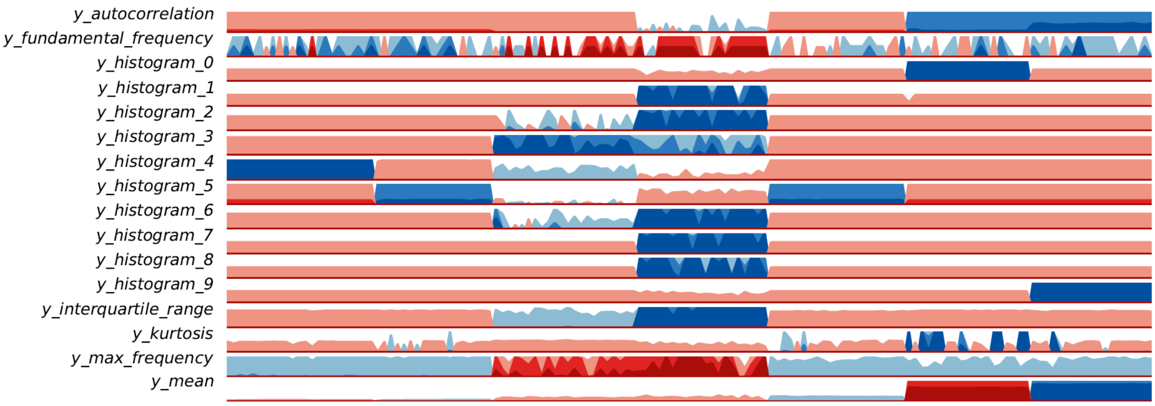


Figure 4.3: Horizon Graph: Time Series Visualization Technique.

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 [6]. Finally, the chart is divided into bands and overlaid, again halving the height.

4.5 Feature Selection Techniques

To systematically assess the usefulness and identify the most important features for discriminating different activities, a simple feature selection technique was implemented. 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. For each signal, different combinations of features, free parameters of that features and signal window size are tested, in order to obtain a better performance of the implemented algorithm. The set of the best features are identified depending on the resulting clustering accuracies for each feature.

4.5.1 Choosing an Appropriate Window-Length

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 [69]. If the subject is performing a single activity, then a long window will include more information about that activity. Choose long windows is based on that the windowed data only contains a single activity, which is how laboratory data usually is collected (long bouts of a single activity) [70]. However, the assumption that the window, no matter how long, will only contain a single activity is likely to be violated in daily life where activity changes are uncontrolled.

In the present study, no window size was stipulated, but a combination of different values which grow in a logarithmic scale. According to Table 4.3, tests were performed with window size ranging from 1000 to 4000 samples, in a logarithmic scale. For each window size, different clustering performances were obtained that allow the choice of an appropriate window size to feature extraction.

4.5.2 Free Parameters of HAR Features

After filtering and windowing the signal, features were extracted from data. In order to make the implemented code versatile, a matrix with the values of all the possible combinations that these parameters can take, was created. Tests were made to determine the free parameters in each activity that allow a better activity recognition performance. **Algorithm 7** computes the free parameters of features available in feature dictionary.

Algorithm 7 Free Parameters and respective Range of Values

$domains \leftarrow dictionary.keys()$ ▷ Goes through each domain
$total\_free\_parameters \leftarrow \{\}$
for i in range(0, length(signal) - 1) **do**
  if $dictionary[feat][parameters]['use'] != 'no'$ **then** ▷ Verify used features
    $free\_parameters \leftarrow dictionary[feat][parameters]['free\_parameters']$
  end if
  if $free\_parameters != [empty]$ **then** ▷ Add information
    $total\_free\_parameters.update(free\_parameters)$
  end if
end for

This algorithm computes all the possible combinations of free parameters. For instance, histogram receives as inputs the range and the number of bins. The value given to these parameters will dictate the clustering algorithm performance. This way, a 486-dimensional free parameter combination vector was obtained.

Table 4.3: Different combinations of Free Parameters and Window Size of the Signal.

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

4.6 Unsupervised Learning

The machine learning algorithms based on the feature representation of accelerometer data have become the most widely used approaches in PA prediction [35]. In this work, 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. The k-means algorithm [39] 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.

4.6.1 Signal Annotation: Creating of a Ground Truth

In this work, an explored aspect of activity recognition is the method applied to annotate sample data that 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 [71, 72]. 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. In other cases, the raw sensor data was manually inspected in order to annotate it with a corresponding activity label [73].

In an unsupervised approach, motion data has to be annotated to compute the performance of the algorithm. 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.

In this 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 . For each signal, an annotation, in JSON format [68], was created, as shown in the Figure 4.4.

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

Figure 4.4: Ground Truth: Annotation Structure.

The annotation dictionary has information about the number and label of the movements that took place and the time intervals that delimit them. The created file, with the annotated signal, has *.ann* extension. 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.

5

Performance Evaluation

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 from waist accelerometer data. The conditions in which data was collected from movements performed in a supervised laboratory setting are described in this chapter.

5.1 System Architecture and Data Acquisition

In the field of wearable sensor based recognition of bodily activities, recognition algorithms can be evaluated on the basis of the complexity of the activities they recognize. People perform a large number of different activities in daily life. The complexity of the activities can vary and depends on different factors including the number of activities, the types of activities and the complexity of the data collected for those activities (collected either in the laboratory or free-living conditions).

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.

In this study, the experiments have been carried out with a group of 8 volunteers within an age range of 16-44 years. The first test consisted in the performing of a gym circuit, in a supervised atmosphere. Each person performed seven activities - standing, sitting, walking, running and lying down (belly up, right side down and left side down) - wearing an accelerometer on the waist. Each activity lasts about one minute. Table 5.1 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. The data acquisition was performed with OpenSignals platform [3] and saved in h5 format. The collected data was processed offline using Python Programming Language [44].

Table 5.1: 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

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 ARI is a function that measures the similarity of the two assignments, ignoring permutations and with chance normalization [39].

Once the clustering algorithm randomly associates the clustering results to non- annotated groups, a function metric that links these groups to their corresponded activity, was implemented. Clustering tests were performed, individually, for each subject and with the respectively concatenated data: in a subject-dependent and a subject-independent context.

5.1.1 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 gather 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 the presented 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 5.2 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.

Figure 5.1 shows the performance behaviour as a function of different window size of motion data.

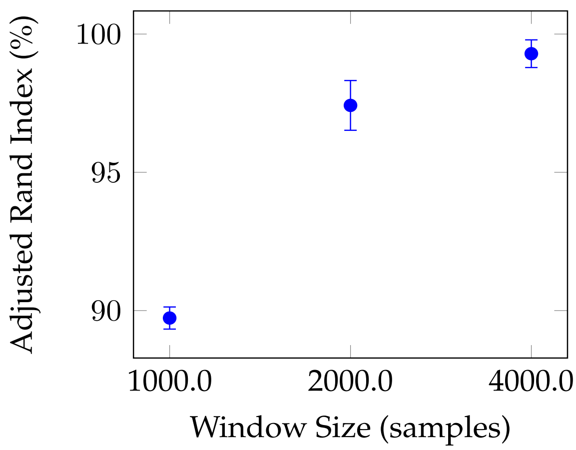


Figure 5.1: Clustering Performance (mean value) as a function of different window length extracted from the best set of features.

Based on these results, the HAR system reaches an accuracy between 89.73% $\pm 0.4\%$ and 99.29% $\pm 0.5\%$, with 1000 and 4000 samples, respectively.

Table 5.2: Clustering Performance (mean value) as a function of different window length extracted from the best set of features.

Window Size of the Motion Data	Clustering Accuracy (%)
1000 samples	89.73% $\pm 0.4\%$
2000 samples	97.42% $\pm 0.9\%$
4000 samples	99.29% $\pm 0.5\%$

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.

5.2 Clustering Evaluation for Subject-independent Context

The subject-independent performance was evaluated with 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% were 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.

A confusion matrix contains information about true and predicted labels done by a clustering system. The performance of such systems is commonly evaluated using the data in the matrix. Table 5.3 shows the confusion matrix for all subjects data, where label i, with i = {1, 2, ...,7}, corresponds respectively to: standing, sitting, walking, running, lying down (belly up), lying down (right side down) and lying down (left side down). This experiment was repeated four times.

The algorithm successfully distinguish all activities, with exception of sitting and standing activities, where in 37% of the time, the algorithm confuses sitting with standing position.

Table 5.3: Confusion Matrix, in percentage, for all subjects, where Lying Down$^{(1)}$ is lying down (belly up), Lying Down$^{(2)}$ is lying down (right side down) and Lying Down$^{(3)}$ is lying down (left side down).

	Standing	**Sitting**	**Walking**	**Running**	**Lying Down$^{(1)}$**	**Lying Down$^{(2)}$**	**Lying Down$^{(3)}$**
Standing	**90**	0	0	0	8	1	1
Sitting	37	**60**	1	0	0	2	0
Walking	0	0	**100**	0	0	0	0
Running	0	0	0	**100**	0	0	0
Lying Down$^{(1)}$	1	2	0	0	**80**	9	8
Lying Down$^{(2)}$	0	0	0	0	1	**90**	9
Lying Down$^{(3)}$	0	0	0	0	0	0	**100**

5.3 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 [39] 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. Table 5.4 shows the obtained clustering performance (mean and standard deviation) for each subject with the best set of features.

Table 5.4: Clustering Accuracy (mean and standard deviation) per subject.

Subject	**Clustering Accuracy (%)**
Subject 1	99.78% ±0.2%
Subject 2	98.91% ±0.3%
Subject 3	99.46% ±0.2%
Subject 4	98.57% ±0.1%
Subject 5	99.86% ±0.1%
Subject 6	99.65% ±0.3%
Subject 7	99.40% ±0.4%
Subject 8	98.72% ±0.2%

Tables 5.5, 5.6, 5.7, 5.8, 5.9, 5.10, 5.11 and 5.12 show the confusion matrix for each subject with the best set of features: root mean square, autocorrelation, mean and MFCC. In order to find out which activities are relatively harder to be recognized, the confusion matrices were analysed.

Table 5.5: Confusion Matrix, in percentage, for the first subject where Lying Down$^{(1)}$ is lying down (belly up), Lying Down$^{(2)}$ is lying down (right side down) and Lying Down$^{(3)}$ is lying down (left side down).

	Standing	**Sitting**	**Walking**	**Running**	**Lying Down$^{(1)}$**	**Lying Down$^{(2)}$**	**Lying Down$^{(3)}$**
Standing	**99**	1	0	0	0	0	0
Sitting	2	**98**	0	0	0	0	0
Walking	0	0	**100**	0	0	0	0
Running	0	0	0	**100**	0	0	0
Lying Down$^{(1)}$	0	0	0	0	**100**	0	0
Lying Down$^{(2)}$	0	0	0	0	0	**100**	0
Lying Down$^{(3)}$	0	0	0	0	0	0	**100**

Table 5.6: Confusion Matrix, in percentage, for the second subject where Lying Down$^{(1)}$ is lying down (belly up), Lying Down$^{(2)}$ is lying down (right side down) and Lying Down$^{(3)}$ is lying down (left side down).

	Standing	**Sitting**	**Walking**	**Running**	**Lying Down$^{(1)}$**	**Lying Down$^{(2)}$**	**Lying Down$^{(3)}$**
Standing	**100**	0	0	0	0	0	0
Sitting	42	**50**	3	2	0	1	2
Walking	0	0	**100**	0	0	0	0
Running	0	0	0	**100**	0	0	0
Lying Down$^{(1)}$	0	0	0	0	**100**	0	0
Lying Down$^{(2)}$	0	0	0	0	0	**100**	0
Lying Down$^{(3)}$	0	0	0	0	0	0	**100**

Table 5.7: Confusion Matrix, in percentage, for the third subject where Lying Down$^{(1)}$ is lying down (belly up), Lying Down$^{(2)}$ is lying down (right side down) and Lying Down$^{(3)}$ is lying down (left side down).

	Standing	Sitting	Walking	Running	Lying Down$^{(1)}$	Lying Down$^{(2)}$	Lying Down$^{(3)}$
Standing	**100**	0	0	0	0	0	0
Sitting	3	**95**	1	0	0	1	0
Walking	0	0	**100**	0	0	0	0
Running	1	11	0	**85**	0	3	0
Lying Down$^{(1)}$	0	0	0	0	**100**	0	0
Lying Down$^{(2)}$	0	0	0	0	0	**100**	0
Lying Down$^{(3)}$	0	0	0	0	0	0	**100**

Table 5.8: Confusion Matrix, in percentage, for the fourth subject where Lying Down$^{(1)}$ is lying down (belly up), Lying Down$^{(2)}$ is lying down (right side down) and Lying Down$^{(3)}$ is lying down (left side down).

	Standing	Sitting	Walking	Running	Lying Down$^{(1)}$	Lying Down$^{(2)}$	Lying Down$^{(3)}$
Standing	**98**	1	1	0	0	0	0
Sitting	3	**95**	1	2	0	0	0
Walking	0	0	**100**	0	0	0	0
Running	0	0	0	**100**	0	0	0
Lying Down$^{(1)}$	0	0	0	0	**100**	0	0
Lying Down$^{(2)}$	0	0	0	0	0	**99**	1
Lying Down$^{(3)}$	0	0	0	0	0	0	**100**

Table 5.9: Confusion Matrix, in percentage, for the fifth subject where Lying Down$^{(1)}$ is lying down (belly up), Lying Down$^{(2)}$ is lying down (right side down) and Lying Down$^{(3)}$ is lying down (left side down).

	Standing	Sitting	Walking	Running	Lying Down$^{(1)}$	Lying Down$^{(2)}$	Lying Down$^{(3)}$
Standing	**98**	1	1	0	0	0	0
Sitting	4	**93**	2	0	0	1	0
Walking	0	0	**100**	0	0	0	0
Running	0	0	0	**100**	0	0	0
Lying Down$^{(1)}$	0	0	0	0	**100**	0	0
Lying Down$^{(2)}$	1	1	0	0	0	**95**	3
Lying Down$^{(3)}$	0	0	0	0	0	1	**99**

Table 5.10: Confusion Matrix, in percentage, for the sixth subject where Lying Down$^{(1)}$ is lying down (belly up), Lying Down$^{(2)}$ is lying down (right side down) and Lying Down$^{(3)}$ is lying down (left side down).

	Standing	**Sitting**	**Walking**	**Running**	**Lying Down$^{(1)}$**	**Lying Down$^{(2)}$**	**Lying Down$^{(3)}$**
Standing	**97**	1	1	0	0	1	0
Sitting	1	**99**	0	0	0	0	0
Walking	0	0	**100**	0	0	0	0
Running	0	0	0	**100**	0	0	0
Lying Down$^{(1)}$	0	0	0	0	**100**	0	0
Lying Down$^{(2)}$	0	0	0	0	1	**98**	1
Lying Down$^{(3)}$	3	1	6	0	2	8	**80**

Table 5.11: Confusion Matrix, in percentage, for the seventh subject where Lying Down$^{(1)}$ is lying down (belly up), Lying Down$^{(2)}$ is lying down (right side down) and Lying Down$^{(3)}$ is lying down (left side down).

	Standing	**Sitting**	**Walking**	**Running**	**Lying Down$^{(1)}$**	**Lying Down$^{(2)}$**	**Lying Down$^{(3)}$**
Standing	**70**	22	6	0	0	1	1
Sitting	8	**90**	0	0	0	1	1
Walking	0	0	**100**	0	0	0	0
Running	0	0	0	**100**	0	0	0
Lying Down$^{(1)}$	0	0	0	0	**100**	0	0
Lying Down$^{(2)}$	0	0	0	0	0	**100**	0
Lying Down$^{(3)}$	0	0	0	0	0	0	**100**

Table 5.12: Confusion Matrix, in percentage, for the eighth subject where Lying Down$^{(1)}$ is lying down (belly up), Lying Down$^{(2)}$ is lying down (right side down) and Lying Down$^{(3)}$ is lying down (left side down).

	Standing	**Sitting**	**Walking**	**Running**	**Lying Down$^{(1)}$**	**Lying Down$^{(2)}$**	**Lying Down$^{(3)}$**
Standing	**99**	1	0	0	0	0	0
Sitting	1	**95**	0	0	2	2	0
Walking	0	0	**100**	0	0	0	0
Running	0	0	0	**100**	0	0	0
Lying Down$^{(1)}$	1	0	0	0	**94**	1	4
Lying Down$^{(2)}$	0	0	0	0	0	**100**	0
Lying Down$^{(3)}$	0	0	0	0	6	3	**91**

5.4 Discussion

The experiment results 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 is a key success in system acceptance. The study investigated several combinations of features used for activity recognition. The achieved findings suggest that using four simple features from time and frequency domains can achieve high recognition accuracy of $99.29\% \pm 0.5\%$ in a person-dependent context and $88.57\% \pm 4.0\%$ in a person-independent context. Compared to the subject-dependent case, the accuracy of all subjects is much lower which can be explained by the variations in human motion for different subjects. Figure 5.2 shows the drawn diagram using Orange Software.

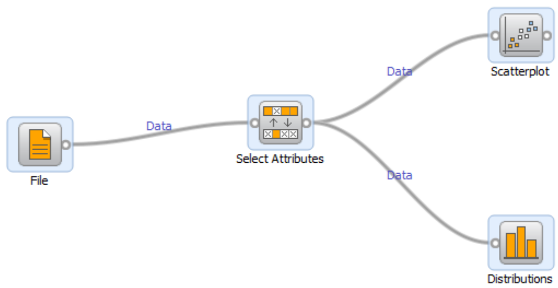
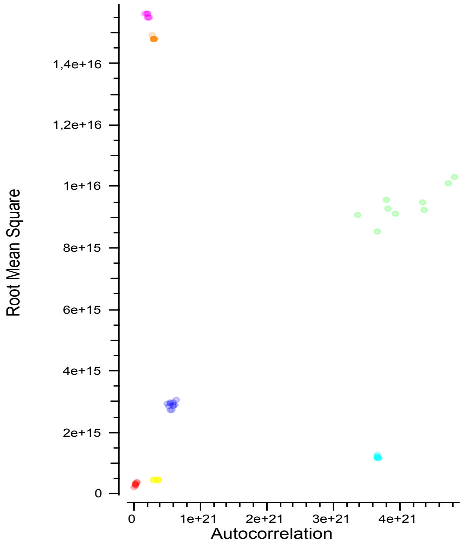


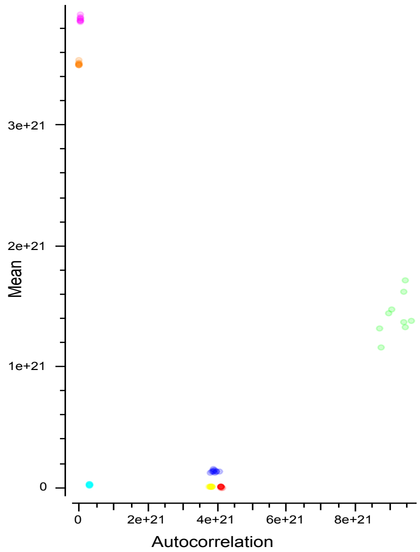
Figure 5.2: Orange Software Visualization Scheme

Figure 5.3 shows the structure of clusters according to the best set of features. In unsupervised learning, unlabelled data is grouping 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 test produces clusters in which the intra-class similarity is high and the inter-class similarity is low. After a feature space clustering procedure, each cluster contains samples of only one activity. This would indicate that the data of the given feature is clearly separable and thus, well-suited as an input for classification.

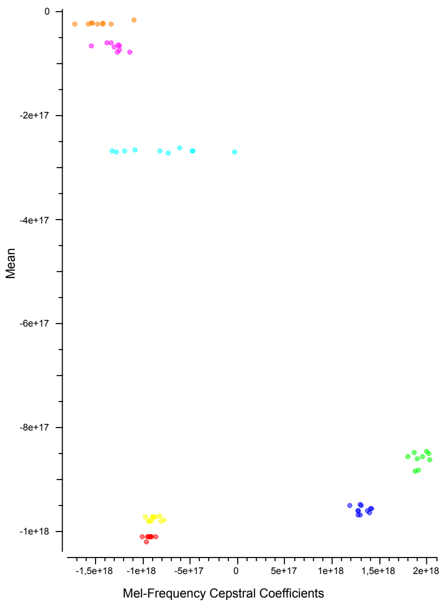
Activities which are static in nature including postures, such as lying and standing, are easier to recognize than the activities which are periodic in nature, such as running and walking. However, postures that are highly similar, such as sitting and standing, are also very hard to discriminate as they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



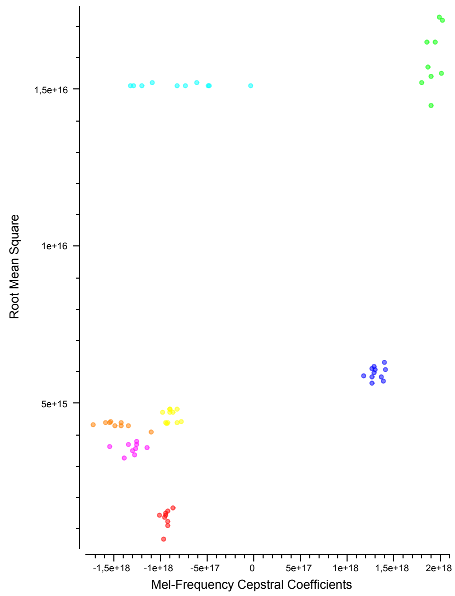
(a) Autocorrelation as a function of Root Mean Square.



(b) Autocorrelation as a function of Mean.



(c) Mean as a function of MFCC.



(d) Root Mean Square as a function of MFCC.

Figure 5.3: Structure of Clusters according to different domain features.

6

Conclusions

To conclude this work, an overview of the general contributions that this research provided for the signal processing area are exposed in this final chapter. A summary of the obtained results is also presented.

6.1 General Contributions and Results

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. A framework to cluster daily physical activities using 3-dimensional ACC data was developed. In the presented experiment, motion data described seven categories of PA and was collected by a wearable sensors worn on subjects' waist to validate the effectiveness of the proposed HAR clustering scheme.

A methodology to search for the best features able to distinguish physical activities was presented. The results indicate that the choice of features is crucial for the success of a recognition algorithm. The techniques that operate on the time domain and frequency domains, as well as on data representations that can be used to discriminate between user activities such as Horizon Plot were described.

The proposed PA clustering scheme, including static and dynamic activity analysis, was capable of cluster the time-series data. The obtained results in clustering accuracy of human activities recognition were very encouraging: an average person-dependent ARI of 99.29% and a person-independent ARI of 88.57% were reached.

Although several systems have been proposed in the past to monitor daily physical activities using accelerometers, the presented framework is promising in several regards: we propose a new set of features extracted from wearable data that are competitive from computational point of view and able to ensure high classification results comparable with the state of the art wearable systems.

The major achievements of the current work, compared to the state of the art are: the presented study performs tests in intra and inter subject context; a set of 180 features is implemented, which are easily selected to test different groups of subjects and different activities and the implemented algorithm does not stipulate, *a priori*, any value for window length, or overlap percentage, but performs a search to find the best parameters that define the specific data.

This study presents features of different domains, one previously studied in the recognition of human activity, and others only used in other areas such as speech recognition.

A clustering metric based on the construction of the data confusion matrix is also proposed.

6.2 Future Work

The presented research leaves a few opened aspects, which are going to be explored in the future. In order to realize the full potential in HAR systems, some topics need further investigation:

- **More Data**: Increase the number of subjects and applications.
- **Concurrent and Overlapping Activities**: The assumption that an individual only performs one activity at a time can not be true. In general, human activities can overlapping and even be concurrent. A person could be walking while brushing their teeth, or watching TV while having lunch.
- **More Computing Power**: Use parallel computing infrastructures on the data collected.
- **Bigger Timespan**: Week long acquisitions of movement.

- **Additional Sensors for Recognition of High-Level Activities**: This work has mainly investigated the use of acceleration sensors for activity recognition. This type of sensor has many advantages, such as being versatile, well-understood, small, lightweight, and proven to lead to good recognition results for many types of physical activities. However, especially when moving towards high-level activities, other sources of information can be helpful.

- **More Discoveries**: Detect the behaviour changes and annotate it.

In the future, this framework has to be tested on other intensity varying activities and across more subjects. For example, test this framework on individuals running and walking at a greater range of intensity levels.

ACC data obtained from wearable accelerometers can be synchronized with the activity of daily living data recorded by such monitoring systems to better describe the information of PA, human mobility, behavioural pattern, and functional ability that encompass the important parameters regarding the overall health status of an individual.

Other objective is to test different clustering algorithms: Affinity Propagation [74], Density-based Spatial Clustering of Applications with Noise (DBSCAN) [75] and Optics [76].

The continuously need to obtain more information, more efficient, more quickly and with less intervention from an expert has led to a growing application of signal processing techniques to motion data. The main challenge for future work in this area will be the development of features and recognition strategies that can work in an ambient assisted living under a wide variety of environmental conditions.

Bibliography

[1] Chandrashekhar H. Human Activity Representation, Analysis and Recognition. *Department of Electrical Engineering , Indian Institute of Technology*, (May), 2006.

[2] Anna Karla Smith. Diagnostics methods for sleep disorders. 27(Supl I):8–15, 2005.

[3] Ricardo Gomes, Neuza Nunes, Joana Sousa, and Hugo Gamboa. Long term biosignals visualization and processing. In *BIOSIGNALS*, pages 402–405, 2012.

[4] Guido Van Rossum and Jelke de Boer. Interactively testing remote servers using the Python programming language. 4:283–304, 1991.

[5] Tomaz Curk, Janez Demsar, Qikai Xu, Gregor Leban, Uros Petrovic, Ivan Bratko, Gad Shaulsky, and Blaz Zupan. Microarray data mining with visual programming. *Bioinformatics*, 21:396–398, 2005.

[6] Jeffrey Heer, Nicholas Kong, and Maneesh 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.

[7] Sasank Reddy, Min Mun, Jeff Burke, Deborah Estrin, Mark Hansen, and Mani Srivastava. Using mobile phones to determine transportation modes. *ACM Transactions on Sensor Networks*, 6(2):1–27, February 2010.

[8] C. Donalek. Supervised and Unsupervised Learning. (April), 2011.

[9] Zhenyu He, Zhibin Liu, Lianwen Jin, Li-xin Zhen, and Jian-cheng Huang. Weightlessness Feature — A Novel Feature for Single Tri-axial Accelerometer based Activity Recognition. *19th International Conference on Pattern Recognition*, pages 0–3, 2008.

[10] National Institute of Mental Health. “Obsessive Compulsive disorder (OCD)". Available at www.nimh.nih.gov/health/topics/obsessive-compulsive-disorder-ocd/index.shtml/.

[11] Luca Palmerini, Sabato Mellone, Guido Avanzolini, Franco Valzania, and Lorenzo Chiari. Classification of early-mild subjects with parkinson’s disease by using sensor-based measures of posture, gait, and transitions. In *AIME*, pages 176–180, 2013.

[12] Bruce J Simon, Joseph P Errico, and John T Raffle. Non-invasive magnetic or electrical nerve stimulation to treat or prevent autism spectrum disorders and other disorders of psychological development, March 3 2013. US Patent App. 13/783,319.

[13] Mi Ryeong Song, Young-Shin Lee, Jong-Duek Baek, and Marci Miller. Physical activity status in adults with depression in the national health and nutrition examination survey, 2005–2006. *Public Health Nursing*, 29(3):208–217, 2012.

[14] Siobhan M Phillips and Edward McAuley. Physical activity and fatigue in breast cancer survivors: A panel model examining the role of self-efficacy and depression. *Cancer Epidemiology Biomarkers & Prevention*, 22(5):773–781, 2013.

[15] Robert Lemoyne, Cristian Coroian, Timothy Mastroianni, and Warren Grundfest. Accelerometers for Quantification of Gait and Movement Disorders: A Perspective Review. *Med. Biol*, 8(2):137–152, 2008.

[16] Roozbeh Jafari, Wenchao Li, Ruzena Bajcsy, Steven Glaser, and Shankar Sastry. Physical Activity Monitoring for Assisted Living at Home. *4th International Workshop on Wearable and Implantable Body Sensor Networks*, page 11, 2007.

[17] E. Villalba, M. Ottaviano, M. T. Arredondo, A. Martinez, and S. Guillen. Wearable Monitoring System for Heart Failure Assessment in a Mobile Environment Life Supporting Technologies group , Technical University of Madrid , Spain. 33:237–240, 2006.

[18] C. V. Bouten, K. T. Koekkoek, M. Verduin, R. Kodde, and J. D. Janssen. A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. *IEEE Transactions on Biomedical Engineering*, 44(3):136–147, 1997.

[19] J. Fahrenberg, F. Foerster, M. Smeja and W. Müller. Assessment of posture and motion by multichannel piezoresistive accelerometer recordings, 1997.

[20] David R. Bassett, Alex Rowlands, and Stewart G. Trost. Calibration and Validation of Wearable Monitors. *Medicine & Science in Sports & Exercise*, 44:S32–S38, January 2012.

[21] Yonglei Zheng, Weng-Keen Wong, Xinze Guan, and Stewart Trost. Physical activity recognition from accelerometer data using a multi-scale ensemble method. In *Twenty-Fifth Annual Conference on Innovative Applications of Artificial Intelligence. IAAI*, 2013.

[22] S.Lee A.Khan. Accelerometer's position free human activity recognition using a hierarchical recognition model. *eHealth Networking Applications and Services Healthcom 2010 12th IEEE International Conference on*, pages 296–301, 2010.

[23] Zhenyu He and Lianwen Jin. Activity Recognition from acceleration data Based on Discrete Consine Transform and SVM. *IEEE International Conference on Systems, Man and Cybernetics*, (October):5041–5044, 2009.

[24] S. Lee A. Khan, Y.K. Lee and T.S. Kim. A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer. *IEEE transactions on information technology in biomedicine a publication of the IEEE Engineering in Medicine and Biology Society*, 14(5):1166–1172, 2010.

[25] Justin J. Kavanagh and Hylton B. Menz. Accelerometry: A technique for quantifying movement patterns during walking. *Gait and Posture*, 28(1):1 – 15, 2008.

[26] Andrea Mannini, Stephen S Intille, Mary Rosenberger, Angelo M Sabatini, and William Haskell. Activity recognition using a single accelerometer placed at the wrist or ankle. *Medicine and science in sports and exercise*, 2013.

[27] Jiahui Wu, Gang Pan, Daqing Zhang, Guande Qi, and Shijian Li. Gesture recognition with a 3-d accelerometer. In *Ubiquitous intelligence and computing*, pages 25–38. Springer, 2009.

[28] R.E. Andersen. Obesity: Etiology, Assessment, Treatment, and Prevention. *Human Kinetics*, pages 111–117, 2003.

[29] A. Godfrey, R. Conway, D. Meagher, and G. OLaighin. Direct measurement of human movement by accelerometry. *Medical Engineering and Physics*, 30(10):1364 – 1386, 2008.

[30] Lei Gao, Alan K Bourke, and John Nelson. Sensor positioning for activity recognition using multiple accelerometer-based sensors. In *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN)*, 2013.

[31] Yunyoung Nam, Seungmin Rho, and Chulung Lee. Physical activity recognition using multiple sensors embedded in a wearable device. *ACM Transactions on Embedded Computing Systems (TECS)*, 12(2):26, 2013.

[32] F. Bugané, M. G. Benedetti, G. Casadio, S. Attala, F. Biagi, M. Manca, and A. Leardini. Estimation of spatial-temporal gait parameters in level walking based on a single accelerometer: Validation on normal subjects by standard gait analysis. *Comput. Methods Prog. Biomed.*, 108(1):129–137, October 2012.

[33] M.W. Whittle. Gait Analysis: An Introduction. *Butterworth-Heinemann*, 1996.

[34] Uwe Maurer, Asim Smailagic, Daniel P Siewiorek, and Michael Deisher. Activity Recognition and Monitoring Using Multiple Sensors on Different Body Positions. *International Workshop on Wearable and Implantable Body Sensor Networks, (Washington, DC, USA), IEEE Computer Society*, pages 4–7, 2006.

[35] I. H. Witten, E. Frank and M. A. Hall. *Data Mining: Practical Machine Learning Tools and Techniques*. 2011.

[36] Stuart P. Lloyd. Least squares quantization in pcm. *IEEE Transactions on Information Theory*, 28:129–137, 1982.

[37] Shi Na, Liu Xumin, and Guan Yong. Research on k-means clustering algorithm: an improved k-means clustering algorithm. In *Intelligent Information Technology and Security Informatics (IITSI), 2010 Third International Symposium on*, pages 63–67. IEEE, 2010.

[38] T. Warren Liao. Clustering of time series data—A survey. *Pattern Recognition*, 38(11):1857–1874, November 2005.

[39] 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. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[40] Z. Ghahramani. Unsupervised learning. *Gatsby Computational Neuroscience Unit, University College London*, 2004.

[41] M. Mathie. Monitoring and interpreting human movement patterns using a triaxial accelerometer. 2003.

[42] T.M. Mitchell. *Machine Learning*. McGraw-Hill, New York, 1997.

[43] Ujjwal Maulik and Sanghamitra Bandyopadhyay. Performance evaluation of some clustering algorithms and validity indices. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(12):1650–1654, 2002.

[44] T. Oliphant. Guide to Numpy. Tregol Publishing. 2006.

[45] Waltenegus Dargie, Juha Plosila, and Vincenzo De Florio. Existing challenges and new opportunities in context-aware systems. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, pages 749–751, New York, NY, USA, 2012. ACM.

[46] Do-Un Jeong, Se-Jin Kim, and Wan-Young Chung. Classification of posture and movement using a 3-axis accelerometer. In *Proceedings of the 2007 International Conference on Convergence Information Technology*, ICCIT '07, pages 837–844, Washington, DC, USA, 2007. IEEE Computer Society.

[47] P. H. Veltink, H. B. Bussmann, W. De Vries, W. L. Martens, and R. C. Van Lummel. Detection of Static and Dynamic Activities using Uniaxial Accelerometers., 1996.

[48] N. Rafiuddin, Y. Uzzaman Khan, and O. Farooq. Feature extraction and classification of eeg for automatic seizure detection. In *Multimedia, Signal Processing and Communication Technologies (IMPACT), 2011 International Conference on*, pages 184–187, 2011.

[49] M. Sekine, T. Tamura, T. Fujimoto, and Y. Fukui. Classification of walking pattern using acceleration waveform in elderly people. *Engineering in Medicine and Biology Society*, 2:1356 – 1359, Jul 2000.

[50] Cliff Randell and Henk L. Muller. Context awareness by analyzing accelerometer data. In *ISWC*, pages 175–176. IEEE Computer Society, 2000.

[51] Albrecht Schmidt and Kristof Van Laerhoven. How to build smart appliances? *IEEE PERSONAL COMMUNICATIONS*, 8:66–71, 2001.

[52] Nishkam Ravi, Nikhil Dandekar, Preetham Mysore, and Michael L Littman. Activity Recognition from Accelerometer Data. *In Proceedings of the Seventeenth Conference on Innovative Applications of Artificial Intelligence (IAAI)*, 2005.

[53] J. L. Rodgers and W. A. Nicewander. Thirteen ways to look at the correlation coefficient. *The American Statistician*, 42:59–66, 1988.

[54] G.E.P. Box and G.M. Jenkins. *Time Series Analysis: Forecasting and Control, Revised Ed.* Holden-Day series in time series analysis. Holden-Day, 1976.

[55] Kevin Englehart, Bernard Hudgins, Philip Parker, and Maryhelen Stevenson. Time-frequency representation for classification of the transient myoelectric signal. In *In ICEMBS'98. Proceedings of the 20th Annual International Conference on Engineering in Medicine and Biology Society. ICEMBS*, pages 2627–2630. Press, 1998.

[56] B. P. Bogert, M. J. R. Healy, and J. W. Tukey. The quefrency alanysis of time series for echoes: cepstrum, pseudo-autocovariance, cross-cepstrum, and saphe cracking. In M. Rosenblatt, editor, *Proceedings Symp. Time Series Analysis*, pages 209–243, 1963.

[57] Judith C. Brown. Computer identification of musical instruments using pattern recognition with cepstral coefficients as features, 1997.

[58] J. Volkmann, S. S. Stevens, and E. B. Newman. A scale for the measurement of the psychological magnitude pitch. *The Journal of the Acoustical Society of America*, 8(3):208–208, 1937.

[59] Todor Ganchev, Nikos Fakotakis, and George Kokkinakis. Comparative evaluation of various mfcc implementations on the speaker verification task. In *in Proc. of the SPECOM-2005*, page 191–194, 2005.

[60] Rajesh Kumar Dubey and Arun Kumar. Non-intrusive speech quality assessment using several combinations of auditory features. *Int. J. Speech Technol.*, 16(1):89–101, March 2013.

[61] Douglas A. Reynolds. Speaker identification and verification using gaussian mixture speaker models. *Speech Commun.*, 17(1-2):91–108, August 1995.

[62] G. Peeters. A large set of audio features for sound description (similarity and classification) in the CUIDADO project. 2004.

[63] Tae Hong Park. Introduction to digital signal processing - computer musically speaking. pages I–XX, 1–429, 2009.

[64] J. Fahrenberg F. Foerster M. Smeja and W. Muller. Assessment of posture and motion by multichannel piezoresistive accelerometer recordings. *Psychophysiology*, page 34(5):607—612, 1997.

[65] F. Foerster and J. Fahrenberg. *Motion pattern and posture: correctly assessed by calibrated accelerometers*, page 32(3):450—457, 2000.

[66] Adil Mehmood Khan, Young-Koo Lee, Sungyoung Y. Lee, and Tae-Seong Kim. A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer. *Trans. Info. Tech. Biomed.*, 14(5):1166–1172, September 2010.

[67] A. Mannini and A.M Sabatini. Machine learning methods for classifying human physical activities from on-body sensors., 2010.

[68] Douglas Crockford. The application/json media type for javascript object notation (json). RFC 4627, IETF, 7 2006.

[69] Konrad Schindler and Luc Van Gool. Action snippets: How many frames does human action recognition require? In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.

[70] Christos Faloutsos, Mudumbai Ranganathan, and Yannis Manolopoulos. *Fast subsequence matching in time-series databases*, volume 23. ACM, 1994.

[71] Lin Liao, Dieter Fox, and Henry Kautz. Location-based activity recognition using relational markov networks. In *Proceedings of the 19th international joint conference on Artificial intelligence*, IJCAI'05, pages 773–778, San Francisco, CA, USA, 2005. Morgan Kaufmann Publishers Inc.

[72] Emmanuel Munguia Tapia, Stephen S. Intille, and Kent Larson. Activity recognition in the home using simple and ubiquitous sensors. In *In Pervasive*, pages 158–175, 2004.

[73] Christopher R. Wren and Emmanuel Munguia Tapia. Toward scalable activity recognition for sensor networks. In *In Lecture Notes in Computer Science*, pages 168–185. Springer, 2006.

[74] Brendan J. Frey and Delbert Dueck. Clustering by passing messages between data points. *Science*, 315(5814):972–976, 2007.

[75] Martin Ester, Hans peter Kriegel, Jörg S, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. pages 226–231. AAAI Press, 1996.

[76] Mihael Ankerst, Markus M. Breunig, Hans-Peter Kriegel, and Jörg Sander. Optics: ordering points to identify the clustering structure. *SIGMOD Rec.*, 28(2):49–60, June 1999.

7

Publications

In this appendix is presented the publication, Human Activity Recognition from Triaxial Accelerometer Data: Feature Extraction and Selection Methods for Clustering of Physical Activities which demonstrates the algorithm that was developed during this dissertation. This article was accepted to BIOSIGNALS 2014, which is a conference – 7^{th} International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2014), held in Paris in March 2014.

Human Activity Recognition from Triaxial Accelerometer Data
Feature Extraction and Selection Methods for Clustering of Physical Activities

Inês Machado[1], Ricardo Gomes[1], Hugo Gamboa[1,2] and Vítor Paixão[3]
[1]*CEFITEC, Physics Department, FCT-UNL, Lisbon, Portugal*
[2]*PLUX - Wireless Biosignals, Lisbon, Portugal*
[3]*Champalimaud Foundation, Lisbon, Portugal*
ipratamachado@gmail.com, ricardo.baptista.gomes@gmail.com, hgamboa@plux.info, vitor.paixao@neuro.fchampalimaud.org

Keywords: Physical Activity Recognition, Signal Processing, Feature Extraction, Feature Selection, Unsupervised Learning.

Abstract: The demand for objectivity in clinical diagnosis has been one of the greatest challenges in Biomedical Engineering. The study, development and implementation of solutions that may serve as ground truth in physical activity recognition and in medical diagnosis of chronic motor diseases is ever more imperative. This paper 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. In this paper, unsupervised learning is applied to the feature representation of accelerometer data to discover the activities performed by different subjects. A feature selection framework is developed in order to improve the clustering accuracy and reduce computational costs. The features which best distinguish a particular set of activities are selected from a 180^{th}- dimensional feature vector through machine learning algorithms. The implemented framework achieved very encouraging results in human activity recognition: an average person-dependent Adjusted Rand Index (ARI) of $99.29\% \pm 0.5\%$ and a person-independent ARI of $88.57\% \pm 4.0\%$ were reached.

1 INTRODUCTION

The constant concern with the human physical and psychological well-being has been the drive for research studies that have led to a promising evolution of medicine and engineering. The study, development and implementation of solutions that may serve as ground truth in physical activity recognition and in medical diagnosis of chronic motor diseases is ever more imperative. In this paper, a Human Activity Recognition (HAR) framework is developed using a wearable 3-dimensional accelerometer sensor. The main focus of this paper is to understanding the signals produced by a Triaxial Accelerometer (TA), interpreting them in the context of human movement and identifying relevant parameters from the data. The versatility of the algorithm enables the identification of relevant features able to recognize simple daily activities. We obtain a 180^{th}- dimensional feature vector from statistical, time and frequency domains. The dimensionality of the feature vector should be as small as possible by reducing the amount of irrelevant and redundant information in the data, not only to reduce the computation complexity, but also to obtain better clustering performance. The remainder of the paper is organized as follows: Section 2 describes the background and related work. The importance of objective monitoring human movement is discussed. That section also presents an overview on other studies about HAR with wearable sensors. Section 3 explains the composition of the TA signal. The signal is made up of several components, and each of these is examined. The difficulties in distinguish between the different signal components are discussed. Section 4 describes the proposed methodology used in this work to extract and select features based on motion data. Section 5 describes the architecture of the acquisition system and the obtained results. Section 6 presents the conclusions obtained from the investigation and some future research directions.

2 BACKGROUND

In recent decades, there has been an increasing interest in the use of Accelerometry (ACC) to monitor human behaviour. Monitoring human movement can provide valuable information on a patient and some parameters of movement can provide information of health status, rate of rehabilitation and other potentially useful clinical data. The advance of technology has helped the development of accelerometers of small size and low cost, making them a very convenient tool for monitoring subjects. One of the key point is the diversity of areas where ACC has been used in the past. The most studied have being: metabolic energy expenditure, Physical Activity (PA), balance and postural sway, sit-to-stand transfers (which is an important indicator for postural instability) and detection of falls. The use of accelerometers has also allowed to help on diagnose of a number of diseases such as Parkinson's Disease (Palmerini et al., 2013), Autism Spectrum Disorder, (Bandini et al., 2013) and Depression (Phillips and McAuley, 2013).

3 TRIAXIAL ACCELEROMETER SIGNAL

The signal measured by each fixed-body accelerometer is a linear sum of, approximately, three components (Mathie, 2003):

- Body Acceleration Component: acceleration resulting from body movement;
- Gravitational Acceleration Component: acceleration resulting from gravity;
- 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 - Body Acceleration Component - is important, however these two components have overlapping frequency spectra. The BA component ranges from above 0 Hz to possibly up 20 Hz, but is mostly contained in the range above 0 and below 3 Hz. This range overlaps the area covered by the GA component, which goes from 0 to several Hertz. It is possible to approximately separate the BA and the GA components with some filtering. A wide range of different filters types with different characteristics and different windowing percentages were tested in previous studies, as in (Mathie, 2003), in order to determine their ability to differentiate the components of the acceleration signal. In the presented study, a cut-off frequency of 0.25 Hz was chosen, as it is consistent with the frequencies used in other research works. (Smeja and Muller, 1997) and (Foerster and Fahrenberg, 2000) choose to use 0.5 Hz, while (Khan et al., 2010) choose 0.1 Hz. In the presented study, in order to isolate the BA component, a second-order Butterworth High-Pass filter with cut-off frequency of 0.25 Hz is used. Figure 1 illustrates each component of a typical recording from the accelerometer showing seven minutes of motion data where the subject is asked to perform seven specific tasks.

The placement of the accelerometer is another important point of discussion. A device that is to be worn over extended periods must be designed to be as simple to put on and comfortable to wear in order to encourage compliance of patients. General body motion can be measured with a single accelerometer placed close to the body's center of mass, which is located within the pelvis (Liu, 2013). The advantage of placing the accelerometer attached to the waist is that it allows the monitoring of accelerations near the center of mass. Any movement of the body will cause the center of mass to shift. This study aims to develop a HAR framework, for a waist mounted accelerometer based system.

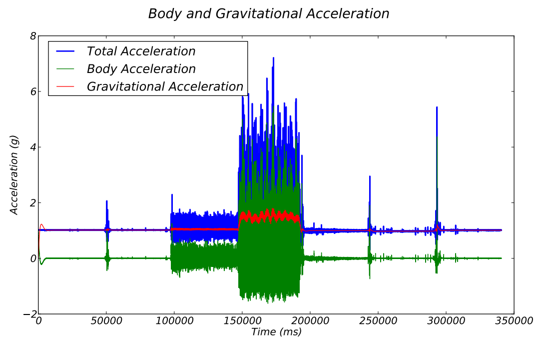


Figure 1: Body and Gravitational Acceleration for each axis of accelerometer sensor.

4 PROPOSED METHODS

Different segmentation methods can be applied to time-series data which enhance signal behaviour and enable the gather of useful information from continuous stream of data such as timing and sliding windows. For activity recognition, where accelerometer data is windowed, the choice of the number of frames is guided by a trade-off between information and resolution. The accelerometer data was collected, cleaned, and preprocessed to extract features that characterize different samples data windows. Clustering mechanisms separate and organize unlabeled data into different groups whose members are similar to each other in some metric. Different approaches generally lead to different clusters. Even for the same algorithm, the parameter identification or the sequence of input patterns may affect the final results. These assessments should be unbiased. In this work, the K-Means Clustering Algorithm (Lloyd, 1982) and a squared Euclidean distance metric were used.

4.1 Feature Design

The HAR strategy depends essentially on the set of features that are extracted from the signal. TA are made up of three separated accelerometer data time series, one time series for acceleration on each orthogonal 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, Equation 1:

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

Each time series ACC_i, with $i = x, y, z$ has been filtered with a second-order Butterworth High-Pass filter with cut-off frequency of 0.25 Hz in order to separate low frequencies component and high frequencies component as suggested in (Mathie, 2003) and (Mannini and Sabatini, 2010). This way, for each time series, three extra time series BA_i are obtained, with $i = x, y, z$, representing the time series with body acceleration component. Finally, features from each one of the time series are extracted.

4.2 Accelerometer Signal Annotation

In unsupervised learning, the motion data has to be annotated to compute the performance of the algorithm. If true class labels are known, the validity of a clustering can be verified by comparing the predicted labels and the true labels. An aspect of activity recognition that has been greatly explored is the method of annotating sample data that can be used to compute the performance of the clustering method. Many experiments use unsupervised learning methods and apply manually annotated test data to evaluate their performances. 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 is manually inspected in order to annotate it with a corresponding activity label (Wren and Tapia, 2006). 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. For each signal, an annotation, in *JavaScript Object Notation* (JSON) (Crockford, 2006) is created, with i the number of activities, Scheme 2:

$$\begin{aligned} &"Labels" : [l_1, , \ldots, l_i], \\ &"Initial\_Times" : [init_1, \ldots, init_i] \\ &"End\_Times" : [end_1, \ldots, end_i] \end{aligned} \tag{2}$$

The 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. The input is an array with the initial and final times of each activity. It also receives as input the window size and the considered overlap percentage.

4.3 Feature Extraction

Recognizing human activities depends directly on the features extracted for motion analysis. A set of features, which will most efficiently and meaningfully represent the information that is important for analysis and the clustering process, is performed. In this section, tests were made in order to assess the following parameters:

- The influence of the signal's 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.

A dictionary of the extracted features from the motion data, was created, in a JSON format (Crockford, 2006). For each feature, the following information was collected: Description, Imports, Use, Metric, Free Parameters, Parameters, Number of Features, Function, Source and Reference. Table 1 shows the high level list of features considered in the presented study. The implemented dictionary divides the features into statistical, temporal and spectral domains. By manipulating this dictionary, the clustering algorithm can be easily tested with a different combination of features. To compute the feature vector the following inputs are needed: motion data, window length of the signal, sampling frequency of the data acquisition, a feature's dictionary, a matrix of free parameter combinations and the considered overlap percentage. For each ACC axis, this function goes through each window of the signal, with the considered window length and overlap percentage and computes a feature matrix with *n*-samples by *m*-features dimension. For each signal, three new files were created: one with the features information per window, one with the names of the features that were extracted for the respective clustering test and another with the label of the activity corresponding to each window. The sensor acceleration signal is made up of three separated accelerometer data time series and complementary to the three axes data, an additional time series have been obtained by computing the magnitude of the acceleration, so four signal vectors are considered. From each window, a vector of features is obtained by calculating features from the statistical, time and frequency domain. This way, a 180^{th} - 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 (MFCC)) with ten outputs each.

Table 1: 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 MFCC Fundamental Frequency Power Bandwidth

Because the scale factors and units of the features described above are different, all the features must be normalized to zero mean and unit variance, before proceed to the feature selection stage.

4.4 Feature Selection for Motion Data

A large number of features can usually be measured in many pattern recognition applications. However, not all features are equally important for a specific task. For each signal, different combinations of features, free parameters of these features and window size of the signal can be tested, in order to evaluate the performance of the implemented clustering algorithm. Optimal features are identified depending on the resulting clustering accuracies for each feature subset.

4.4.1 Free Parameters of Features Set

In order to make the implemented code versatile and the least subjective as possible, a matrix with the values of all the possible combinations that these parameters can take, was created. No window size value was stipulated, but a combination of different values from a growing logarithmic scale can be tested. According to Table 2, tests were made in which the window size ranged from 1000 to 4000 samples, in a log scale. For each window size, different performances were obtained. Tests were made to determine the free parameters in each activity, that allow a better activity recognition performance. Examples of free parameters are the number of bins or the range of the implemented histogram. The values given to these parameters will dictate the performance obtained by the clustering algorithm. In this way, a 486-dimensional free parameter combinations vector was obtained.

Table 2: Possible combinations of free parameters and window size values.

Free Parameter	Range	Combinations
Window Size	[1000 ; 4000]	3
Bins of Histogram	[10 ; 20]	3
Range of Histogram	[1 ; 3]	2
Cepstral C.	[1 ; 11]	3
MFCC	[10 ; 20]	3
Power Bandwidth	[10 ; 20]	3

4.4.2 Graphical Perception of Features Visualizations

A technique for the visualization of time series data and evaluate their effect in value comparison tasks was described in (Heer et al., 2009). In order to visually analyse each feature's behaviour throughout different activities, horizon graphs are used. This procedure ensures a visual perception of the features that better separate certain activities, those which do not change their value between activities and those which only add redundant information. Figure 2 shows an example of a horizon graph generated for a matrix of features, resulting from an ACC signal composed by seven distinct activities. Each activity lasts about one minute and we consider 4000 samples for the window size of the signal. It is possible to quantitatively compare 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. Finally, the chart is divided into bands and overlaid, again halving the height.

4.5 Unsupervised Learning

Machine learning algorithms based on the feature representation of accelerometer data have become the most widely used approaches in PA prediction (I. H. Witten, E. Frank and Hall, 2011). In this work, 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. 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, a priori, from the designed protocol of the performed activities. A good clustering methodology will produce clusters in which the intra-class similarity is high and the inter-class similarity is low. The K-Means Clustering Algorithm (Lloyd, 1982) gives a single set of clusters, with no particular organization or structure within them.

5 DATA ACQUISITION AND RESULTS

The experiments have been carried out with a group of 8 volunteers within an age range of 16-44 years. The test consists in performing of a gym circuit. Each person performs seven activities in sequence lasting about one minute each - standing, sitting, walking, running, lying down (belly up), lying down (right side down) and lying down (left side down), wearing an accelerometer on the waist. Using this system, data with 3-axial acceleration at a constant rate of 800 Hz and 12 bits of resolution was acquired. The data acquisition was performed with OpenSignals platform (Gomes et al., 2012) and saved in a h5 format. The collected data was processed offline using Python Programming Language (Oliphant, 2006). Clustering tests are performed, individually, for each subject and with the respectively concatenated data: in a subject-dependent and a subject-independent context. To evaluate the subject-dependent accuracy of the proposed algorithm, the K-Means Clustering Algorithm (Lloyd, 1982) was performed for each subject data. Given the knowledge of the ground truth class assignments (labels true) and the clustering algorithm assignments of the same samples (predicted labels), the adjusted Rand index (ARI) is a function that measures the similarity of the two assignments, ignoring permutations and with chance normalization. The ARI was calculated to obtain the performance of the clustering method. An average person-dependent accuracy of 99.29% and standard deviation of 0.5% were obtained, with a window size of 4000 samples and the best set of features: mean, autocorrelation, root mean square and MFCC. High accuracies are reached for all subjects. The subject-independent performance was also evaluated with K-Means Clustering Algorithm (Lloyd, 1982). A person-independent accuracy of 88.57% and standard deviation of 4.0% were obtained, with a window size of 4000 samples and the best set of features: mean, autocorrelation, root mean square and MFCC.

Table 3: Clustering Performance (mean value) as a function of different window length extracted from the best set of features.

Window Size	**Adjusted Rand Index (%)**
1000 samples	89.73% $\pm$0.4%
2000 samples	97.42% $\pm$0.9%
4000 samples	99.29% $\pm$0.5%

Table 3 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 HAR system reaches an accuracy between 89.73% $\pm$0.4% and 99.29% $\pm$0.5%, with 1000 and 4000 samples, respectively.

5.1 Classification-based Evaluation: Proposed Metric

A new metric for assessing the obtained results from unsupervised techniques, a classification-based

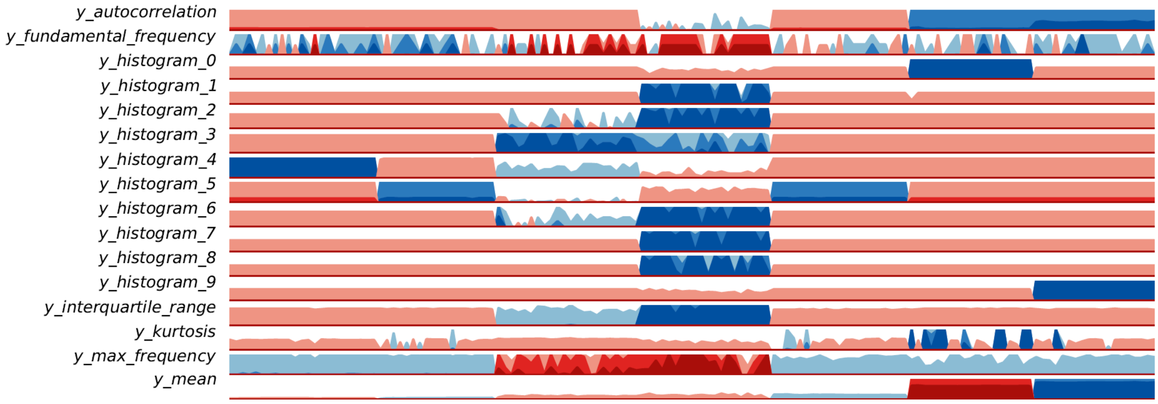


Figure 2: Horizon Graph - Time Series Visualization Technique.

evaluation metric, was developed. Initially, a confusion matrix that contains information about true and predicted labels done by a clustering method was constructed. Once the clustering algorithm randomly associates the clustering results to non-annotated groups, the **Algorithm**, *Best Cluster Permutation*, that links these groups to their corresponded activity, was implemented. The presented **Algorithm** receives the confusion matrix with a random assignment and goes through each row of the matrix and stores the index that contains the maximum value of each row.

Algorithm: `Best Cluster Permutation`
`Input: Input: confusion matrix with a random assignment.`
`Output: confusion matrix correctly assigned.`

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 swap vector, the matrix with the labels already associated is reconstructed. Table 4 shows the confusion matrix for this study where label i, with $i = \{1, 2, ...,7\}$, corresponds respectively to: standing, sitting, walking, running, lying down (belly up), lying down (right side down) and lying down (left side down). For the concatenated data, the algorithm successfully distinguish all activities.

6 CONCLUSIONS AND FUTURE WORK

The continuously need to obtain more information, more efficient, more quickly and with less intervention from an expert has led to a growing application of signal processing techniques to motion data. During the experiment, acceleration signals were collected from a waist mounted accelerometer based framework. In the presented study, a methodology to search for the best features able to classify different physical activities was presented. 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. The obtained results in clustering accuracy of HAR were very encouraging: an average person-dependent ARI (Santos and Embrechts, 2009) of 99.29% and a person-independent ARI of 88.57% were reached. The major achievements of the current work, compared to the state of the art are: the presented study performs tests in intra and inter subject context; a set of 180 features was implemented, which are easily selected to test different groups of subjects and different activities and the implemented algorithm does not stipulate, a priori, any value for window length of the signal or overlap percentage, 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 was also proposed. The presented research leaves a few opened questions, to be explored in the future:

- **Bigger Timespan**: Week Long Acquisitions of Movement.
- **More Data**: Increase the Number of Subjects and Applications.
- **More Computing Power**: Use Parallel Computing Infrastructures on the Data Collected.
- **More Discoveries**: Detect the Behaviour changes and annotate those changes.

Table 4: Confusion Matrix, in percentage, for concatenated data, where Lying Down$^{(1)}$ is lying down (belly up), Lying Down$^{(2)}$ is lying down (right side down) and Lying Down$^{(3)}$ is lying down (left side down).

	Standing	**Sitting**	**Walking**	**Running**	**Lying Down$^{(1)}$**	**Lying Down$^{(2)}$**	**Lying Down$^{(3)}$**
Standing	**92.1**±**3.2**	0.0±0.0	0.0±0.0	0.0±0.0	5.4± 2.3	1.3±0.9	1.1±0.8
Sitting	28.3±6.9	**68.0**±**5.9**	1.1±0.6	0.3±0.7	0.1± 0.3	1.6±0.7	0.6±1.3
Walking	0.0±0.0	0.4±0.5	**99.5**±**0.5**	0.1±0.3	0.0±0.0	0.0±0.0	0.0±0.0
Running	0.0±0.0	0.0±0.0	0.3±0.4	**99.4**±**0.7**	0.3±0.4	0.1±0.3	0.0±0.0
Lying Down$^{(1)}$	0.9±0.6	2.0±1.1	0.1±0.3	0.0±0.0	**82.1**±**1.9**	7.5±1.4	7.4±1.3
Lying Down$^{(2)}$	0.0±0.0	0.0±0.0	0.1±0.3	0.5±1.0	1.1±0.0	**90.4**±**0.9**	8.0±1.3
Lying Down$^{(3)}$	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.1±0.3	0.4±0.5	**99.5**±**0.5**

In the future, this framework should be tested on other intensity varying activities and across more subjects. For example, test it on individuals running and walking at a greater range of intensity levels. ACC data obtained from wearable accelerometers can be synchronized with the activity of daily living data recorded by such monitoring systems to better describe the information of human mobility, behavioural pattern and functional ability that encompass the important parameters regarding the overall health status of an individual.

The main challenge for future work in this area will be the development of features and recognition strategies that can work in an ambient assisted living under a wide variety of environmental conditions.

REFERENCES

Ankerst, M., Breunig, M. M., Kriegel, H.-P., and Sander, J. (1999). Optics: ordering points to identify the clustering structure. *SIGMOD Rec.*, 28(2):49–60.

Bandini, L. G., Gleason, J., Curtin, C., Lividini, K., Anderson, S. E., Cermak, S. A., Maslin, M., and Must, A. (2013). Comparison of physical activity between children with autism spectrum disorders and typically developing children. *Autism*, 17(1):44–54.

Bouten, C. V., Koekkoek, K. T., Verduin, M., Kodde, R., and Janssen, J. D. (1997). A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. *IEEE Transactions on Biomedical Engineering*, 44(3):136–147.

Bugané, F., Benedetti, M. G., Casadio, G., Attala, S., Biagi, F., Manca, M., and Leardini, A. (2012). Estimation of spatial-temporal gait parameters in level walking based on a single accelerometer: Validation on normal subjects by standard gait analysis. *Comput. Methods Prog. Biomed.*, 108(1):129–137.

Crockford, D. (2006). The application/json media type for javascript object notation (json). RFC 4627, IETF.

Donalek, C. (2011). Supervised and Unsupervised Learning. (April).

Ester, M., peter Kriegel, H., S, J., and Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. pages 226–231. AAAI Press.

Foerster, F. and Fahrenberg, J. (2000). *Motion pattern and posture: correctly assessed by calibrated accelerometers*, page 32(3):450457.

Frey, B. J. and Dueck, D. (2007). Clustering by passing messages between data points. *Science*, 315(5814):972–976.

Gomes, R., Nunes, N., Sousa, J., and Gamboa, H. (2012). Long term biosignals visualization and processing. In *BIOSIGNALS*, pages 402–405.

Heer, J., Kong, N., and Agrawala, M. (2009). 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)*.

I. H. Witten, E. Frank and Hall, M. A. (2011). *Data Mining: Practical Machine Learning Tools and Techniques*.

Khan, A. M., Lee, Y.-K., Lee, S. Y., and Kim, T.-S. (2010). A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer. *Trans. Info. Tech. Biomed.*, 14(5):1166–1172.

Liao, L., Fox, D., and Kautz, H. (2005). Location-based activity recognition using relational markov networks. In *Proceedings of the 19th international joint conference on Artificial intelligence*, IJCAI'05, pages 773–778, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Liu, A. D. (2013). *THE EFFECT OF SENSOR MASS, SENSOR LOCATION, AND DELAMINATION LOCATION OF*. PhD thesis, California Polytechnic State University.

Lloyd, S. P. (1982). Least squares quantization in pcm. *IEEE Transactions on Information Theory*, 28:129–137.

Mannini, A., Intille, S. S., Rosenberger, M., Sabatini, A. M., and Haskell, W. (2013). Activity recognition using a single accelerometer placed at the wrist or ankle. *Med Sci Sports Exerc*.

Mannini, A. and Sabatini, A. (2010). Machine learning methods for classifying human physical activities from on-body sensors.

Mathie, M. (2003). Monitoring and interpreting human movement patterns using a triaxial accelerometer.

Oliphant, T. (2006). Guide to Numpy. Tregol Publishing.

Palmerini, L., Mellone, S., Avanzolini, G., Valzania, F., and Chiari, L. (2013). Classification of early-mild subjects with parkinson's disease by using sensor-based measures of posture, gait, and transitions. In *AIME*, pages 176–180.

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

Phillips, S. M. and McAuley, E. (2013). Physical activity and fatigue in breast cancer survivors: A panel model examining the role of self-efficacy and depression. *Cancer Epidemiology Biomarkers & Prevention*, 22(5):773–781.

Pour, G., Troped, P., and Evans, J. (2013). Environment feature extraction and classification for context aware physical activity monitoring. In *Sensors Applications Symposium (SAS), 2013 IEEE*, pages 123–128.

Santos, J. M. and Embrechts, M. (2009). On the use of the adjusted rand index as a metric for evaluating supervised classification. In Alippi, C., Polycarpou, M. M., Panayiotou, C. G., and Ellinas, G., editors, *ICANN (2)*, volume 5769 of *Lecture Notes in Computer Science*, pages 175–184. Springer.

Simon, B. J., Errico, J. P., and Raffle, J. T. (2013). Noninvasive magnetic or electrical nerve stimulation to treat or prevent autism spectrum disorders and other disorders of psychological development. US Patent App. 13/783,319.

Smeja, J. F. F. F. M. and Muller, W. (1997). Assessment of posture and motion by multichannel piezoresistive accelerometer recordings. *Psychophysiology*, page 34(5):607612.

Song, M. R., Lee, Y.-S., Baek, J.-D., and Miller, M. (2012). Physical activity status in adults with depression in the national health and nutrition examination survey, 2005–2006. *Public Health Nursing*, 29(3):208–217.

Wren, C. R. and Tapia, E. M. (2006). Toward scalable activity recognition for sensor networks. In *In Lecture Notes in Computer Science*, pages 168–185. Springer.

Z. Ghahramani (2004). Unsupervised learning. *Gatsby Computational Neuroscience Unit, University College London*.

Zheng, Y., Wong, W.-K., Guan, X., and Trost, S. (2013). Physical activity recognition from accelerometer data using a multi-scale ensemble method.