Integrated Accelerometry-Based System for Functional Mobility Assessment and Fall Detection

Cristina Soaz González

Integrated Accelerometry-Based System for Functional Mobility Assessment and Fall Detection

Cristina Soaz González

To my *Doktorvater*.

Abstract

Miniaturized modern accelerometers offer the possibility to unobtrusively monitor the functional mobility status of patients for prolonged periods of time, both in clinical settings and in uncontrolled environments. Monitoring of balance and gait quality is of special interest in applications such as medical diagnostics, assessment of the risk of falling, and evaluating patient response to specific rehabilitation programs. Another advantage of this technology is its ability to work simultaneously as a fall detector.

The focus of this thesis is the development and validation of algorithms and methods for gait and postural stability analysis, fall risk assessment, and fall detection using a single waist-worn accelerometer. In this work, the analysis of gait consists of the automatic detection of steps and their parameterization. The step detection algorithm is divided in three parts: recognition of regions of interest, estimation of initial contacts, and template matching. The parameterization of detected steps is carried out by identifying fiducial points. These points indicate relevant changes in the gait phases with the potential to determine underlying walking dysfunctions. The analysis of postural stability is based on the measurement of the center of mass displacement. Here, two methods to transform the acceleration changes into displacement are discussed and validated using gold standards.

For fall risk assessment, a data collection protocol and a prediction model were built. The patient data contain clinical and functional test records of prospective fallers and non-fallers, which are the labels for the classes used in the prediction. The model accounts for the challenge of class imbalance, large class overlapping, and high dimensionality of the data. The results show that prediction is considerably accurate, and that acceleration-derived features improve the predictive value more than conventional features do.

Lastly, the fall detection algorithm of this work is based on detection of an impact and the subsequent body orientation in relation to the ground. The algorithm was tested on simulated falls, real falls, and real long-term physical activity records. The results show high sensitivity and the lowest false alarm rate of the algorithms published to date.

Acknowledgement

Contents

Contents

1. Introduction

1.1. Vision

The vision of this research is to provide an accelerometry-based integrated system for the assessment and monitoring of physical functional status and detection of adverse events, particularly in persons with impaired gait and frail elderly living alone. The major aims are:

- Assessment of the gait and balance quality,

- Early warning of an increase in the risk of falling,

- Outcome monitoring in exercise therapy programs or clinical interventions using functional tests and long-term measurements of physical activity (in a clinic or at home), and

- Automatic fall detection with alarm generation.

A system of this nature is intended to promote a healthy lifestyle, improve healthcare quality, and facilitate independent living by: giving feedback to the users about their functional status and setting goals; faciliting the practice of evidence-based medicine for the clinicians by delivering timely, objective patient-specific information; and providing peace of mind to people who live alone as well as to their families.

The work contained in this thesis provides the foundations for attaining the aforenamed objectives, including selection of monitor type, suitable sensor location, methodology, and, in particular, algorithms for extracting significant clinical parameters and fall detection.

1.2. Motivation

1.2.1. Sustainability of the Health Care System

Healthcare costs associated with the elderly and people with common chronic diseases - including cardiovascular disease, cancer, stroke, and diabetes - is a rising concern for the sustainability of the healthcare system, especially in developed countries where life expectancy is estimated to grow significantly in upcoming years [47] and where the top four causes of death are related to cardiovascular disease [4].

The main economic costs in the case of the elderly are comprised of admission to a nursing home or hospitalization and long-term rehabilitation as a result of injurious falls. Moreover, in most cases the elderly never fully recover from the injury, which has profound consequences for their independence, psychological health, and overall quality of life [9].

In persons with chronic disease, direct costs include payments for health services and medications, while indirect costs come from loss of productivity as a result of absenteeism and inability to work as well as the cost of household members caring for those who are ill [144].

Some of the factors that put us at risk for these chronic diseases are beyond our control, such the age and genetic make-up. But the risk factors irrespective of inherited conditions that explain the vast majority of chronic disease deaths are indeed preventable and modifiable. These common risks include: tobacco use, unhealthy diet, and physical inactivity. Experts agree that changes towards a healthy dynamic lifestyle reduce the risk of disease and injury, and that in order to sustain the current healthcare system there is an urgent need to shift the focus from

reactive hospital-based care to long-term, proactive, home-based care [8]. This new care model, focused on promoting an independent and active life style, also demands new technologies that will permit early identification and monitoring of populations at risk of functional decline and will help persons with walking difficulties to live autonomously at home.

1.2.2. Limitations in Current Methods of Functional Assessment

Functional assessment (or functional mobility assessment) is an examination of an individual's mobility, transfer skills, and activities of daily living [1]. It is usually performed by a physiotherapist for community-dwelling patients irrespective of medical diagnosis and consists of an evaluation of some standardized gait and balance tests. Examples of common functional tests include the 10-meter walk test (time to cover 10 m in distance), Timed Up and Go test (time needed to stand up from a chair, walk 3 meters, come back, and sit down in the chair), standing on one leg for a specific number of seconds. In clinical settings, poor scores typically trigger functional interventions like muscle strengthening, gait and balance training, or aerobic exercise [150]. These tests have been shown to constitute a valuable instrument in identifying persons at high risk of falling and evaluating therapeutic interventions [50, 135, 40]. However there are some shortcomings in the traditional outcomes and methods used to conduct such tests:

Not objective   In some cases, the assessment is self-reported, carried out via questionnaires. In cases involving a large undifferentiated mid-range score in which most individuals are distributed, a patient's functional status must be subjectively judged by the clinician.

Floor and ceiling effects   In clinical settings, balance tests are commonly judged according to a binary outcome ("success" or "failure"), which merely identifies patients at very high or very low risk of functional decline. Deterioration in a group with bad performance will not be detected, nor will improvement in a group with good performance.

Impractical equipment   Equipment used in the gait laboratory, like force plates or 3D infra-red camera systems, are objective, precise, and overcome the aforementioned drawbacks, but due to their complexity, non-portability, and high cost they become impractical for use as clinical screening tools.

Time-consuming   The duration of some tests is too long for the resource and time constraints imposed in clinical settings, which does not allow for a comprehensive functional analysis.

Poor ecological validity   Ecological validity refers to the extent to which the findings of a research study can be generalized in real-life settings. Measurements in a clinic are usually affected by the "white-coat effect," a state of anxiety in response to the presence of a health provider, usually clad in a white coat, which may alter various measurements. Long-term monitoring and functional assessment at home would provide more reliable results.

No identification of underlying dysfunction   The primary purpose of functional tests is to identify the presence of gait or balance abnormalities, and not to establish the cause of the problem. Assessment tools that differentiate among types and reasons of functional anomalies could help to determine more effective treatment for a specific disorder.

1. Introduction

Burdensome data management   In general, rehabilitation centers store and update patient information and test outcome monitoring using spreadsheets or non-standardized forms. In this context, an integrated tool providing automatic data storage and simple management would considerably decrease the workload for the clinicians.

1.2.3. Challenges of Fall Detection Monitors

Various methods for automatically detecting falls have been developed over past years, comprised of video cameras [84, 11, 110], floor vibrations sensors [116, 204, 10], pressure mats [12], and inertial sensors [111, 26, 95, 114]. Context-aware systems based on video analysis, floor vibration recognition, or pressure sensors usually require installation work, they are costly and cumbersome, raise concerns in users about being observed, must be placed at specific locations on the floor to achieve acceptable accuracy, or they just do not work when the person is not at home. In contrast, inertial sensors offer advantages in terms of cost, size, weight, unobtrusiveness, and portability in particular. With wearable sensors, data collection is no longer confined to a laboratory environment, thus leading to ubiquitous health monitoring [15]. Among wearable sensors, waist-worn accelerometers seem to be the optimal option for monitoring falls [115, 6, 93]. Recently, Bagala et al. [15] examined real falls to determine the performance of 13 fall detection algorithms for waist-worn monitors, and although the overall specificity and sensitivity turned out to be worse than the results reported by the original authors using simulated falls, they still presented a good trade-off, with up to 83% sensitivity and 97% specificity in the best-case performance. However, the great weakness was found in the amount of false alarms. None of the reviewed algorithms achieved a false alarm rate lower than 4 false alarms per day on average for the various participants involved in the study. In summary, wearable sensors appear superior to context-aware systems, but they still present some challenges:

Performance under real-life conditions   Most of the research on fall detection has been performed using data from simulated falls recreated by young people and according to distinct protocols [86]. This methodology does not appear to have high ecological validity in light of recent findings showing that fall detection algorithms present lower accuracy values when run on real falls compared to data generated in controlled settings [15]. However, this methodology is still in use, the main reason being that although falls occurs very often - especially in elderly populations - capturing real-world falls is extremely difficult [15]. One of the major challenges in this context is the creation of a standardized procedure and a public real falls database including subjects with different ages and clinical history.

Reduction of False Alarm Rate   Although several commercial products are available on the market, the reality is that fall alarms are not widely used and they do not have an actual impact on seniors' lives yet [86]. In some cases, potential users are simply unaware of their existence, while in other cases users reject wearing them after having tried the alarms. The major reason for rejection, by the wearer and also by remote monitoring centers, is the high number of false alarms, which result in inappropriate alerts [140, 86]. False alarms rate, although it seems to be a good indicator of the performance of a fall detection algorithm, it is still not so widely used in the literature as benchmark, but the specificity. The usefulness of this last parameter is limited because it does not provide any information about the prevalence of false alarms. In order to validate whether a specific algorithm would be applicable to a real scenario, it would be recommendable to also provide the value of the false alarm rate under real-life conditions.

Improvement of Usability and Acceptance   A key factor for success in a fall detection monitor is the level of acceptance to wear it, which is closely linked to the usability of the sensors. Here,

usability means the extent to which the fall monitor can be used by a specific user to detect falls with effectiveness, efficiency, and satisfaction [7]. It mainly depends on the characteristics and position of the sensor. For example, a sensor placed on the back of the neck may detect falls efficiently, but the obtrusiveness and the discomfort when wearing it limits the usability of the device. Another aspect consists of battery life, i.e. the need to recharge the battery very often may compromise the user's satisfaction. And, lastly, the user interface, which needs to adapt to user needs. Fall alarms are mainly intended for monitoring chronically ill patients and the elderly. Thus, the user interface should be very simple and intuitive.

If the user is satisfied with a system's usability, then the chances for acceptance will be very high. Yet there are further factors that may still influence acceptance, like the high cost of some monitors [90] or stigmatization of the fragility of the old person [140]. Making fall alarm systems affordable and discrete would increase acceptance considerably.

1.3. System Requirements

To ensure the practical utility of the system, this one should be comprised of only a single monitor and, depending on the application, a display device like a smart phone, smart watch, or tablet. Naturally, placing more sensors across the body will generate more information, but it is highly likely that doing so will increase complexity and cost significantly, and usability will suffer. As far as the algorithmic aspect is concerned, there are no critical requirements in terms of response time and memory capacity, except for the fall detector, in which case the algorithm should run on the monitor's microcontroller.

1.3.1. Monitor Requirements

In mobile healthcare systems, the most important part is the monitor, as it is the basis for clinical diagnosis. In this case, the monitor consists of a wearable triaxial acceleration sensor. The prerequisites that the sensor should fulfill for the proposed system are the following:

Unobtrusive and portable   In pervasive sensing and monitoring, the appearance of the sensor is an important factor. The presence of the system should be undetectable in order to avoid stigmas or excessively draw attention from others. It should be also sufficiently small and lightweight to be carried for prolonged periods of time, optionally beneath clothing.

Long battery life   The battery should run for at least two to four weeks under continuous use (24/7) during the day as well as during sleep, as frequent battery changes would likely hinder user acceptance [104] or endanger the reliability of the system. High power batteries may be quite large. Thus, to not compromise the size of the device, the sensor has to be extremely power efficient.

Security   The stored and transmitted acceleration data should be encrypted and/or hashed to ensure user privacy and data integrity.

Sampling frequency and amplitude range   In order to assess daily physical activity and detect falls, the accelerometer must be able to measure accelerations up to $\pm 6g$ when attached at waist level, and frequencies between 0 and 20 Hz [25, 172].

Wireless communication   The sensor should be equipped with an embedded wireless module that allows connectivity with a smart phone or similar mobile technology to send an alert in case of adverse events, such falls, and for transmission of raw data or relevant health status information.

Adequate computational capability and memory capacity   The computational capabilities of the sensor are linked to computational power, the price and, to some extent, to the size. If we want to make the sensor accessible to as many people as possible, we will need to use low-cost, low-power, miniaturized acceleration sensors. Although computational capability and memory capacity are limited in this context, the device should at least support the execution of un-demanding algorithms for fall detection applications. Internal storage should also be big enough to store at least a weeklong acceleration recording.

Robust   The sensor must endure everyday use in low(< 0 °C) and high (up to 45 °C) temper-ature environments. Ideally, the device should also be waterproof so that it can be used during showering, a time of high risk of falling.

Easy to use   The sensor should be attached to the waist with a comfortable, hypoallergenic attachment that non-irritating to the skin. The interface should be as simple and intuitive as possible. The device must be easy to handle by elderly or persons with mobility impairment.

1.3.2. Algorithm Requirements

The prerequisites for the algorithms differ depending on the application:

Fall Detection

Fall detection algorithms usually consist of detecting an impact, checking body orientation for a specific time and, depending on some threshold-based parameters, deciding whether the event resembles a fall [15], in which case the activation of an alarm should be triggered. More sophis-ticated algorithms may use pattern recognition techniques like support vector machines [159]. Regardless of the algorithm rules, the fall detection algorithm should satisfy the following:

Short response time   It is not necessary for the algorithm to run in real-time, but as soon as all necessary raw data for the decision-making are collected, the output should be processed in a matter of milliseconds.

Efficient   The algorithm will run in a low-cost, low-power consumption microcontroller with some computational and memory constraints, thus, the computational load should be minimized.

Functional Assessment

For functional assessment purposes, the patient's raw data are transmitted via wireless or USB to a medical server where more computational power and storage capacity are available, which means no big restrictions will be imposed in terms of resource efficiency. Although there are also no tough restrictions as concerns response time, nested conditions and loops should be avoided, in particular if we take into account that the input data may be considerably big ($\approx$ 200 MBytes for a weeklong record).

1.4. Related Work

This chapter presents the previous work done in the field of accelerometry-based clinical applicati-ons for gait and balance analysis, fall risk assessment, and fall detection. Each section summarizes state-of-the-art systems, algorithms, and models in terms of their capabilities and limitations.

1.4.1. Functional Assessment Review

Functional mobility assessment comprises two main evaluations, the first being analysis of walking quality, and the second being analysis of the subject's balance skills.

The first thing to do in order to analyze walking is to develop the capacity to distinguish the elemental units of the gait cycle, i.e. the steps. There have already been several studies that have investigated the potential of inertial devices to recognize steps, but since this field is still relatively new, there are no widespread standard algorithms or methods. Table 1.1 presents a list with the most relevant studies on step detection using a single waist-worn accelerometer. These publications have been retrieved through a systematic review explained in Appendix A. The major critical remarks are indicated in red color.

Publication	Reference	Validation Metrics	Remarks
Zijlstra, 2003 [207]	treadmill with force sensors	$\max(\epsilon) = 0.103$ s, $\min(\epsilon) = 0.002$ s	-
Dijkstra, 2008 [49]	video	OA: $\bar{\epsilon} = 7.4 \pm 3.1$, PD: $\bar{\epsilon} = 6.9 \pm 3.0$	-
Umemura, 2008 [184]	not specified	$\max(\eta) = 5.6$, $\min(\eta) = 0.0$	no HS recognition
Marschollek, 2008 [122]	video	$\max(\eta) = 60.0$, $\min(\eta) = 8.4$	no HS recognition
Lee Hyo-Ki, 2009 [109]	optical camera	$\max(R) = 0.99$, $\min(R) = 0.92$	no HS recognition
Lee Hyo-Ki, 2010 [108]	optical camera	$\max(\bar{\epsilon}_c) = 0.27 \pm 0.24$, $\min(\bar{\epsilon}_c) = 0.21 \pm 0.21$	no HS recognition
Foglyano, 2011 [61]	optical camera	left $D = 99 \pm 56$ms, right $D = 25 \pm 83$ms	only one patient

$\epsilon = |GRF - acceleration|$ is the absolute error of the foot contact instant
$\bar{\epsilon}[\%] = \frac{1}{n} \sum \frac{|video - acceleration|}{video} * 100$ is the mean absolute percentage error of number of steps
$\eta[\%] = \frac{|reference - acceleration|}{reference} * 100$ is the relative error of number of steps
R is the correlation coeficient of estimated step time and stride time with the reference
$\bar{\epsilon}_c$ is the mean cadence error measure in steps/minute
D is the mean delay in comparison to the true heel strike in miliseconds

Table 1.1.: Studies on step detection and postural stability using a single waist-worn triaxial accelerometer.

The principal limitations found in the aforenamed work concern the group of subjects employed in algorithm development, the validation method, and the type of event detected by the algorithm.

Only two authors, Dijkstra [49] and Marschollek [122], included a reasonable minimum amount of data from persons with impaired gait (e.g., frail seniors, Parkinsonian patients) in their study. Lee et al. [109, 108] included data from hemiplegic patients and children with cerebral palsy, however the amount of subjects was low (6 and 8 subjects, respectively). The same occurred in Unemura's [184] study with only 3 young subjects. Among the publications that have been validated in an elderly population [122, 49], none reported suitable results for low speeds. The relative errors in the number of steps detected for this population varied from 28.1% to 62.1%. The more customized the algorithms were to specific walking patterns, the worse they performed on different samples [122]. Another remark is that only Zijlstra [207] reported detailed information about the accuracy of the algorithm relative to the walking speed, although as a negative aspect, he restricted his study to healthy subjects.

With the exception of Dijkstra [49] (who employed the algorithm developed by Zijlstra), all authors used autovalidation in their studies. This usually leads to an overestimation in algorithm accuracy in those cases where an algorithm is tuned according to parameters derived from a specific group of subjects, especially when the database is small.

Concerning the type of event detected by the algorithm, most of the publications included in Table 1.1 detect steps by identifying a peak (usually a maximum) in the raw acceleration signal. This is fine for counting steps, particularly for monitoring physical activity levels. However, to

analyze gait quality in terms of gait asymmetry, for example, it is necessary to recognize a step as such, i.e. from the moment in which the heel of one leg contacts the ground until the next consecutive heel strike of the other leg. Zijlstra and Foglyano [207, 61] were the only ones who developed algorithms to specifically detect these particular events.

Finally, the lack of standards is reflected by the variety of validation measures and reference tools to quantify the performance of step detection algorithms. Naturally, the absence of a benchmark makes the comparison of results exceptionally difficult.

1.4.2. Fall Risk Assessment Review

Fall risk assessment is an important tool for prevention. It helps to determine the most appropriate interventions and to reduce or eliminate falls. Over the last decade, wearable sensors have been used to capture and analyze objective mobility data with the aim of improving fall risk assessment.

Howcroft et al. [82] recently published a review of studies employing inertial sensors to assess the risk of falling in geriatric populations. They provided a critical and comprehensive examination about study methodologies, sensor locations, analyzed variables, and validation methods. In the list of most relevant studies, some limitations have been identified and emphasized using red in Table 1.2. In essence, these limitations consist of incorrectly labeled data, class imbalance, and validation methods.

Concerning class labels, the main weakness found in the previous work is that most publications classify "fallers" as persons at high risk of falling (identified using the traditional fall risk scores) or persons who are retrospective fallers, as opposed to using prospective fallers. For example, Giansanti [67] and Gietzelt [67] used the Tinetti and STRATIFY score respectively as the criterion for dividing the dataset into "fallers"and"non-fallers." Both achieved good prediction performance because the outcome obtained with the acceleration-based data from the functional tests are similar to the traditional functional scores. Thus, they did not predict actual falls, but rather the respective scores. Another scale commonly used as a gold standard for identifying potential fallers in clinical settings is the Tinetti score, even though it has been shown to have some limitations, like possible ceiling effects [157].

Caby [31], Greene [69] and Kojima [99] used retrospective fallers to predict the risk of falling, leading to fallacious results because what they "predict" is the likelihood that a person has fallen and not the likelihood that a person will fall. Also, classifying retrospective fallers is easier than classifying prospective ones since previous falls tend to deteriorate a subject's functional performance substantially, especially when the fall occurred within the recent past and when the person suffered a fracture as a consequence of the fall.

Another limitation that may affect the accuracy of the fall risk assessment model is the imbalanced size of the classes. This is related to the fact that, given a similar group of individuals, the number of subjects expected to fall is 3 times smaller than the non-fallers [59, 94, 143]. Taking that into consideration, Caby [31], König [100] and Marschollek [121] used rather balanced datasets. Weiss et al. [190] even used more fallers than non-fallers. In their study, it seems that the authors included a predefined quantity of healthy non-fallers after having collected a sample of fallers. While this may improve predictability, it does so at the expense of distorting the authenticity of the class sizes. Greene [69] included also more fallers than non-fallers. The reason for this being that he and his colleagues looked at a 5-year falling history, exchanging prediction accuracy for temporal accuracy in doing so.

Other remarks related to selection of the database are the following: Weiss [190] applied multiple exclusion criteria to improve prediction at the cost of narrowing the generality of the models, and Caby [31] edited 10% of class labels to line up with an expert's opinion. Proceeding in this manner can considerably improve results, especially when involving so few subjects (20 subjects).

Publication	Validation	TPR	TNR	ACC	Remarks
Caby, 2011 [31]	LOOCV	1	1	1	only retrospective falls, 20 subjects, hospital inpatients, edited class labels
Doi, 2013 [52]	not specified	0.69	0.84	0.81	–
Giansanti, 2008 [67]	fixed split	0.98	0.97	0.97	Class label: Tinetti score
Greene, 2010 [69]	10-fold CV	0.77	0.76	0.77	only retrospective 5-year fall history
Gietzelt, 2009 [68]	not specified	0.89	0.91	0. 90	Class label: STRATIFY score
König, 2014 [100]	not specified	0.74	0.76	0.75	–
Kojima, 2008 [99]	not specified	0.61	0.68	0.62	only retrospective falls
Marschollek, 2011 [121]	10-fold CV	0.58	0.96	0.80	hospital inpatients
Weiss, 2011 [190]	not specified	0.91	0.83	0.88	41 subjects, fallers are majority, multiple fallers, no p-value correction

TPR is the True Positive Rate
TNR is the True Negative Rate
ACC is the Accuracy

Table 1.2.: Best classification results of authors performing fall risk assessment with inertial sensors in geriatric populations [14].

Finally, many authors do not state anything about their validation method, as seen in Table 1.2, which leads us to assume that they simply used auto-validation. This validation does not show overfitting and can therefore extremely overestimate the predictive power of a model. They also do not always include p-value correction. When exploring the usefulness of a set of variables, the family-wise error rate has to be taken into account and countered by a suitable control method, like Bonferroni correction.

1.4.3. Fall Detection Review

Falls are the main cause of morbidity and mortality among elderly, also representing a major concern among persons with mobility impairment. Automatic detection and communication of falls may initiate prompt medical assistance, thus increasing the sense of security these persons have and reducing some of the adverse consequences of falls. Several algorithms have been developed to detect falls by analyzing the acceleration patterns of the individuals recorded with a body-worn inertial sensor. Table 1.3 comprises the most relevant fall detection algorithms published that have been independently evaluated under real-life conditions [15]. Main criticisms have been marked in red.

These algorithms use the magnitude of the vector given by the three acceleration axes and compare it to a certain threshold. Based on that comparison and on previous/later states, such as changes in body orientation, they decide whether the event resembles a fall or not. All of them performed worse in real-life conditions than in simulated falls. Nevertheless, one positive aspect is that they have low computational cost and low complexity, which makes them suitable for implementation in microcontrollers for real-time applications.

Chen's algorithm [35] uses a high threshold for the comparison of the impact magnitude reducing the number of false positives at the cost of missing falls with low peak values. The authors did not provide any value for the specific angle change that constitutes a change in body orientation. In their independent evaluation, Bagalà et al. [15] set this angle to the value for which they achieved the optimal specificity and sensitivity but the results did not discriminate well between real-world falls and ADL.

Publication	SEN	SPE	FP [average per day]	Remarks
Chen, 20016 [35]	0.76	0.94	8.3	fails with low magnitudes, high FP rate
Kangas, 2008 [92]	0.55	0.95	6.5	poor sensitivity, high FP rate
Bourke, 2007 [22]	1.00	0.19	60.8	poor specificity, very high FP rate
Bourke, 2008 [24]	0.72	0.87	19.8	high FP rate
Bourke, 2010 [23]	0.83	0.97	4.4	best result, fails with low magnitudes, still high FP rate

Table 1.3.: Sensitivity (SEN), specificity (SPE), and false positives (FP) of relevant falls detection algorithms using real-falls data from trunk-/waist-worn acceleration sensors.

Kangas et al. [92] investigated three different sets of algorithms with increasing complexity. The threshold values for the impact were low, which allowed detection of most of impacts. Thus, the real fall detection was made depending on the posture monitoring. They achieved high specificity at the expense of a low sensitivity. Table 1.3 shows the best result for the set of investigated algorithms.

Bourke's first algorithm [22] uses two thresholds, set according to the database to achieve 100% of sensitivity. A fall is detected if the magnitude vector is outside the range between those two values. However, the specificity obtained was considerably bad (SPE = 0.19). The reason for the poor specificity is that many ADLs (e.g. sitting on a chair or bed) show values for magnitude acceleration lower than the bottom threshold due to the phenomenon of weightlessness.

The second algorithm suggested by Bourke et al. [24] provides results comparable to Chen's algorithm. The difference with respect to their first algorithm is that this one provides posture monitoring after fall. Adding information about posture after impact improves results in terms of specificity. Yet the false alarm rate is still too high to turn the system into a commercial solution.

Finally, Bourke et al. [23] presented a third algorithm. This algorithm provides the best trade-off in terms of sensitivity (83%) and specificity (97%), although the results still differ from those obtained by the authors in their simulated fall database (100% SEN and SPE). Bagalà reported that the algorithm mostly fails to detect falls with low impact magnitude (especially forward falls, falling onto a bed/sofa, and against a wall).

1.5. Research Questions

The research questions that have guided this research are the following:

1 Can the use of a single waist-worn triaxial accelerometer overcome the limitations of the current methods for physical functional assessment? Can the use of the accelerometer-derived data improve fall risk assessment when compared to the results obtained using only conventional fall risk indicators in a group of seniors with osteoporosis?

2 Can we detect falls with a high sensitivity and low false alarm rate using a single waist-worn triaxial accelerometer?

1.6. Contributions

In this work, I have defined the constraints and requirements for an acceleration monitor and the requisite algorithms for assessing functional mobility and detecting falls.

I have collaborated with clinical partners to develop an original data acquisition protocol. As a quality assurance method in data collection, I have developed a test detection algorithm integrated on a dedicated server to give automatic feedback about correct administration of the test protocol when a user uploads data onto the system.

I have developed and validated new algorithms to extract parameters from an acceleration signal to provide relevant and comprehensible deductions about a subject's mobility status.

For gait assessment, I have created an algorithm that detects steps with high accuracy in healthy adults and seniors, even when walking at very low speeds, and with a considerably low false detection rate. The lack of standard metrics has rendered direct comparison of this work's results with similar studies difficult. However, my study does take into account two main limitations of previous works: the small size of the subject database and the validation methods. I have made a new contribution to the field by providing the first instance of a detailed step pattern analysis based on an acceleration signal recorded near the body's center of mass. This analysis includes parameterization of the signal by identifying fiducial points. Based on the characteristics of these points, I was able to determine differences between normal and impaired walking, as well as the potential underlying dysfunctions. For the assessment of balance, I created a specific type of scatter plot to estimate a subject's postural sway in a measurable manner, thus overcoming the ceiling and floor effects afflicting current methods. I also validated the accelerometer's capacity to measure body trunk displacement, something not yet accomplished in previous published studies.

Using the data collected as well as the aforementioned algorithms, among other items, I have extracted parameters to build a model for fall prediction. This model has shown that accelerometer-derived features improve fall prediction when compared to the results obtained using conventional features from the same subject database. These results are useful in predicting a person's risk of falling and perform exceptionally well in doing so, especially when taking into account the flawed characteristics of the data sets on which the model was set: large class overlap and imbalance. In contrast to most previous studies, this model successfully overcomes those pitfalls.

I have developed a fall detection algorithm that, to date, delivers the best performance in terms of false alarm rate when compared with other published algorithms tested on real data. My algorithm also achieved 100% sensitivity with data recorded under laboratory conditions, as well as successful detection of real-world falls.

I have been fully involved in the data collection employed for developing and validating the algorithms presented here. No synthetic data have been used in this thesis.

1.7. Outline

In continuation, the structure of the thesis starts with an introduction to the acceleration monitor and signal in Chapter 2. The order of the chapters to follow is congruent with the stages depicted in Figure 1.1, which represent the main applications of the proposed system.

In the event of evidence suggesting a person may be at risk of falling, an *assessment* should be performed regarding functional mobility status and other medical factors. Chapter 3 and Chapter 4 explain how functional status can be evaluated according to a subject's gait quality and balance ability through accelerometer-derived parameters. The algorithms and methods for extracting these parameters are presented in combination with the corresponding validation. The next step is to leverage this information to estimate the risk of falling using real patient data. To build a proper prediction model, an acquisition protocol of functional tests was set up for data collection. Both the model and the protocol are described in Chapter 5. Following assessment, the patients predicted as being at high risk of falling should ideally take part in a fall *prevention* program. This part of the procedure has been indicated in gray in Figure 1.1 because it mainly concerns physician work. Nevertheless, the system can still be used in this stage as an outcome monitoring tool. Finally, in the event the prevention program does not achieve the expected

results or the person suffers an adverse accident resulting in a fall, the system can automatically call for help if the user is unable to do it by himself/herself. Chapter 6 presents a fall *detection* algorithm and alarm system to that end. Finally, Chapter 7 lays out a summary of the principal thesis findings.
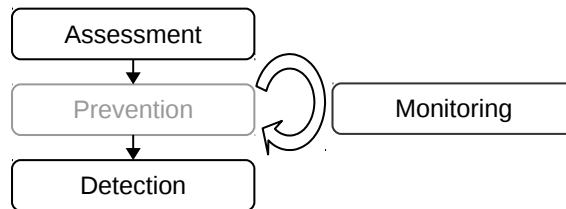
Figure 1.1.: Main applications of the proposed accelerometry-based system.

2. The Acceleration Sensor and Signal

2.1. Technical Characteristics

The acceleration sensor used in this thesis for the data collection and experimental set-up is the actibelt$^{\circledR}$ [43, 133], a triaxial accelerometer (ADXL 345 BCCZ Analog Devices) placed inside a belt buckle. The sensor (Fig. 2.1) has been developed at the Sylvia Lawry Centre for Multiple Sclerosis Research e.V. - The Human Motion Institute and Trium Analysis Online GmbH. The design is unobtrusive and ensures that the device is located close to the subject's center of mass. The acceleration sensor measures up to $\pm$ 6 g (g = 9.81 m/s/s) with a resolution of 0,0024g and sample frequency $f_s = 100Hz$. The power supply is provided by a rechargeable LiPo battery that lasts approximately one month and can be recharged via USB in 2 hours. The data are stored in an internal flash memory with a capacity of 512 MB allowing to record uninterruptedly for about 10 days. Data are encrypted at the time of download.
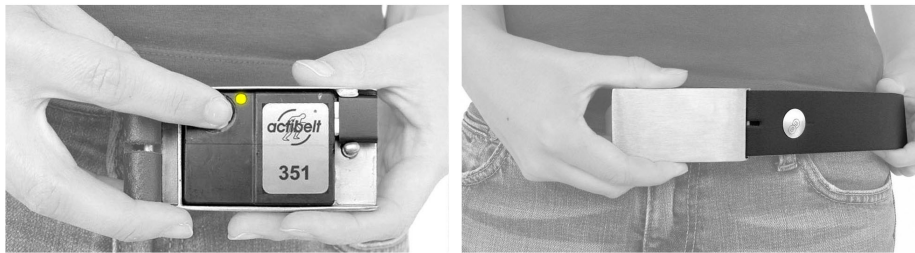
Figure 2.1.: Photos of the actibelt$^{\circledR}$ accelerometer used for the data acquisition. Courtesy of SLCMSR-The Human Motion Institute and Trium Analysis Online GmbH.

The sensor records static acceleration of gravity and dynamic accelerations relative to the trunk along the vertical, mediolateral, and anterioposterior directions. In anatomical position (see Fig. 2.2), the orientation of the accelerometer is the following: positive X values correspond to up acceleration (in red), positive Y values correspond to left acceleration (in green), and positive Z values correspond to anterior acceleration (in blue).

There is a wireless version of the sensor – the actibelt$^{\circledR}$-BLU – that includes a temperature and barometric pressure sensor (Bosch BMP085), a bluetooth$^{\circledR}$ module (LMX9838 National Semiconductors), and a 125 KHz RFID module. The wireless device is equipped with a micro USB port and a removable 8 GB micro SD card. The rest of technical characteristics are the same ones like in the non-wireless version. The wireless functionality allows connection to smart phones and tablets.

2.2. Sensor Position

The major factors that determine the ideal body location of activity monitors are the type of movement being studied and the wearer compliance. According to Yang [201], most studies adopted waist-placement for motion sensors because of the fact that the waist is close to the Center of Mass (CoM) of the human body. This implies that the accelerations measured by a single sensor at this location can better represent the major human motion. From an ergonomic
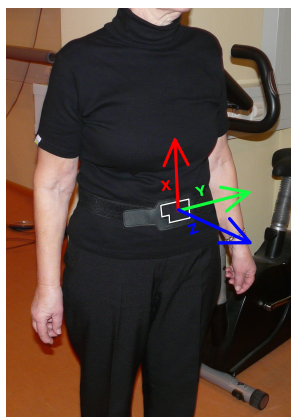
Figure 2.2.: actibelt® axis. Image adapted from [5]

point of view, the torso can better bear extra weight when carrying wearable devices. Sensors or devices can be easily attached to or detached from a belt around waist level. A range of basic daily activities, and also falls, can be recognized using acceleration patterns measured with a waist-worn accelerometer [95, 136, 51, 207]. Accelerometers can also be attached to the wrist, the thigh, or the ankle. Accelerometers attached to the thigh or ankle have been used to study gait-related parameters and leg movements [88, 62], and accelerometers attached to the wrist have been used to measure tremors in patients with Alzheimer's and Parkinson's disease [185]. But overall, the accuracy in activities recognition with accelerations sensors placed at these positions is lower than with accelerometers placed at the waist, as shown in Table 2.1.

Activity	Chest	Lower Back	Left Foot	Left Hip	Left Thigh	Left Wrist
Lying	1	1	0.997	1	0.972	0.967
Running	1	1	1	1	1	1
Sitting	0.966	0.992	0.924	1	0.972	0.966
Stairs down	0.940	0.920	0.915	0.935	0.925	0.926
Stairs up	0.928	0.906	0.920	0.929	0.929	0.902
Standing	0.969	0.993	0.929	1	1	1
Walking	0.981	0.973	1	0.990	1	0.961
Average	0.969	0.968	0.955	0.978	0.971	0.965

Table 2.1.: Balanced F-measure for each location, detailed by class, when using Neural Networks and a single 3-axial accelerometer, according to [201]. A higher F-measure value indicates improved detection of the investigated activity.

2.3. Sensor Calibration

In order to guarantee the validity of the measured acceleration data, the sensor must be calibrated regularly. This step is specially important when monitors are intended to be used over time, e.g. in intervention studies [192]. The procedure to calibrate the actibelt® consists in placing the sensor to a number of orientations with respect to gravity. In each position, only the gravitational component is measured by the accelerometer. Afterwards, this information is used to estimate the ± 1 g values rendering a ratio scale with 1 g as unity. This calibration procedure has been semi-automated by means of a fine crafted metal cube (Fig. 2.3) and a Java based client-server application. Each time the sensor is calibrated, the client sends a calibration log to the server via

web with information useful for the creation of statistics, like for example, actibelt$^{\circledR}$ and server time at the calibration instant, device identifiers, and calibration quality metrics.
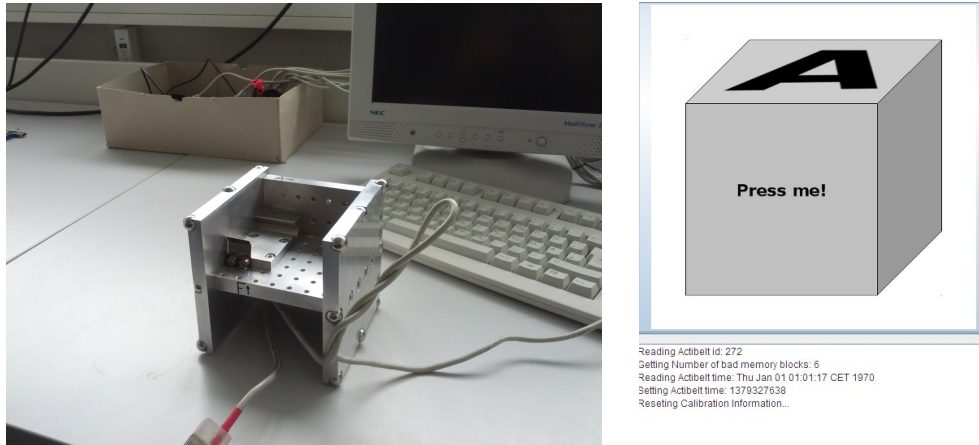
Figure 2.3.: Picture of the actibelt$^{\circledR}$ calibrator cube and screenshot of the wizard software. Photos courtesy of SLCMSR-The Human Motion Institute and Trium Analysis Online GmbH.

2.4. Acceleration Signal

The signal measured with the accelerometer can be decomposed in a *gravitational component*, a *movement component*, and *noise* as

$$\mathbf{a}_n = \mathbf{a}_{Gn} + \mathbf{a}_{Mn} + \eta_n, \tag{2.1}$$

where $\mathbf{a}_n = (a_n^x, a_n^y, a_n^z)$ is the total acceleration vector in the vertical (x), mediolateral (y) and sagittal (z) axis, $\mathbf{a}_{Gn} = (a_{Gn}^x, a_{Gn}^y, a_{Gn}^z)$ is the gravitational vector, $\mathbf{a}_{Mn} = (a_{Mn}^x, a_{Mn}^y, a_{Mn}^z)$ is the movement acceleration vector and η_n the noise component.

The sources of accelerometer *noise* are mechanical thermal noise and electrical thermal noise with the mechanical thermal noise being the dominant one above 10 kHz [113]. The high frequency noise can be removed or significantly diminished by low pass filtering the signal at a specific cut off frequency. Details about the denoising process are described in Section 4.2.4 and Section 6.3.1.

The *movement component* mainly comprises the acceleration forces originated by the body movements. Other extraneous signals, such as artifact due to soft tissue movement and vibrations imposed on the body when travelling in a motor vehicle, may also be present in the signal, and they can be minimized through careful instrument placement and signal filtering [125].

The *gravitational component* measures the inclination of the accelerometer. Assuming that the sensor is attached tight to the waist, the inclination depends on the body orientation and the body anatomy. The body orientation may change according to a particular physical activity or body posture and it is of special interest in applications such as activity recognition and fall detection. It can be estimated by filtering the gravitational component of the acceleration signal (for more details about filtering techniques refer to Section 6.3.1 and Section 4.2.1). Anatomical characteristics, such as body shape, can also produce an inclination of the sensor. In this case, we called sensor tilt. The sensor tilt causes a constant drift in the acceleration signal, undesired in certain situations, like for example, during the measurement of balance control in quiet standing [130].

2.4.1. Correction of Sensor Tilt

When standing upright or walking on a level surface, the accelerometer axis are pretty well aligned to the body axis for most healthy young and adults. But in the elderly, overweight people, or persons affected by spine curvature disorders, the difference between the body and the sensor axis may be no longer negligible.

For clarity, I will refer from now on to the *body axis* as the longitudinal, sagittal and transverse axes on a person's body standing upright and with the trunk straight, as in Figure 2.4. The body axes coincide with a horizontal-vertical coordinate system.
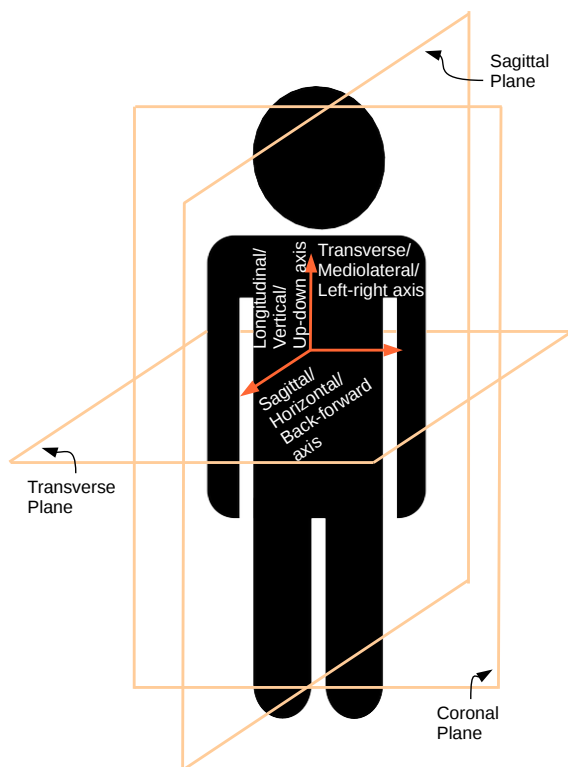
Figure 2.4.: Body axis and planes.

The effect of a misaligned axis can be observed in the center of mass (CoM) acceleration by looking to the signal offset. Usually for a healthy adult (Fig. 2.5a), the average value of the acceleration near the CoM in the left-right and backward-forward axis is close to zero and in the up-down axis it is near to -1 (without subtraction of gravity effect); whereas for a person with abnormal body posture, the sensor axis are shifted. In Figure 2.5a the backward-forward and the left-right axis are shifted upwards and downwards respectively, which indicates that the body leans forwards and rights.

With the correction of sensor tilt (Fig. 2.5c), the data is normalized by transforming them into a horizontal-vertical coordinate system. This correction is important to eliminate variability caused by other sources than the phenomenon being studied [129]. In order to do this, I adapted the formulas proposed by Moe-Nilssen [129] according to the direction of the acceleration along the actibelt$^{®}$ axis as follows:

Anterior-posterior direction

$$a_A = a_a \cos(\arcsin(\bar{a}_a)) + a_v \bar{a}_a$$

where a_A is the estimated anterioposterior acceleration, a_a the measured anterioposterior acceleration, and a_v the vertical measured acceleration.

Vertical direction

$$a_V = -a_m \bar{a}_m + a_V{}' \cos(\arcsin(\bar{a}_m))$$

where a_V is the estimated vertical acceleration and a_m is the measured mediolateral acceleration. $a_V{}'$ is an estimated provisional vertical acceleration equal to:
$a_V{}' = -a_a \bar{a}_a + a_v \cos(\arcsin(\bar{a}_a))$

Mediolateral direction

$$a_M = a_m \cos(\arcsin(\bar{a}_m)) + a_V{}' \bar{a}_m$$

When calculating the average of the measured accelerations it is important to use only the signal intervals during which the person is standing or walking in an upright position.
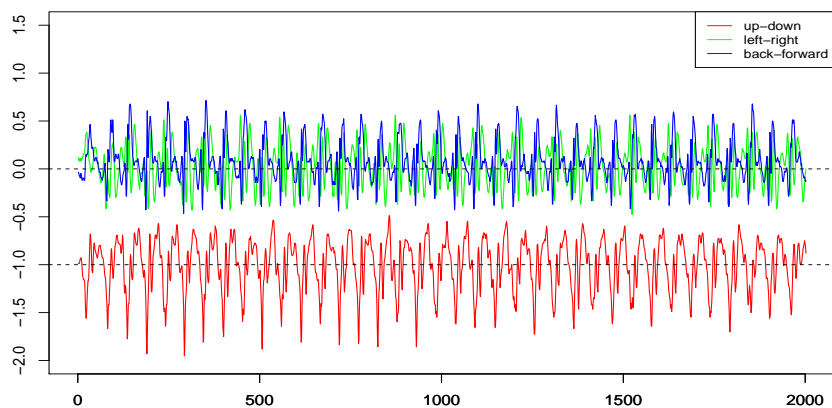
(a)
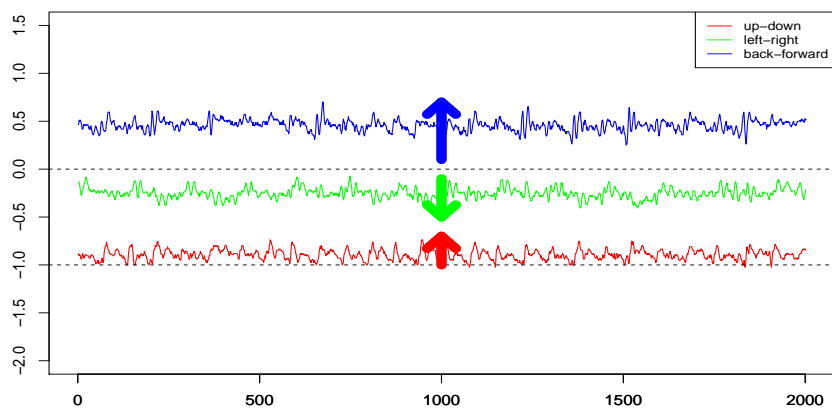
(b)
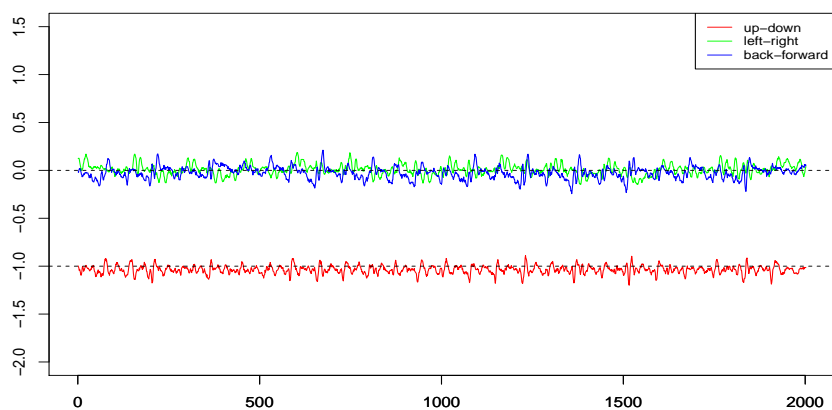
(c)

Figure 2.5.: 3D acceleration signal (in g units) recorded near the body's COM of: (a) a healthy adult and (b) an elderly female with spine curvature, both walking at self-selected speed. Subfigure (c) shows the signal in b) after tilt correction. The unit of the horizontal coordinates are samples (100 samples = 1 second).

3. Gait Analysis

3.1. Introduction

Gait ability is a primary outcome to assess functional mobility [166]. In clinical settings, gait is typically assessed by a physiotherapist via visual observation of patients, while in gait laboratories the assessment is done using sophisticated systems, such as infrared cameras and force plates. Recently, waist-worn accelerometers have been presented as a portable and reliable alternative to assess gait, as well as to monitor patients in rehabilitation therapy [165, 40, 135, 191]. They can be used in clinical settings and to record long-term acceleration data in the free-living environment for higher ecological validity. A fundamental measure of a person's ambulatory ability using accelerometers is step counting [102]. There exist numerous published studies using a waist-worn accelerometer to detect steps [207, 184, 109, 108, 85, 122, 61], but none of them has thoroughly investigated the performance of data from people with severely impaired gait; although it is known that many algorithms present poor accuracy in detecting the steps of those persons [122].

In this chapter, the aim is to develop an algorithm that detects steps accurately, both in healthy and in seniors with impaired walking, as well as to investigate the potential of the device to distinguish between normal and impaired walking via gait pattern analysis.

3.2. Gait Analysis Principles

Walking is a complex task involving nervous, somatosensory and musculoskeletal systems and most parts of the body. The contribution of each body segment is determined by the gait speed. Normal walking speed primarily involves the lower extremities with the arms and trunk providing stability and balance [168]. This limb movement produces acceleration changes near the Center of Mass (CoM) of the body distinctive for normal and pathological gait. The gait cycle is comprised of a complete stride, or equivalently a sequence of two steps. A step is defined as the interval between two consecutive heel strikes and is the elementary periodic signal measurable with the accelerometer near the CoM during walking. A step can be divided in the following gait phases (Fig. 3.1):

Initial contact (or heel strike)   The initial contact (IC) occurs in the first 0-2% period of the gait cycle when the foot first contacts the ground [151]. In this instant the vertical component of the ground reaction force (GRF), mainly responsible for the forces acting on the CoM, goes to a maximum of 120-150% of the body weight (BW) [152, 138]. During running, these forces can go even up to 500% of BW [83]. The position of the ankle joint at this instant determines the force response and therefore the way this IC manifests itself in the acceleration signal.

Loading response   The loading response at normal walking is a bipedal phase which occupies about 10% of the gait cycle, from the IC to toe off of the opposite leg. During the loading response, the foot comes into full contact with the floor, the knee flexes slightly absorbing some power in order to reduce the IC on the floor and the body weight is transferred onto the stance limb. A smooth correct transfer guarantees the stability of the upper body. It is in this phase where a running step can be distinguished from a walking step. If there is no bipedal stance the person no longer walks but runs. In the actibelt$^{®}$ acceleration signal, during running, the

acceleration in the vertical direction crosses the 0g line (g: the gravitational acceleration, 9.81 m/s2) twice, once when the body accelerates upwards along the positive direction of the vertical axis propelled by the calf muscle and again when it descends in free fall for a very short time right before reaching the ground.

Single Limb Support During this phase (10-50%), the body is supported by one single leg and the forces drop in average to about 60-80% of the BW. [151]. The body begins to move from force absorption of impact – mid stance (10-30%)– to force propulsion forward – terminal stance (30-50%) [168].

   This brief summary of gait analysis theory in relation to the acceleration changes on the body CoM while walking provide the basis for the step detection methodology and the interpretation of the results from a biomechanics perspective.
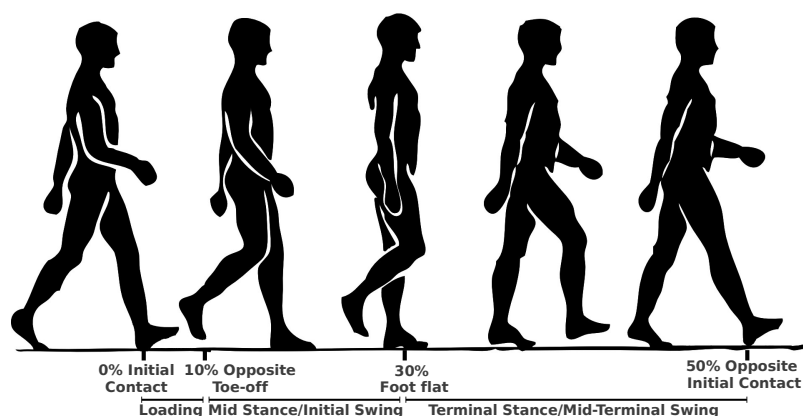
Figure 3.1.: Important phases for the detection of steps using the actibelt® sensor. Image adapted from [3]

3.3. Experimental Setup

3.3.1. Acceleration Sensor

The acceleration data was collected with the actibelt® sensor described in section 2.1

3.3.2. High Speed Video Camera

A high speed camera (Go Pro HERO3), configured to record at 100 Hz with the highest resolution (1280x960 pixels), was used as the gold standard for acceleration data annotation. Synchronization of camera and accelerometer was accomplished by filming a tap on the sensor and subsequently matching matching the peak acceleration value in the vertical acceleration component with the frame that displays the instant of the impact.

3.3.3. Data Collection Protocol

Standardized walking tests of healthy adults from a wide age spectrum (range 22 to 64 years) and elderly were included to produce a heterogeneous set of gait patterns. Adults were recruited among relatives and university students, and seniors were recruited at a nursing home. An experimenter supervised that the accelerometer was oriented correctly, firmly fastened around the waist, and centered at the middle of the mediolateral axis before the participants started

the test. The test consisted in walking a minimum distance of 10 meters at self-selected normal speed along a straight path with marks on the floor to track the speed. The experimenter walked slightly behind the test subjects, in order to influence their walking speed as little as possible, and holding the camera attached to a metal rod at ground level to record the contacts of the foot with the ground. Participants were allowed to use a walking frame, crutches or grab hold of somebody's arm for support. The local ethics committee approved the study and all subjects signed the informed consent prior to participation.

The instants of the initial contacts for a series of 10 consecutive steps were identified using the gold standard and annotated on the acceleration signal together with information about the average speed. Although I can not exactly quantify the accuracy of the gait speed estimated from the video data using the tracking marks on the floor, I assume that in average the maximum error that can be made in the annotation of the data is ± 3 frames (equal to ± 0.03 seconds), since out of this interval the foot tip has clearly crossed the mark on the floor, and therefore the potential deviations from the real gait speed are assumed not to have a big impact in the final results.

3.3.4. Database

Training and Validation Data Set

For the main analysis, and in order to identify possible differences in the walking patterns of healthy adults and seniors, I divided the data in two groups according to age ≤ 65 (n=10) and > 65 years old (n=21) respectively. In general, the age of 65 is considered as the age of entering old age [160, 189]. Afterwards, each group was separated randomly into a training part and a validation part for testing and evaluation of the algorithm. Among the seniors, 11 were very frail elderly who needed the use of walking frame to ambulate. A summary is presented in Table 3.1.

	Training			Validation		
	f:m	Age [yrs]	Speed [m/s]	f:m	Age [yrs]	Speed [m/s]
Adults	3:2	35.2$\pm$17.6	1.30$\pm$0.19	3:2	39.4$\pm$ 21.2	1.12$\pm$0.38
Seniors	8:2	82.0$\pm$6.9	0.54$\pm$0.21	10:1	82.5$\pm$ 6.1	0.40$\pm$0.16

Table 3.1.: Gender of participants (f=female,m=male), mean (SD) age and walking speed

Data Subset for Exploratory Analysis

The purpose of this exploratory analysis was to gain understanding about how the walking acceleration pattern changes with age (or equivalently with gait speed) to determine the optimal methodology for robust step detection, and as well as to examine the validity of the acceleration data recorded at the body's CoM and its consistency with the considerations made in Section 3.2 about the forces acting in the different gait phases.

I conducted the exploratory analysis on a random small data subset composed of a group of adults (4f:1m; 26-40 years old), "active seniors" (3f;74-80 years old) and "functionally dependent seniors" (2f:1m; 84-89 years old). The "active seniors" were those healthy seniors who can function on their own and can complete their daily tasks independently, and the "functionally dependent seniors" were those ones who depend on specific services due to declining health or diseases.

Walking-like Data

In addition, I recorded 9 so called "walking-like" activities or movements described in the Table 3.5 to estimate the upper limit of the false step rate. They are called walking-like movements because their acceleration pattern is similar to the pattern generated while walking, in terms of signal

amplitude and periodicity. Some of them involved physical exercise, like doing push-ups or sit-ups, while others were included to challenge the capacity of the algorithm to distinguish between real steps and patterns originated intentionally to "cheat" the step detector, as for example, shaking the sensor to emulate a walking pattern. These activities were executed by the same individual (a 30 year old female) and recorded in parallel with a high speed video camera.

3.3.5. Step Detection Methodology

As shown in the gait analysis section, the initial contact (IC) of the foot with the ground marks the beginning of a step. According to the measurements obtained in previous clinical studies with the use of force plates, the IC can generate ground reactions forces (GRF) in the vertical axis up to 1.5 times the body weight (equivalent to 1.5 g-force units) when walking at normal speed [138]. In the actibelt$^{®}$ signal this translates into an acceleration peak in the X axis at every contact of the foot with the ground as shown in Fig. 3.2a. In the preliminary experiments in healthy adults those peaks went from -1.4 g to -2.2 g for slow and fast walking respectively, and down to around -5 g for running. At walking, this means a maximum net acceleration change of 1.2 g (from -1 g to -2.2 g). This value is slightly lower than the 1.5 g but still consistent with the results from the literature in particular if we take into account that GRFs are measured at foot level whereas the actibelt$^{®}$ measures the acceleration near the CoM and therefore some of the impact which propagates upwards is absorbed by the limb muscles before it reaches the waist level. The initial contacts annotated using the video recordings may be shifted two to three milliseconds with respect to the minima in the acceleration signal (see Fig. 3.2). This may be caused by the propagation delay of the impact until reaching the waist and/or to a small systematic error introduced by the experimenter when annotating the data. The data annotation can be a subjective task depending on the quality of the image and the lighting conditions during the recording.
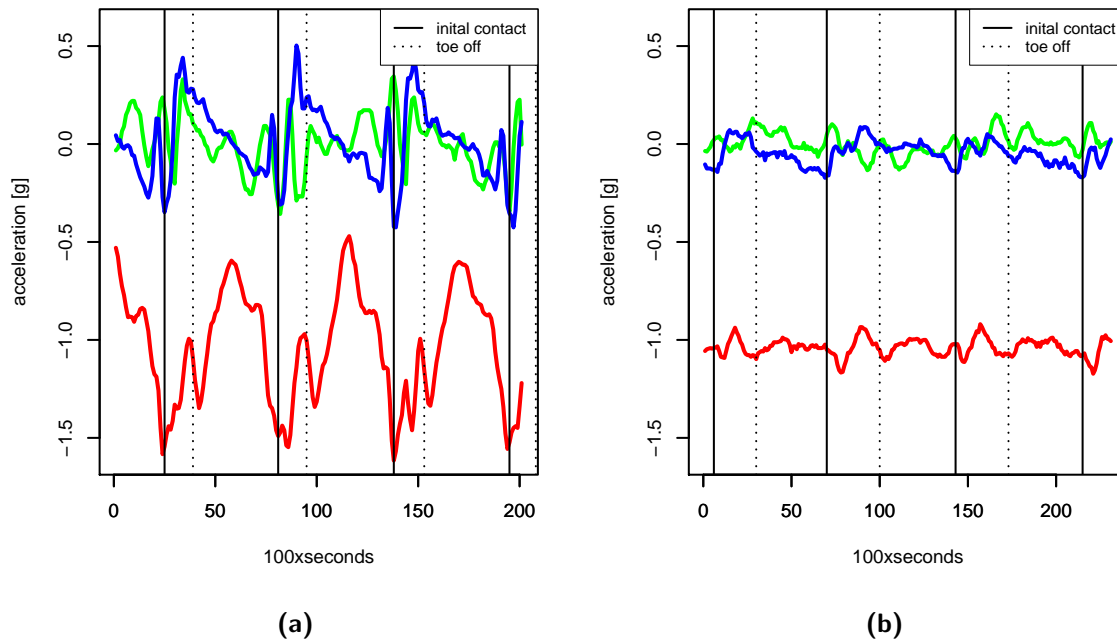

(a)   (b)

Figure 3.2.: Acceleration near the body's CoM of (a) a healthy young male and (b) an elderly woman (right) walking at normal self-selected speed.

Hence, the first observations of data from healthy individuals were plausible and they suggested a direct method to detect steps by determining the local minima along the vertical acceleration. However, when the data from the elderly group was explored, I found that the amplitude of those minima turned out much lower and noisy (Fig. 3.2b) making difficult to pinpoint the IC peaks. These changes in the amplitude were associated with the reduction in gait speed and the flat-footed landing usually caused by a decreased ankle plantarflexion in old walkers [198].

In order to find out whether the mediolateral or the anteroposterior axes could yield better results in the identification of steps in a broad spectrum of gait patterns, I compared the low frequency components of the average step among the group of adults, "active seniors" and "functionally dependent seniors" belonging to the exploratory data set. I found that the patterns in the mediolateral axis were very idiosyncratic for each individual, which confirmed the results from a previous study [207]. In contrast, the sagittal accelerations were more consistent. The variability increased with age but the graph still showed the existence of a common acceleration-deceleration pattern in the anterioposterior axis for all groups, more regular than for the vertical axis and in particular for the most functionally impaired group. In this axis the heel strikes produce a minima followed by a sudden increase of the acceleration related to the propulsion of the body upwards and forwards during the loading response phase. This acceleration is proportional to the amplitude of the horizontal GRF but exerted in the opposite direction [137].

In short, the results of the exploratory analysis suggested that the most robust method to detect steps could be achieved by identifying the local minima which precede the maximum forward acceleration of the body in the anterioposterior axis of the actibelt$^{\circledR}$.

3.4. Step Parameterization

3.4.1. Characteristic Step Graph

The first task in the process of step Parameterization is to obtain a normalized waveform corresponding to a standard representation of a step. To that effect, I used a so-called Characteristic Step Graph (CSG) for normal walking and a Characteristic Step Graph for age-related impaired walking. The CSG is calculated from a series of normalized Average Step Graphs.

Average Step Graph (ASG)

The Average Step Graph represents the average person's step and displays the average and standard deviation of each acceleration sample in a step interval for a series of single steps along the vertical, anteroposterior and mediolateral axes (Fig. 3.3).

In the same manner that the time differences in step duration offer information about the temporal variability, the ASG can be used to measure the spatial variability which has been associated in the clinical literature to risk of falling, mobility function and pathological gait [76, 27, 75]. Fig. 3.3 shows the differences in the average step for a healthy adult and a senior. In the exploratory analysis the amplitude of the ASG acceleration in healthy adults was on average more than twice as high as in the elderly group and the step variability increased up to three times as much with age.

Normalization of ASG

To make all the steps comparable to each other they were normalized, first, in time via a linear interpolation adjusting their length to the median value of all the step durations in the walking series and, second, in amplitude between -1 and 1 by the maximum absolute value.
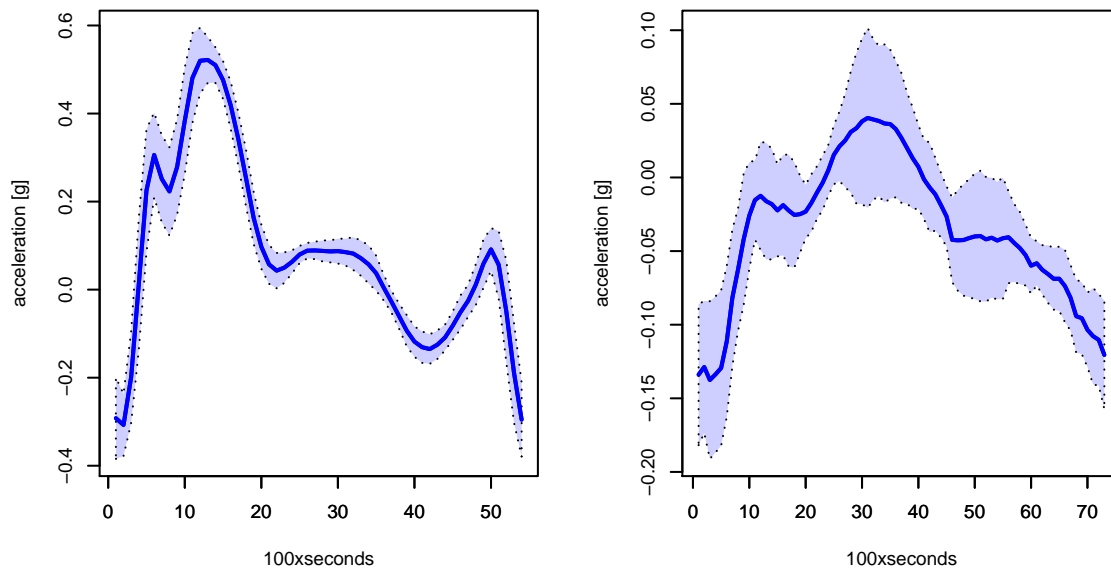
Figure 3.3.: Comparison of the average step acceleration pattern and standard deviation in the anterioposterior axis for (a) a healthy young individual and (b) an elderly walking at normal self-selected speed.

K-means Clustering

The exploratory analysis suggested the existence of slightly different patterns in the anterioposterior acceleration for the adults and seniors, or equivalently for normal and slow gait speeds. To corroborate this assumption, I applied K-means clustering [72] over the normalized ASGs and grouped them in two clusters: cluster A, later identified with "age-related impaired walking" (from now on, abbreviated as "impaired walking") and cluster B, characteristic of "normal walking".

From the total number of participants in the training data set (n=15, 22-90 years old), participants with gait speed under 0.75 m/s (n=10, 77-90 years old) were assigned to group A and the 80% of people with gait speed higher than 0.75 m/s (n=5, 22-36 years old) were assigned to group B. Only one middle-aged adult (age 65 years old) with speed = 1.01 m/s was assigned to the group A.

The Characteristic Step Graphs for impaired and for normal walking were chosen as the A and B cluster means respectively. The estimation of CSG from the clustered data is age-independent, and relies only in the acceleration step patterns which are directly associated to the gait speed. The parametrized CSGs are shown in black in Fig. 3.4.

3.4.2. Parametric Description

Five main fiducial points were identified in the anterioposterior direction of the normalized ASG, each of them associated with a particular gait event. For their identification I inspected in parallel to the acceleration data the video recordings in slow motion.

The **C** point (**Contact**) corresponds to the initial contact of the heel with the ground. In that moment the ankle is dorsiflexed with toes pointing up and the foot begins to land on the middle to outside of the heel. As the foot continues landing the ankle plantarflexes and finally the forefoot comes down.

In normal walking, when the heel of one foot, for example the right foot, touches the ground

the ankle of the left foot begins to plantarflex immediately bringing the left heel off the ground, and by the time the right forefoot reaches the ground the arch of the opposite foot recoils bringing the toes off. These actions push the body upwards and forwards bringing a peak on the anterioposterior acceleration, the **R-A** (**Rise** and **Advance**) peak.

The interval between the **C** point and the **R-A** peak corresponds to the double support phase described in the gait theory as the period in which both feet are in contact with the ground. In this analysis I found that this phase occupies the 22% of the step time or equivalently, 11% of the stride cycle for normal walking, what is in agreement with the results from the clinical literature [151].

In the elderly, the duration of the double stance is usually bigger. Winter [198] reported a difference of more than 6% with respect to young adults indicating that the elderly adapt toward a safer and more stable gait pattern. I found that on average the time spent by the elderly in double stance was in comparison with the healthy adults of around 11% bigger. The elderly tended to wait until the forefoot lands on the ground to start moving the rear foot forwards, which causes a drop in the forward acceleration followed by an increase again when the calf muscles of the rear foot begin to bring the heel off. This discontinuity in the loading response split the **R-A** peak in two, **R** and **A**. The **R** (**Rise**) point coincides with the instant at which the forefoot of the stance leg comes down to the ground and the **A** (**Advance**) point marks the instant which immediately precedes the toe off. I observed that for a few elderly the amplitude of the **R** peak, Δ, was higher than that of the **A** peak. This is indicated in figure 3.4 with gray dotted lines. I also drew in semi-transparent text the **R** peak found in the pattern of several young adults that was filtered out when averaging all step patterns. This fine **R-A** separation present sometimes in the younger group is not related to any pathological gait pattern, since the loading response appeared totally normal and smooth and the participants did not report any medical condition affecting the walking ability. This fluctuation is also not necessarily present in all the steps taken by the same individual. The minimum and maximum difference in acceleration within the 22% first interval found in the training data was: $\min \Delta = 0.09$ [g] for an elderly and $\max \Delta = 1.05$ [g] for a young adult.

After the **R-A/A** instant, the acceleration begins to decrease until it reaches a point of **Inflection**, **I**, which marks the beginning of the mid-swing phase. In that moment the hip is pulled forward by the concentric hip flexor and the knee extends rapidly. The end of this phase coincides with the end of the leg swinging towards a maximum extension, the so called **S** (**Swing**) peak in the step graph.

Again, the possible variations on the mid-swing phase were depicted in gray in Fig. 3.4. The smoothing caused by the average of the step graphs makes the **I** and **S** points to vanish. However, when inspecting each single pattern the **I** and **S** peaks still exist for a minority of elderly without reduced knee extension. The absence of **I-S** interval seemed to be closely related to the use of walking aid, especially walking frame.

3.5. Step Detection Algorithm

The step detection algorithm is divided into three main parts: recognition of regions of interest, estimation of ICs and template matching. Each of them is described in detail in the following sections. The algorithm was implemented in R language [174]. The raw data was processed in overlapped blocks of length equal to 1000 samples (empirically calculated as the optimal length [63]) and although there were not constraints in terms of execution time the algorithm was written to completely avoid the use of loops.

Mathematical notation The sensor measures the acceleration at a rate of 100 samples per second in three orthogonal axes: vertical (X), mediolateral (Y) and sagittal (Z). On this basis,
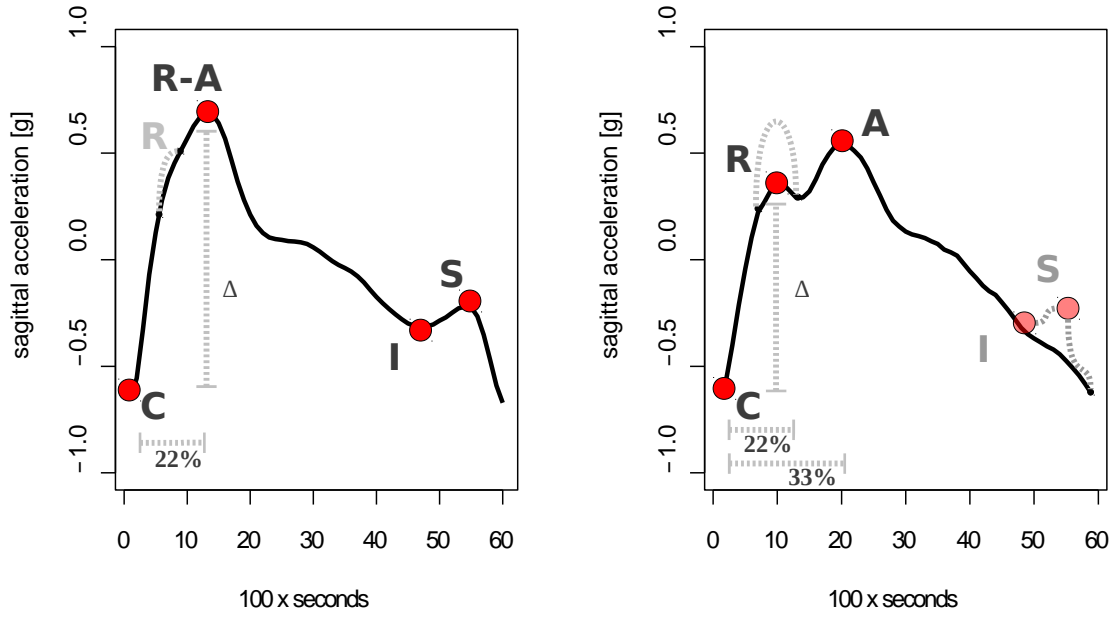
Figure 3.4.: On the left the parametrized Characteristic Step Graph for normal walking (cluster B) and on the right for impaired walking (cluster A). Dotted lines show possible variations of the pattern.

the acceleration signal is considered a discrete multivariate time series where each acceleration measurement can be defined as an element of a three dimensional Euclidean space in Cartesian coordinates as

$$\mathbf{a}_n = (a_n{}^x, a_n{}^y, a_n{}^z) \tag{3.1}$$

with $n = 1, 2, \ldots, N$ and $\mathbf{a}_n \in \mathbf{R}^3$.

3.5.1. Recognition of Regions of Interest

Walking implies the realization of a certain level of physical activity and to maintain a more or less upright standing posture. Using the capacity of the device as inclinometer and measuring the magnitude of the accelerations near the CoM we can estimate which regions may contain walking segments. To recognize the regions of interest, I first excluded from the analysis all those intervals where the body orientation was not upright. We consider an upright posture when the angle formed by the trunk with the sagittal axis (ϕ_B) is smaller than 45° (ignoring possible sensor tilt).

Second, I used a threshold-based segmentation of activity levels in 2 zones. The activity level, g_n, is obtained by differentiation of the acceleration vector in the sagittal plane,

$$g_n = |v_{n+1} - v_n| \tag{3.2}$$

where $v_n = \|(a_n{}^x, a_n{}^z)\| = \sqrt{(a_n{}^x)^2 + (a_n{}^z)^2}$.

Before calculating the activity threshold, this signal g_n was smoothed with a Gaussian filter ($\frac{1}{\sigma} = 2.5$) of length $L = 400$ samples to guarantee that adjacent regions with similar activity level were merged together. The threshold was selected in such a manner that all walking segments in the training data set were included inside the region of interest. Since some of the elderly had a very impaired gait (in particular those using walking frame) their activity while walking was

very low and likewise the threshold. All those values under the threshold were not considered for further analysis.

Hence, the activity profile was divided in two zones, with one of them corresponding to very low activity or motionless and/or non-upright orientation, and the other one to upright position and medium-high activity levels. This procedure does not guarantee that the values over the threshold correspond to walking phases but restricts the area to search for steps improving the algorithm's performance time and specificity.

3.5.2. Estimation of Initial Contacts

As explained in section 3.3.5, the step detection is based on the detection of the greatest local minima in the anterioposterior axis which coincide with the instants of the initial contacts. These local minima are sometimes difficult to discriminate due to undesired fluctuations in the signal usually produced by abdominal fat oscillations or by the hip extension. To remove these fluctuations, I applied a low pass Butterworth filter $(n = 4, \ f_c = 2Hz)$ to the anterioposterior acceleration $a_n{}^z$. After that, I calculated the minima n_{min}

$$n_{min} = \{n \in 1, 2, \dots, N \mid \tilde{a}_{n-1}^z - \tilde{a}_n^z > 0 \land$$
$$\tilde{a}_{n+1}^z - \tilde{a}_n^z < 0\} \tag{3.3}$$

in the filtered signal $\tilde{a}_n^z$ by differentiation.

Each local minima marks off the interval I

$$I = [n_{min}, n_{min} + T_{min}) \tag{3.4}$$

where to search for the initial contacts in the anterioposterior axis a_n^z. The duration of the interval is equal to the minimum step duration (T_{min}) that I have estimated in $1/4.2$ seconds, the inverse of the usual maximum step frequency during running [20]. For healthy adults and elderly walking faster this limit decreases down to 2.5 steps per second [205, 19].

The first approximation of the initial contacts, $\hat{c}_i$, are the minima of the acceleration signal in the interval I along the anterioposterior axis, calculated again as in equation 3.3.

The next task was to verify whether each segment

$$\hat{\mathbf{s}} = \{a_n^z \mid n \in [\hat{c}_i, \hat{c}_{i+1})\} \tag{3.5}$$

between two consecutive estimations of the initial contacts fits in the parametric description of a step.

The first condition that the vector $\hat{\mathbf{s}}$ needs to fulfill to be considered as a step is to present a maximum of amplitude $\Delta \geq \min \Delta$, with $\min \Delta = 0.09$ [g], in the first 22% of the interval. This maximum usually corresponds to the **R-A/A** point in normal walking and to the **R** point in impaired walking (Fig. 3.5). The **I** and **S** fiducial peaks were not considered for the step detection because they are not always present in the signal. The second condition is that the length of the vector $\hat{\mathbf{s}}$ in seconds should not be bigger than a maximum step duration $T_{max} = 2$ s. This value is approximately equal to the double of the biggest step duration measured in the training data set (1.06 seconds).

The vectors that satisfy both conditions are the estimations of the real steps and they are denoted as s.

3.5.3. Template Matching

One of the most challenging tasks in step detection under uncontrolled environment is the ability of the algorithm to distinguish between real steps and activities with similar patterns. In digital
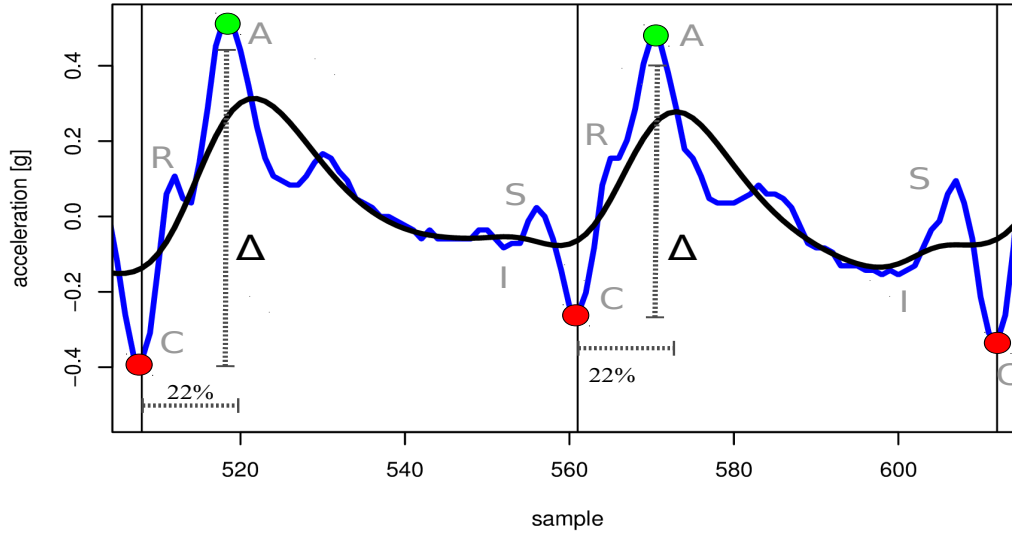
Figure 3.5.: Initial Contacts (ICs) detected using the algorithm (in red). The blue line is the anterieopos-terior acceleration (g units) and the black line is the low-pass filtered version. Vertical lines correspond to the annotated ICs.

image processing a common technique to identify characteristic patterns is template matching. I applied the same principle here in order to reduce the number of false positives by measuring the similarity of every acceleration pattern **s** with a template **t** and depending on this value decide whether the estimated step fits the shape of a standard step pattern or not. The normalized cross-correlation between both signals, defined as

$$\gamma_{\mathbf{s},\mathbf{t}} = \frac{1}{L}\sum_{j=1}^{L} \frac{(s_j - \mu_{\mathbf{s}})(t_j - \mu_{\mathbf{t}})}{\sqrt{\sigma_{\mathbf{s}}\sigma_{\mathbf{t}}}}, \tag{3.6}$$

was used as the similarity measure. In the equation 3.6, the scalars μ_x and σ_x denote in general the average and the standard deviation of any signal x, and L is the length of the template.

The Characteristic Step Graph for normal ($\mathbf{t_1}$) and for impaired walking ($\mathbf{t_2}$) were used as templates. And the length of the estimated step signal **s** was adjusted to the length of the template by linear interpolation.

The steps with a maximum normalized cross-correlation

$$\max(\gamma_{\mathbf{s},\mathbf{t_1}}, \gamma_{\mathbf{s},\mathbf{t_2}}) < 0.70 \tag{3.7}$$

were considered not to fit the standard. The value 0.70 corresponds to the lowest value of the normalized cross-correlation between the templates $\mathbf{t_1}$ and $\mathbf{t_2}$ and the steps annotated in the training data set.

3.6. Validation

The annotated steps in the validation data set were again used to calculate a Characteristic Step Graph for normal and impaired walking and to cross-check the fiducial points determined in the first analysis.

To validate the step detection algorithm I measured the accuracy in the validation data set and an estimation of the specificity using the walking-like data.

The accuracy of the step detector is determined by the average trueness $\overline{T}$

$$\overline{T} := \frac{1}{n} \sum_{i=1}^{n} 1 - \frac{|N_i^{Alg} - N_i^{Gold}|}{N_i^{Gold}}, \tag{3.8}$$

where N_i^{Alg} and N_i^{Gold} are equal to the number of steps detected with the algorithm and with the gold standard for each person i, and by the average error $\overline{E}$

$$\overline{E} := \frac{1}{n} \sum_{i=1}^{n} \left(\frac{1}{l} \sum_{j=1}^{l} |c_{ij}^{Alg} - c_{ij}^{Gold}| \right), \tag{3.9}$$

where c_{ij}^{Alg} are the initial contacts detected with the algorithm and c_{ij}^{Gold} are the ones annotated with the gold standard for each step j and participant i.

The specificity was measured in terms of the maximum false step rate ($\max FS$), as the number of false steps per minute when running the algorithm over a sequence of walking-like activities. The value for the specificity given here is considered an upper limit because in a realistic scenario the rate at which these kind of activities occur over the time interval measured with the sensor is very much lower than in the experimental walking-like sequences. To obtain an actual value of the false step rate one should consider the relative frequency of occurrence of the walking-like activities described in Table 3.5 for each age group or/and mobility status. For example, in general the probability that a person affected by a severe functional impairment rides a bicycle regularly may be smaller than for a healthy individual.

3.7. Results

3.7.1. CSG Cross-Check

The K-means clustering applied to the average anterioposterior accelerations to determine the Characteristic Step Graph yielded good results when discriminating between groups with gait speed $v \geq 0.75$ m/s (labeled as "normal" walking) and those with $v < 0.75$ m/s (labeled as "impaired" walking) in the training data set, as seen on the confusion matrix in Table 3.2.

	impaired	normal
impaired	9	0
normal	1	5

Table 3.2.: Confusion Matrix for the training data set

However, the quality of the clustering using the same criterion was not so accurate when applied to the validation data (Table 3.3a), with 5 persons with "impaired" walking classified as "normal". Two of those five persons, presented the highest speed in the actual "impaired" group (0.62 m/s and 0.66 m/s) and the three remaining presented extreme to very low speeds (between 0.12 m/s and 0.36 m/s), which evidenced their condition as outliers. This assumption was confirmed when the three misclassified samples with low gait speed were grouped together in a new cluster after applying again the K-means algorithm and setting the number of clusters k to 3. Table 3.3b shows the Confusion Matrix after removal of outliers. The still misclassification of the two patterns corresponding to gait speeds in the range 0.6-0.7 m/s suggests that the step graph may still present characteristics of a normal walking pattern at those speeds.

Combining the results of the training and validation analysis, 90% of acceleration patterns belonging to healthy adults (n=10) were classified in the group "normal", unlike only 9.5% of step patterns of seniors (n=21) which were also assigned to this group.

The cluster means of the validation data set produced almost identical Characteristic Step Graphs as the ones depicted in Fig. 3.4 for the training data set. The Average Step Graph of the

	impaired	normal
impaired	7	5
normal	0	4

(a) k =2

	impaired	normal
impaired	7	2
normal	0	4

(b) k=3

Table 3.3.: Confusion Matrix for the validation data set

elderly who use a walking frame (Fig. 4.14b) in comparison with the elderly who do not use one (Fig. 4.14a) revealed that for the first group the interval **I-S** tends to disappear due probably to the reduced knee extension characteristic for this type of walk [13], and the **R** peak amplitude is on average smaller than the one for the **A** peak.
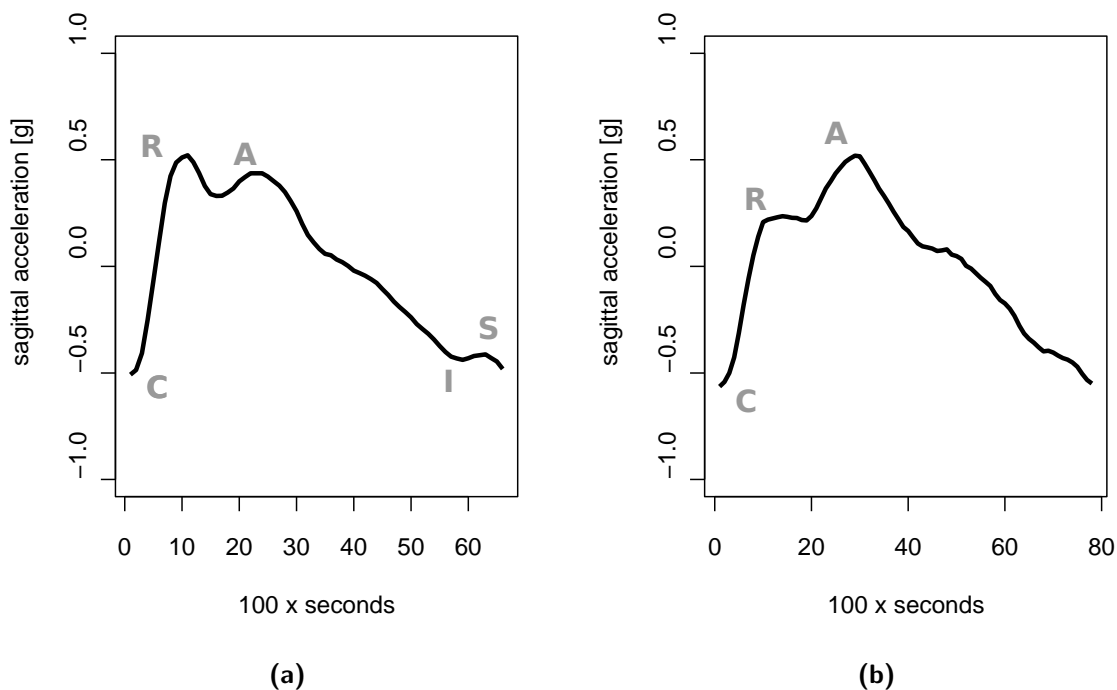

(a)

(b)

Figure 3.6.: The Average Step Graph of the elderly (a) without a walking frame (n=10) and (b) with a walking frame (n=11)

3.7.2. Algorithm Performance

The accuracy of the step detection algorithm was dependent on walking speed. The spectrum of gait speeds in the validation data set ranged from 0.12 m/s for the most functional impaired elderly to 1.47 m/s. Without the use of template matching, the average trueness in this data set was almost excellent ($\overline{T} \approx 1$) for speeds between 0.20-1.5 m/s but it reduced down to near the 50% for gait speeds under 0.20 m/s. The minimum and maximum error in the IC detection was of the order of 1 and 6 hundredths of a second respectively. The precision of this error $\overline{E}$ may vary depending on the precision with which the experimenter synchronized the acceleration data with the gold standard data and annotated the ICs and, as mentioned already in section 3.3.3, this difference should not be bigger than 0.03 seconds. On average, trueness was reduced by 12.3% when using template matching. The detailed results of the algorithm performance are illustrated in Table 3.4.

Except for cycling, the false step rate of the algorithm without applying template matching was very good for most of the eight activities presented in Table 3.5, with nearly zero false steps for seven of the activities and 20 false steps/min detected in one of them. When applying template matching the false step rate during cycling was reduced by 70% and went down to zero for all the rest of activities.

	Speed [m/s]	n	$\overline{T}$	$\overline{E}$ [s]	$\overline{T}_T$
Normal	$1.00 \leq v < 1.50$	3	1.00	0.01	0.97
Slow	$0.50 \leq v < 1.00$	6	0.98	0.04	0.82
Vslow	$0.20 \leq v < 0.50$	5	1.00	0.06	0.80
Xslow	$0.10 \leq v < 0.20$	2	0.45	0.06	0.35

Table 3.4.: Average error ($\overline{E}$) and trueness with ($\overline{T}_T$) and without ($\overline{T}$) template matching

	Movement description	max FS [steps/min]	max FS_T [steps/min]
1	Sitting and moving heel up-and-down	0.0	0.0
2	On tiptoes and lowering back down	20.0	0.0
3	Sit-ups	0.0	0.0
4	Push-ups	0.0	0.0
5	Standing toe-touch	3.2	0.0
6	Shaking sensor up-down	0.0	0.0
7	Shaking sensor backward-forward	0.0	0.0
8	Shaking sensor left-right	0.0	0.0
9	Cycling	96.0	26.0

Table 3.5.: False step rate in walking-like movements with ($\max FS_T$) and without ($\max FS$) template matching

A small ad-hoc analysis was done to test the algorithm performance with running data. The exploratory results of the step detection algorithm without the use of template matching showed a good performance ($\approx 100\%$ of steps detected) when applied on a set of running sequences from 3 different young individuals. To use template matching, new templates for running steps should be generated.

3.8. Summary of Findings

The results of the step pattern examination from the CoM accelerations were consistent with the qualitative and quantitative description of the gait phases presented in the clinical literature [151, 198] and with other previous studies using force plates [152, 138] and accelerometers [207].

I found that the axis with a more congruent pattern for all ages and speeds is the anterioposterior axis and the minima along this direction coincide with the initial contacts.

The significant differences in the amplitude and the standard deviation of the Average Step Graph between a random subset of elderly and adults suggest that the ASG could be a good measure to assess functional decline. The K-means algorithm misclassified only 3 out of 31 Average Step Graphs (after removal of outliers in the validation data set) when they were categorized in two groups, labeled as "impaired" and "normal" walking, according to walking speeds lower and equal or greater than 0.75 m/s respectively. Among the step graphs classified as "normal", approximately the 82% belonged to the healthy adults and the rest to fit seniors. These results evidence the correlation among the acceleration step pattern, the age and the gait speed.

Apart from the expected changes in the maximum amplitude of the acceleration with the speed, the device was able to detect little fluctuations in the acceleration at the CoM level

originated with the heel strike, the forefoot loading, the plantar flexion of the foot and the knee extension. The Parameterization of these gait events and posterior analysis revealed differences in the biomechanics of gait among healthy adults and elderly with and without walking frame. The acceleration-deceleration pattern of the signal envelop was also coherent with the inverted pendulum model of the CoM trajectory [206].

To what extent the variations in the acceleration step pattern may result from the effect of the walking speeds or from other factors related to aging, like loss of strength and flexibility, should be further investigated. Some gait movements, as for example the knee flexion, has been shown to be related to walking speed [96], therefore it is essential that normal ranges for gait parameters are defined with reference to walking speed.

The step detection algorithm can be reliably used for the study population when walking at speeds higher than 0.20 m/s independently of the use of walking aid. At lower speeds approximately only half of the steps taken can be detected. For high speeds the ad-hoc exploratory analysis showed promising results with almost all steps detected, however, in order to assess the validity of the algorithm at running speeds an additional analysis is required. In uncontrolled settings, the utilization of template matching techniques to reduce potential false positives due to walking-like activities, like cycling, may reduce the accuracy of the step counter. Therefore, the decision between applying this additional data processing or not, depends in part on a trade-off between the required accuracy and the specificity. This compromise should be assessed once the frequency with which these kind of activities are performed has been estimated. Other factor to take into account is the time performance of the algorithm, since the template matching may increase considerably the time needed to process the data.

To the author's knowledge, this is the first study using a single waist-worn accelerometer that a) has developed a reliable algorithm for step detection in a healthy population of adults/youngs and in seniors walking at very low speeds, and b) that identified fiducial points in the acceleration pattern with clinical significance for the assessment and monitoring of age-related functional decline.

Although varied step patterns and gait speeds have been used in this study for the algorithm development, additional validations in bigger data sets and with different pathological gaits (e.g. MS or Parkinson disease) would be needed, such as the assessment of the ecological validity of the step detector in uncontrolled environments like in the community or at home, where walking is not limited to a straight walk.

Systems using a single waist-worn 3D accelerometer are a promising technique for reliable long-term evaluation of physical functioning and monitoring of exercise therapy or medication interventions. In particular the placement of the sensor inside a belt buckle guarantees a limited obtrusiveness and the fixed location near the center of mass of the body provides accurate information about body orientation and other parameters such as gait symmetry or trunk sway, as will be shown in the next chapter. Analogous to electrodes that measure the electrical activity of the heart to produce an Electrocardiogram (ECG), waist-worn accelerometers could measure the motions near the body's center of mass to trace out an acceleration graph of the gait, an "AcceleroCoMgraph" (ACG).

4. Postural Stability Analysis

4.1. Introduction

Postural stability is typically assessed by measuring the body sway during the performance of some standardized functional tests using force-sensing plates. Most of the parameters that measure body sway are based on the excursion of the body Center of Pressure (CoP) in the transverse plane, the most common being the area of the ellipse that contains the 95% of the samples in this plane, the path length and the velocity [156, 158].

Winter [196] suggested that the displacement of the body Center of Mass (CoM) could be a better indicator of body sway than the CoP. The disadvantage is that the estimation of the CoM is laborious since the position of the CoM is not fixed but it changes depending on distribution of the body segments. There exist complex models that estimate the CoM position based on the location, magnitude and mass distribution of the body segments [73, 199, 106], but the costly 3D infra-red technology required to obtain these data makes this procedure impractical in clinical settings. Other approaches are based in the low-pass filtering of CoP data [106, 17, 33] or tracking a point in the body surface near to where the CoM is estimated to be during the performance of the balance test [21, 58, 97]. The last two mentioned methods do not estimate the CoM displacement in such an accurate manner as the segmental approach [56] but they present an important advantage, they simplify substantially the data collection procedure which is a crucial factor for the management of time and resources in the clinical practice.

Winter [197] also showed that, in an inverse pendulum model, the difference between CoP and CoM was proportional to the horizontal acceleration of CoM and that this difference could be seen as an "error" signal in the balance control system which causes this CoM's acceleration, therefore of considerable importance to understand the biomechanics and muscular synergies of human upright posture. Until the introduction of accelerometers in clinical research, the CoM was usually estimated by double differentiation of CoP data, a procedure that introduces considerable noise in the results, but in the last years with the growing advance in MEMS technology, numerous groups have studied the feasibility of using miniaturized inertial sensors to measure body sway [126, 130, 145, 183, 146, 123, 51, 48]. These studies have shown good results in the identification of populations at risk of falling [145, 51] and to characterize early mild Parkinson's Disease patients [146]. Nevertheless, accelerometry is still not an standard tool for the assessment of postural control, probably because of a limited validation and the lack of normal ranges for healthy subjects and for patients suffering from specific balance disorders.

In this section, the purpose is to 1) estimate and validate the CoM displacement using a single waist-worn accelerometer and 2) compare the parameters derived from the acceleration data and the same parameters derived from the CoP excursions in order to investigate the validity of the sensor as an alternative tool for the quantitative assessment of postural stability.

4.2. Estimation of the Center of Mass Displacement during Standing

In this section two methods for the estimation of the Center of Mass displacement are presented. The first method consists in the double integration of the movement component of the acceleration and the second method is based on the inverted pendulum model. Both methods have been

used in previous works, but it is still open which one approximates the empirical results better. This question will be treated in later sections.

4.2.1. Method 1: Double Integration of Pre-Detrended Acceleration Signal

Theoretically the center of mass displacement could be estimated by attaching an accelerometer on the body waist near the body's CoM and integrating the measured acceleration twice.

Mathematically, the calculation of displacement $x(t)$ from a measured acceleration $a(t)$ is

$$x(t) = d_0 + v_0 t + \int_0^t dt \int_0^\tau a(\tau) d\tau \tag{4.1}$$

where d_0 is the initial displacement and v_0 the initial velocity. This formula is used for continuous functions but nowadays almost all signals are discrete and therefore the integration must be numerical. The numerical integration per se is not a complicated task and several methods exists that minimize truncation errors [44]. When working with real sampled acceleration data the main source of errors are given in general by the unknown initial conditions, bias, noise, and low sampling resolution.

Initial conditions For the calculation of the CoM displacement during the performance of balance tests, the initial conditions, and the low sampling resolution did not pose a problem. The initial velocity and position are assumed to be equal to zero since the patient should begin the measurement in a steady state posture and the initial location of the trunk corresponds to the origin of the coordinates system.

Sampling resolution We also consider that the error introduced by the discretization of the analog signal does not affect the results significantly. We can make this assumption because the sampling rate of the sensor (100Hz) is in this case much higher than the maximum expected trunk sway frequency under quasi-static conditions ($\approx$ 3Hz) [45, 154] guaranteeing a sufficiently good approximation of the real acceleration.

Bias The integration of a biased signal adds a linear drift to the result, and a double integration produces a parabolic drift. Figure 4.1 shows an example of double integration over a sinusoidal acceleration to which a bias of 0.01 m/s^2 was added. Although the sinusoid is still present, the result is clearly overwhelmed by the parabolic drift. To minimize the bias error it is important to detrend the measured acceleration, separating properly the movement-related acceleration from the gravity component.

Noise The integration process acts as a low pass filter, thus in acceleration signals affected by noise the double integration tends to smooth the irregularities in the estimated displacement. However, the numerical double integration has an undesirable effect due to its cumulative nature, it introduces a variable drift of the actual displacement, particularly in the regions where the noise amplitude is larger in comparison to the signal amplitude, as it is shown in Figure 4.2. This is most likely a result of the noise being more positive or negative in certain regions, which essentially acts as a small DC bias in the signal after the first integration.

Numerical Integration Methods

Displacements from numerical integration of digital signals can be computed with different methods [139]. Usually, all the available numerical techniques for integration calculate the area under the graph of the discrete function over time. The simplest method to estimate that area
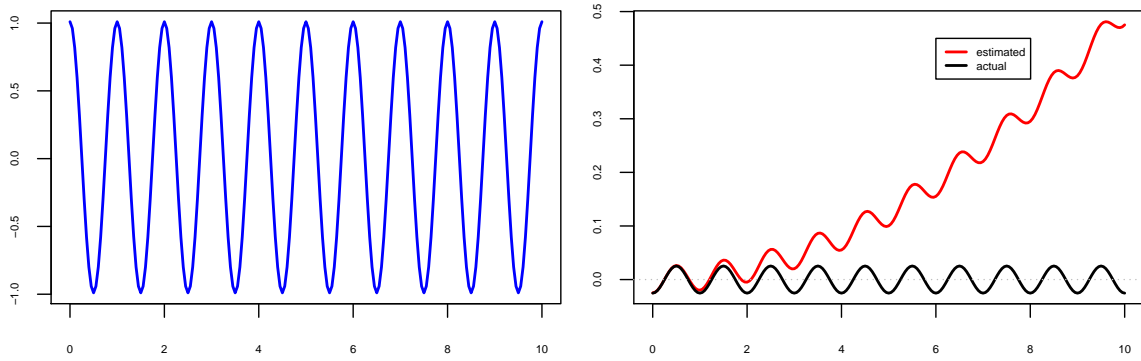
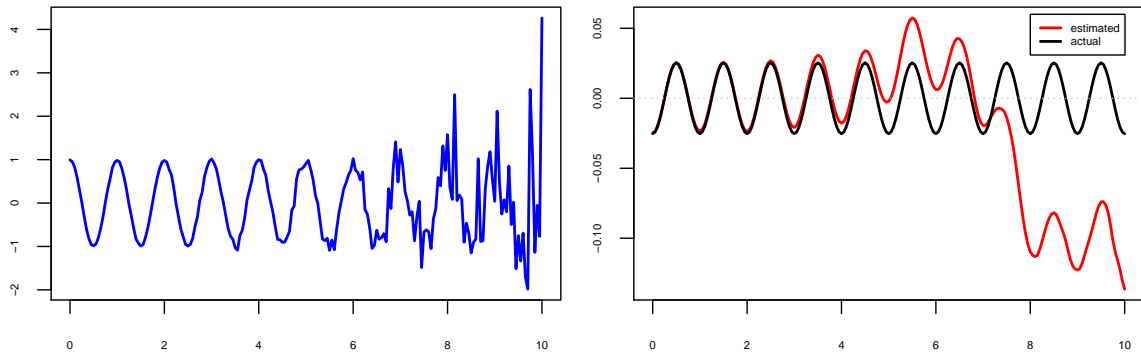**Figure 4.1.:** Sinusoidal acceleration with bias (left) and estimated and actual displacement (right)

Figure 4.2.: Sinusoidal acceleration with noise (left) and estimated and actual displacement (right)

is the **rectangular integration** method. This method consists of summing up all past samples and the current input and divide the result by the sampling rate. Rectangular integration can be represented by the difference equation

$$y[n] = \frac{1}{f_s} \sum_{k=0}^{n} x[n-k] = y[n-1] + \frac{1}{f_s} x[n] \tag{4.2}$$

where $x[n]$ is the signal to integrate and f_s the sampling frequency.

Another simple and more accurate method for integration uses the **trapezoidal rule**. The output of the integrator using this method can be written as

$$y[n] = y[n-1] + \frac{1}{2f_s}(x[n] + x[n-1]) \tag{4.3}$$

where again $x[n]$ is the integrand and f_s the sampling frequency.

The **Simpson's rule** approximates the integrand with a sequence of quadratic parabolic segments in contrast to the previous methods that use picewise linear approximations. The output of the Simpson's integrator can be defined as

$$y[n] = y[n-2] + \frac{1}{3f_s}(x[n] + 4x[n-1] + x[n-2]) \tag{4.4}$$

with $x[n]$ the signal to integrate and f_s the sampling frequency.

Because it uses quadratic approximations, when the underlying function is smooth Simpson's rule is more accurate than the trapezoidal rule, specially for low sampling frequencies. Figure 4.3 illustrates the three mentioned methods.
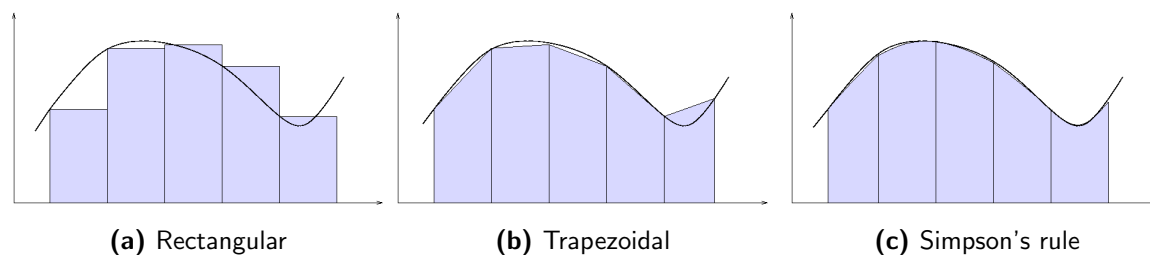

(a) Rectangular   **(b)** Trapezoidal   **(c)** Simpson's rule

Figure 4.3.: Illustration of the rectangular, trapezoidal and Simpson's rule numerical integration. Rectangular rule acts as a zero-order hold, trapezoidal rule acts as a first-order hold and Simpson's rule uses quadratic interpolation to approximate the integrand. Image source: [2]

Under quasi-static conditions the bandwidth of the acceleration signal (0-3Hz, [45, 154]) is much lower than the sampling frequency of the accelerometer (100Hz). Therefore, even the worst of the approximations – the linear approximation – would be enough accurate for our purpose.

Acceleration Detrend in Quasi-Static Conditions

Most of the acceleration sensors, including the one used in this work, measure the gravity acceleration and the acceleration due to body movements (see Equ. 2.1). In order to get only the movement-related component, the measured acceleration should be detrended.
Under steady state conditions the gravity component manifests itself as an offset in one or more sensor axes and it can be used to detect the sensor inclination with respect to the vertical plane [186]. In this case, it is enough to subtract the average from the recorded data in order to obtain the movement acceleration vector.
In the presence of rotational movements the frequency domains of the movement-related acceleration vector and the gravitational acceleration can overlap, thus making a perfect separation of both components difficult [77].

Humans tend to rotate when swaying, so if the sensor is fixed to the human body, the device which is designed to measure purely translational movements will also be subjected to rotations that will severely distort the data if not accounted for. If the acceleration signal is not properly detrended the residual gravity vector will act as a bias that after the double integration process transforms into a drift on the displacement signal.

Using a Savitzky-Golay filter to detrend acceleration   The Savitzky-Golay (S-G) filter is a smoothing filter based on a local least-squares polynomial approximation method. It was originally proposed by Savitzky and Golay [162] and used to smooth noisy data obtained from chemical spectrum analyzers.
The S-G filter smoothes a given signal $x[n]$ by fitting a polynomial

$$p(n) = \sum_{k=0}^{N} a_k n^k \tag{4.5}$$

of a specified order N to a set of input samples of length $2M + 1$. The coefficients of the chosen polynomial should minimize the mean-squared approximation error for the group of input samples
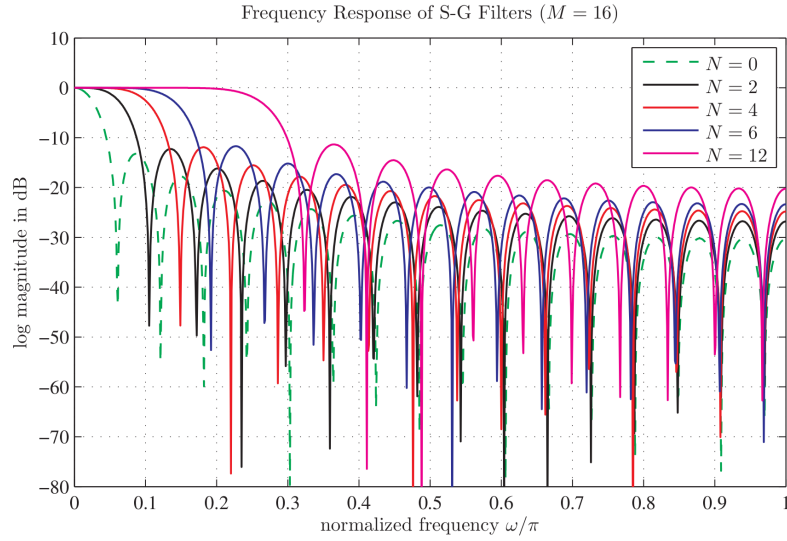
Figure 4.4.: Frequency response of the Savitzky-Golay filter for a window of length M =16 and different polynomial order N. Image source: [163]

centered on $n = 0$,

$$\varepsilon_N = \sum_{n=-M}^{M} (p(n) - x[n])^2 = \sum_{n=-M}^{M} (\sum_{k=0}^{N} a_k n^k - x[n])^2. \tag{4.6}$$

Figure 4.4 shows the filter response for different polynomial orders. A reasonably accurate approximation for the 3dB cutoff frequency is this equation [163]

$$f_c = \frac{N+1}{3.2M - 4.6} \qquad M \geq 25 \text{ and } N \leq M. \tag{4.7}$$

The advantage of using Savitzky-Golay filtering over simple digital filtering is the ability to preserve higher order moments around inflection points, like local maxima and minima. Also, unlike a moving average, in estimating the value of the fit at a certain point, it does not factor in the values on the polynomial fit around it, therefore not introducing a bias at inflection points [155].

Savitzky-Golay filters have been used in applications like electrocardiogram (ECG) analysis [71] and in processing of ECG signals in combination with accelerometers to remove the slower variations in motion-induced components of the signal and preserve the heart beats [29, 147]. Some publications in motion and balance analysis using accelerometers mention the use of S-G filters or the basic concept of least-squares polynomial filtering for smoothing the acceleration raw data [136, 58] and to separate the motion and gravitational components for the estimation of trunk sway [130, 21, 51, 97]. However, none of them reported the accuracy in the estimations using these type of filtering and there is still not a general rule for choosing the suitable values for the window length and the filter order that reduce the error between the estimated trunk sway and the real displacement.
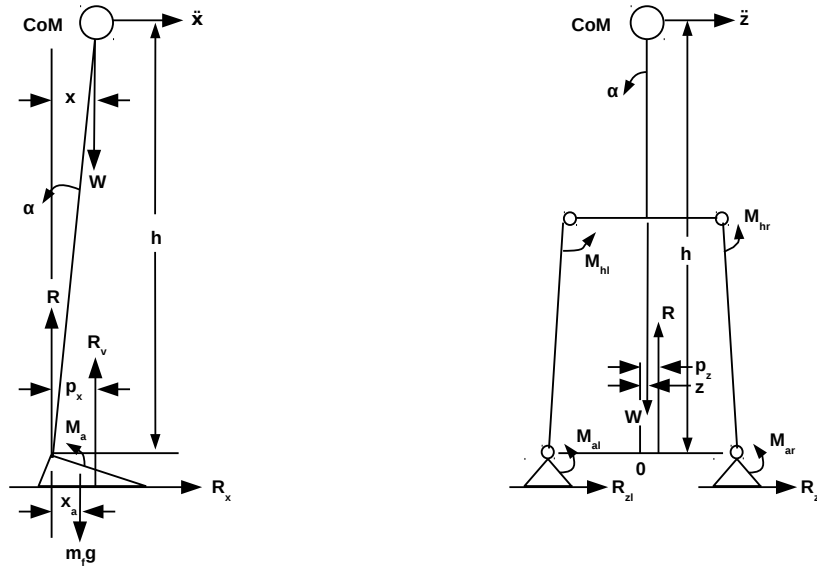
Figure 4.5.: Inverted pendulum model for standing posture in the sagittal (left) and frontal (right) planes according to Winter et al.[199]

4.2.2. Method 2: Anthropometric Filter Based on the Inverted Pendulum Model

The inverted pendulum model for standing posture originally proposed by Winter et al.[199] relates the controlled variable, CoM, with the controlling variable, CoP. Such a model provides an analytic relationship between these two commonly measured variables and the horizontal acceleration of the CoM. This relationship shows that the difference CoP-CoM is proportional to the horizontal acceleration of the CoM in both the sagittal (anterior/posterior direction, A/P) and frontal (medial/lateral direction, M/L) planes. The equations that capture this relationship are

$$p_x - x = (-I_{sa}/Wh)\ddot{x} \tag{4.8}$$

where p_x and x are the CoP position and CoM position with respect to the ankle joint in the A/P direction, $\ddot{x}$ is the CoM horizontal acceleration, I_{sa} is the inertia of the body about the ankle joint in the sagittal plane, W is the weight of the body (minus the weight of the feet), and h is the CoM height above the ankle joint, and

$$p_z - z = (-I_f/Wh)\ddot{z} \tag{4.9}$$

where z refers to displacements in the M/L direction and I_f is the inertia of the body about the ankle joint in the frontal plane.

Winter et al. validated the proposed model (Fig. 4.5) by demonstrating a high correlation between the CoP-CoM error signal and the respective horizontal accelerations of the CoM in each plane. The correlation for large sways ranged between -0.96 and -0.99, whereas for small amplitudes turned out to be slightly lower. This reduction was attributed to the decrease of the precision of the CoM estimate at very low amplitudes.

Calculation of CoM displacement   Based on the inverted pendulum model and under the assumption that the sway angles are small in quiet standing conditions, Palmerini et al. [146] computed the equation that relates the acceleration to sway angle and CoM height in the sagittal plane (see Fig. 4.6)

$$a(t) = h\ddot{\theta}(t) - g\sin\theta(t) \approx h\ddot{\theta}(t) - g\theta(t) \tag{4.10}$$

where g is the gravitational acceleration, $a(t)$ is the acceleration near the CoM in the AP direction, $\theta(t)$ is the sway angle with respect to vertical, and h is the height of the inertial sensor (which is assumed to be the same height as the CoM estimate). The corresponding transfer function is

$$H(s) = \frac{\theta(s)}{a(s)} = -\frac{1}{g - hs^2} \tag{4.11}$$

Equation 4.11 can be written in the frequency domain as

$$H(jw) = -\frac{1}{g + hw^2} = -\frac{1}{g}\frac{1}{1 + (w/w_n)^2}, \qquad w_n = \sqrt{\frac{g}{h}}. \tag{4.12}$$

On average one can assume h approximately equal to 1 meter for an adult person (the average leg length for adult male humans ranges between 0.88-0.92 m [128, 195]). Therefore the displacement of the CoM on the AP direction ($x = h\sin\theta$) can be approximated as $x = \sin\theta \approx \theta$.

Based on equation 4.12 Palmerini et al. [146] used a low-pass filtering with a cut off frequency of 0.5 Hz and static gain of $-1/g$ in order to estimate x. They derived some features from the calculated CoM displacement using this model although they did not validate it. They mentioned that the intention was not to achieve a precise estimation of the CoM position but rather a signal which approximates the characteristics of the CoM displacement. In their paper they did not provide the type of filter used in the analysis.

Here, the purpose is to validate this model using different filter techniques and show how precise it is in comparison with the double integration method.
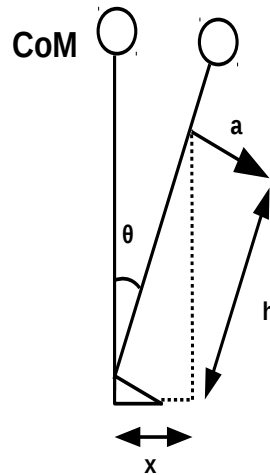

Figure 4.6.: Simplified inverted pendulum model for standing posture. Adapted from: [146]

4.2.3. Experimental Setup

Video Tracking System

The gold standard for the displacement of CoM in the mediolateral axis was obtained by tracking the movement of the sensor over time using a high speed camera (Go Pro HERO3, 100 fps,

1280x960 pixels resolution) and a video tracking software (Tracker v4.85 [28]) as shown in Figure 4.7.

Synchronization of camera and accelerometer was accomplished by filming a tap on the sensor and subsequently matching the maximum in the acceleration signal with the frame that displays the instant of the impact.
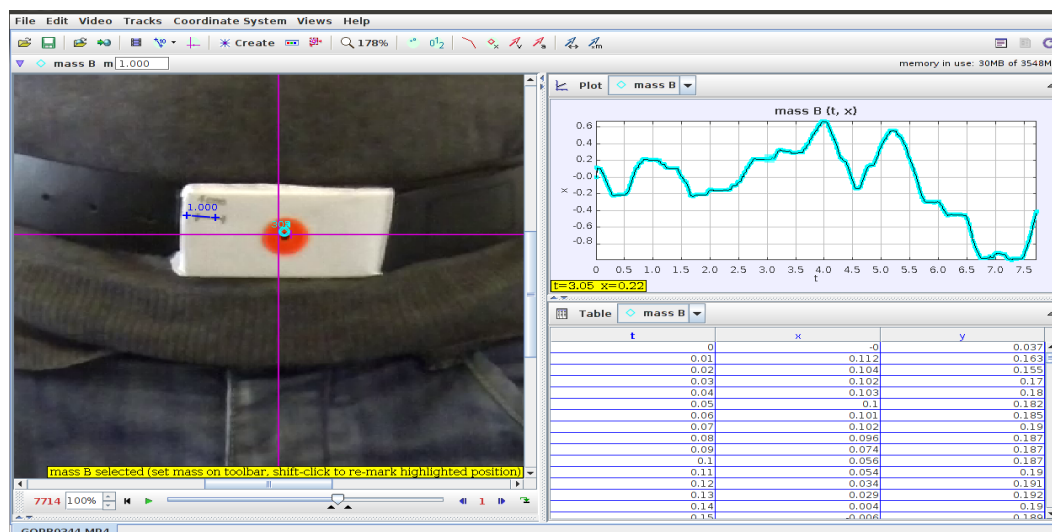

Figure 4.7.: Video tracking of estimated CoM displacement in the mediolateral direction during quiet standing

Data Collection Protocol

Four standardized clinical tests: Romberg test (T1), Tandem test (T2), One-legged with Eyes Open (T3) and One-legged test with Eyes Closed (T4), were performed during 10 seconds by two healthy subjects, one 30-year-old female (S1) an one 31-year-old male (S2). Beginning and end of each test was indicated by a single and a double tap on the sensor. The sensor was firmly fastened around the waist and centered at the middle of the mediolateral axis near the center of mass of the body. The high speed camera was fixed a few centimeters away in front of the sensor.

4.2.4. Data Processing and Analysis

All video and acceleration data were annotated, synchronized and stored in single files corresponding to each balance test. The duration of all tests was adjusted to 8 seconds, discarding the first and last second to ensure that the participant was in a stable position and avoid undesired effects, like for example in the one-legged test, the leg lifting movement at the beginning of the test. In addition to that, the reference signal extracted from the video recording was smoothed and the raw acceleration signals were corrected for sensor tilt (see section 2.4.1 for more details) and denoised by applying a low-pass Butterworth filter (4th order). The cut frequency of the filter, $f_{cutoff} = 4$ Hz, was chosen to preserve the frequency spectrum of the trunk sway, estimated to occupy the frequencies in the band 0-3.5 Hz [45, 154, 146], and get rid of eventual electrical-thermal sensor noise [113] or muscle tremors typically present in patients with neurological disorders like for example Parkinson's disease [146]. The data analysis was entirely done in R-2.14.1 [174] and two different types of data processing were performed depending on the method used to calculate the CoM displacement. Both methodologies are described in the following paragraphs:

Method 1: De-trend and double integration   To estimate the CoM displacement according to the first method – double integration of the movement-related acceleration – the raw data was first detrended employing a Savitzky-Golay polynomial filter of high order ($N = 6$) by subtracting the low-frequency trend to the raw data. Filters of lower and higher order were employed but discarded from the final analysis since the CoM estimation was less accurate. After the filtering, the de-trended acceleration was integrated twice using Simpson's rule approximation, first to obtain the velocity $v(t)$ and a second time to calculate the displacement $x(t)$. The flowchart in Figure 4.8 shows the data processing performed on the raw data.
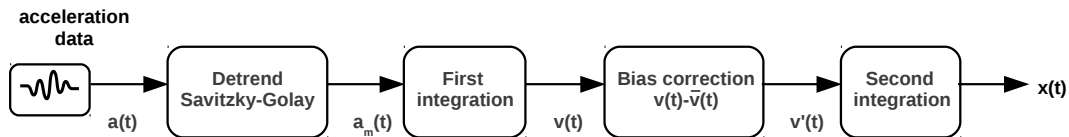

Figure 4.8.: Estimation of CoM displacement using to the double integration method

In order to find the optimal 3dB cutoff frequency that maximizes the similarity between the estimated displacement and the actual one, an exhaustive search was done for frequencies in the interval $[0.11, 0.25]$ Hz, in steps of 0.01 Hz. Frequencies out of this interval produced a lower similarity and therefore they were excluded. The similarity was measured with the metrics specified in table 4.1.

Similarity metric	Definition
DP2P	Absolute value of the difference between the maximum peak to peak amplitude of the reference signal and the maximum peak to peak amplitude of the estimation. Units: millimeters
MNCC	Maximum value of the normalized cross correlation. Calculated for a lag of $\pm$ 1 second and normalized between 0 and 1

Table 4.1.: Similarity metrics calculated to compare the reference and the estimated CoM displacement .

Method 2: Anthropometric low-pass filter based on the inverted pendulum   The second method to estimate the CoM displacement consists of performing low-pass filtering with a filter that approximates the frequency response

$$H(jw) = -\frac{1}{g}\frac{1}{1 + (w/w_n)^2}, \qquad w_n = \sqrt{\frac{g}{h}}.$$

In order to do that, two different low-pass filters were used, a second order Butterworth and an Elliptical filter. The filters were chosen for their different characteristics of the frequency response, the Butterworth has a smooth response with a slow roll off whereas the Elliptical has a near-instant cut off but it presents ripples in both the pass-band and the stop-band. The frequency response of these filters can be seen in Table 4.2 and Figure 4.9.

For both filters the range of frequencies that best approximated the reference signal lay in the interval $[0.45, 1]$ Hz. The optimal frequency was found using the same procedure as in the first method and the search was done in steps of 0.05 Hz.

Filter	Frequency response
Butterworth	$$H(jw) = \frac{1}{\sqrt{1 + (w/w_c)^{2n}}}$$ where n represents the filter order and w_c is the cuttoff frequency at 3dB
Elliptical	$$H(jw) = \frac{1}{\sqrt{1 + \epsilon^2 R_n^2(\xi, w/w_c)}}$$ where R_n is the nth-order elliptic rational function, ϵ represents the ripple factor, ξ is the selectivity factor and w_c is the cuttoff frequency

Table 4.2.: Frequency response of a Butterworth and Elliptical filter.
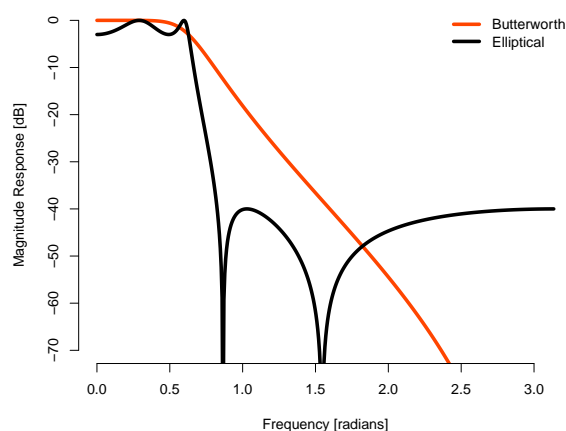

Figure 4.9.: Frequency response graph of a Butterworth and Elliptical filter.

4.2.5. Results

Optimal Cuttoff Frequency

To estimate the optimal cuttoff frequency (f_c), I used the difference of maximum peak to peak amplitudes (DP2P) and the maximum normalized cross correlation (MNCC) described in Table 4.1. Figure 4.12 shows the results for each frequency interval and methodology. Looking at the graphs one can conclude that there is not a unique frequency that overall maximizes the similarity, but the values along the different frequencies either present a increasing/decreasing step-like response or an almost flat response. Table 4.3 presents the median values of the frequencies that maximize the cross correlation and the median values that minimize the difference in peak to peak amplitudes.

Based on these values and taking into consideration that to give a reliable assessment of the postural stability the priority is to detect accurate changes in the amplitude of the CoM displacement rather than to obtain the best waveform match, the optimal cut off frequencies were estimated as 0.16 Hz for the Savitzky-Golay filter and 0.50 Hz for the Butterworth and Elliptical filter (the average of the DP2P optimal f_c of both filters).
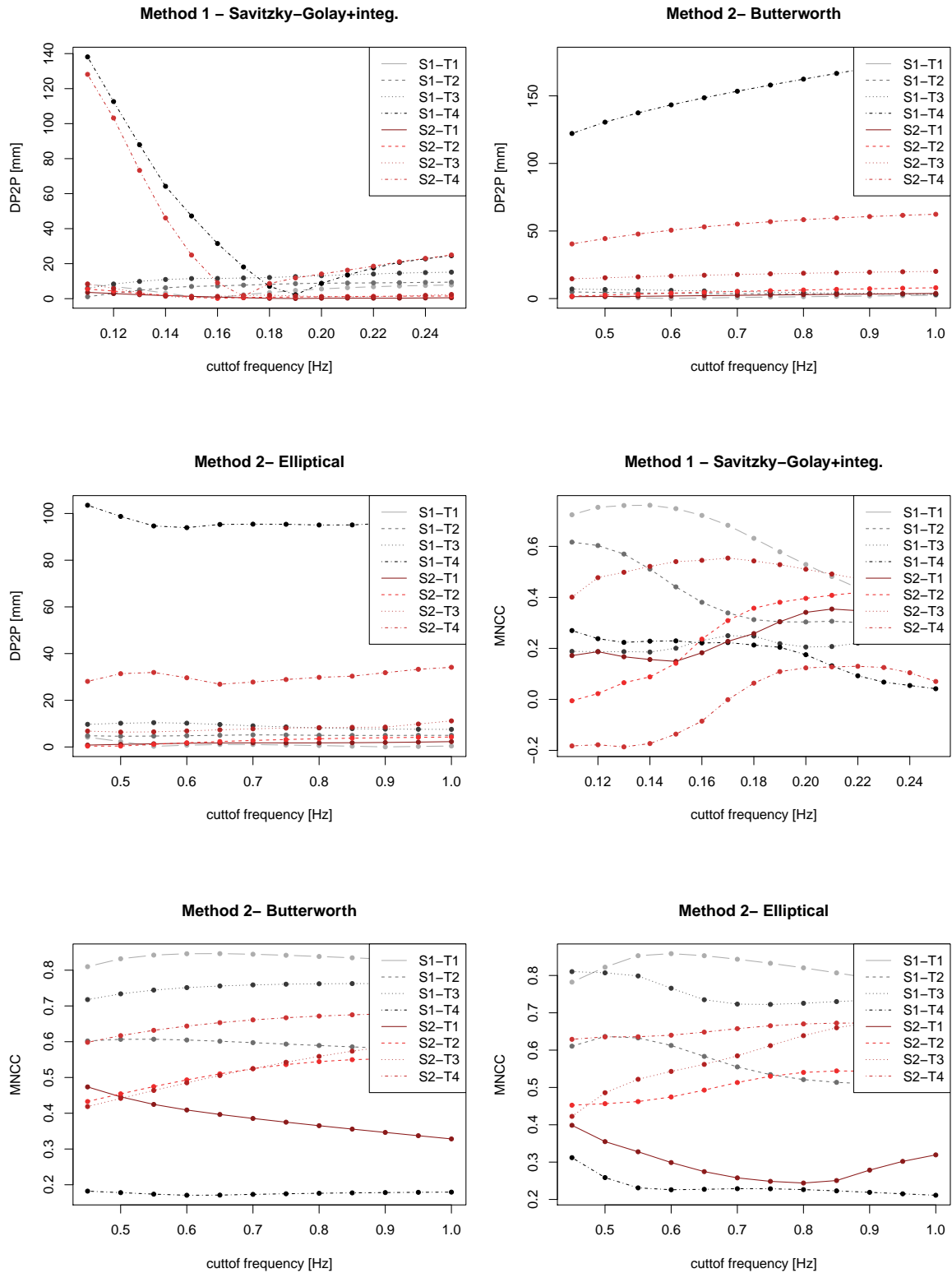
Figure 4.10.: Similarity measures between the estimated and measured CoM displacement for each subject (S) and test (T) and for different cuttoff frequencies and methods.

Method	median(argmax $MNCC(f)$)	median(argmin $DP2P(f)$)	Estimated optimal f_{cutoff}
1 - Savitzky-Golay+integ.	0.18 Hz	0.16 Hz	0.16 Hz
2 - Butterworth	0.65 Hz	0.45 Hz	0.50 Hz
2 - Elliptical	0.55 Hz	0.55 Hz	0.50 Hz

Table 4.3.: Optimal cuttoff frequencies for the maximum normalized cross correlation and the peak to peak difference between the estimated and the measured signal

Estimation Error

The error made in the estimation of the CoM displacement was measured using the DP2P and the MNCC. The DP2P represents the error concerning the amount of trunk sway and the MNCC measures the deviation of the estimated CoM trajectory with respect to the real one (ignoring the differences of the offset between the estimated and the gold standard signal). Table 4.4 presents the average and standard deviation of both parameters for the optimal frequencies determined in the previous paragraph, $f_{cutoff} = 0.16$ Hz and $f_{cutoff} = 0.50$ Hz for Method 1 and Method 2 respectively.

To measure the level of difficulty of each test the maximum trunk sway has been calculated (i.e. the peak to peak amplitude of the measured CoM displacement). The bigger this parameter is, the more complicated to maintain the equilibrium. Figure 4.11 shows how the difficulty increases with the test number.
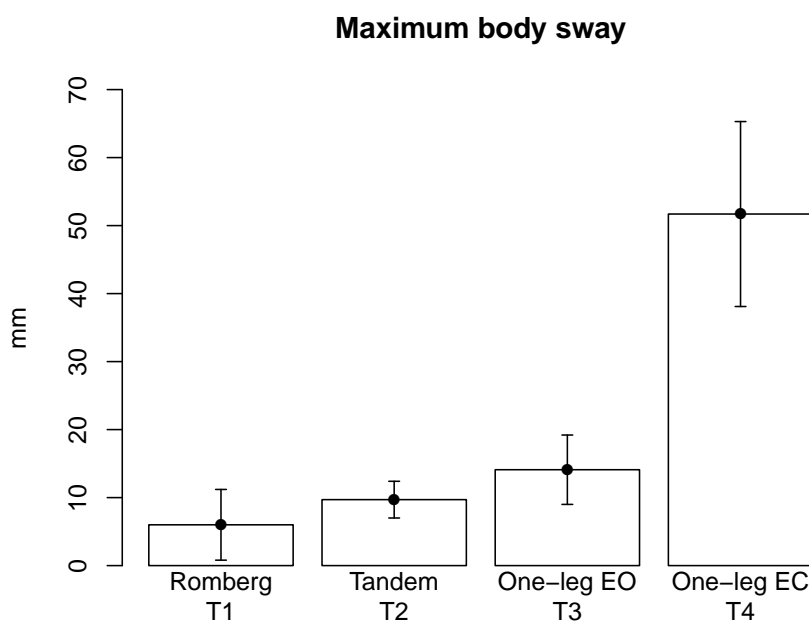

Figure 4.11.: Average and standard deviation of the maximum trunk sway calculated using the gold standard CoM displacement

In the diagram bar of Figure 4.12 one can see how the average error in the estimation of the amount of sway increases exponentially with the difficulty of the test. This is equivalent to say that the bigger the trunk sway is, the bigger the error made in the estimation (Fig. 4.13). On average, the difference in the peak to peak amplitude between the real and the estimated

CoM displacement was in the range of 0.001-1 cm for the Romberg, Tandem, and One-legged test with Eyes Open (EO) and in between 2 cm and 8.7 cm for the One-legged test with Eyes Closed (EC). Overall the most accurate method to calculate the trunk sway was the Method 1 (double integration method) whereas for the Method 2 (anthropometric filtering) the Elliptical filter performed better in general than the Butterworth filter.

The maximum normalized cross correlation (MNCC) between the estimated and the measured signal was nearly constant for the three first balance tests with an average value of around 0.4 for the Method 1 and 0.55 for the Method 2. In the most difficult test, the One-legged EC, the similarity decreased in average 0.25 units regardless of the method used.

Method	Romberg (T1) n=2		Tandem (T2) n=2		One-legged EO (T3) n=2		One-legged EC (T4) n=2	
	DP2P mean ±sd	MNCC	DP2P mean ±sd	MNCC	DP2P mean ±sd	MNCC	DP2P mean ±sd	MNCC
Method 1-Savitzky-Golay	0.8±0.2	0.5±0.4	3.7±5.0	0.3±0.1	6.4±7.3	0.4±0.2	20.3±15.2	0.1±0.2
Method 2-Butterworth	1.4±0.2	0.6±0.3	3.5±1.4	0.5±0.1	11.1±6.1	0.6±0.2	87.5±60.9	0.4±0.3
Method 2-Elliptical	1.7±0.9	0.6±0.3	2.5±3.0	0.5±0.1	8.3±2.7	0.6±0.2	65.1±47.6	0.4±0.3

DP2P and DMEAN are given in millimeters
MNCC is unitless

Table 4.4.: Similarity values between the estimated and the measured CoM displacement

The standard deviation measures the consistency of the results. In this experiment, the error of the maximum trunk sway for the easier tests (Romberg, Tandem and One-legged EO) has a relatively small standard deviation indicating that our estimation is consistent regardless of the participant and as long as the trunk displacement is below approximately 2 cm, the maximum trunk displacement for this set of tests. On the contrary, the trunk sway error for the most complicated test (One-legged EC) presents a high standard deviation which means that the precision of the estimation can vary considerably depending on the subject or trial. The same occurs for the standard deviation of the cross correlation in all tests.

4.3. Comparison of Accelerometer versus Pressure Plate

4.3.1. Experimental Setup

Four participants, 3 males (34.7 ± 11.9 years old) and 1 female (29 years old) stood barefoot on a pressure plate with both feet together and the accelerometer attached to the waist near the CoM. In this position, they performed 5 balance tests simulating 5 increasing levels of difficulty during approximately 10 seconds. Level 1 consisted on standing still (Romberg test) and level 5 consisted of generating as much body movement as possible but still without losing balance or stepping out of the base of support. The duration of the tests was determined by the maximum recording time permitted by the pressure plate (RSscan-0.5m Advanced footscan®) configured at 100Hz to match the acceleration sensor sample frequency. To synchronize acceleration and pressure data, the participants jumped on the pressure plate at the beginning of each test in order to generate a maximum on the net force signal as well as in the vertical acceleration component with the first contact of the feet with the floor. The first 1,5 seconds of recording after the maximum peak where discarded of the analysis and both signals length was truncated to 6 seconds. All participants gave written informed consent.
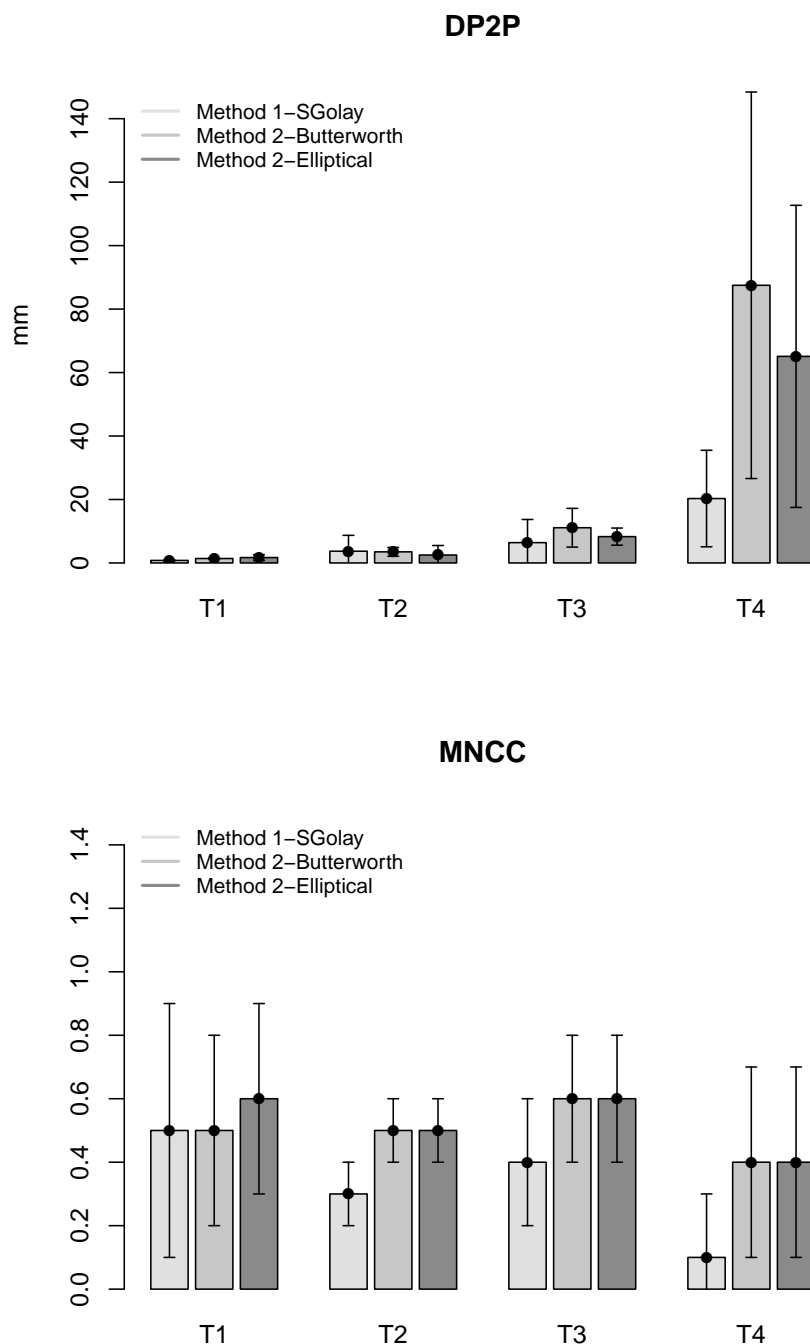
DP2P

MNCC

Figure 4.12.: Graphs of the similarity values (means and standard deviation) between the estimated and the measured CoM displacement

4.3.2. Extraction of Stabilo-Parameters

The parameters chosen for the comparison between the accelerometer and the pressure plate are: a) the overall area covered by the CoP displacement and b) the distance travelled. These parameters are among the major outcomes of the commercial force-sensing plates. The overall area covered by the CoP is an area containing the 95% of the CoP samples lying in the transverse plane, often reported in the clinical literature as predictor of the risk of falls [120, 175], and the

CoM position [mm] – Romberg test

(a)

CoM position [mm] – One–legged test EC
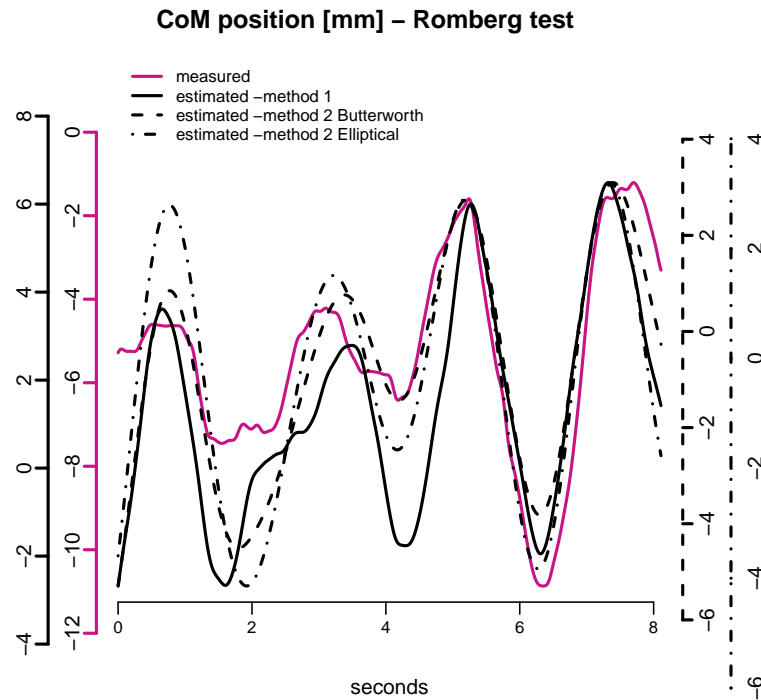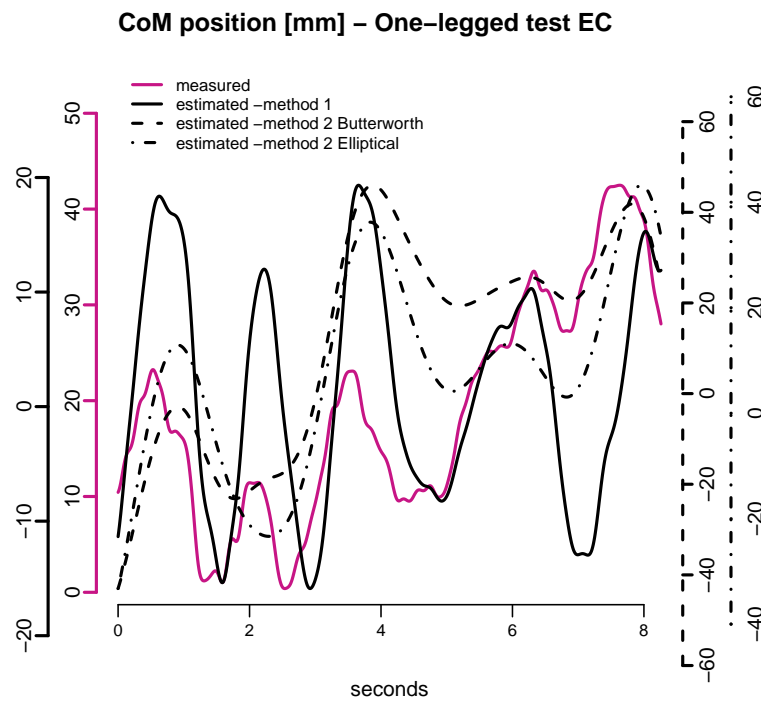
(b)

Figure 4.13.: Estimated and measured CoM position (in milimeters) along the mediolateral axis for (a) Romberg test and female subject and (b) one-legged EC test and male subject (right).

distance travelled is the length of the path described by the CoP. Both parameters are closely linked to the stabilogram, an illustrative graph of the postural sway.

Stabilogram and Stabilo-Ellipse

The trajectory of the center of pressure along the transverse plane during a balance test is commonly called a stabilogram (Fig. 4.14). Even without numerical analysis, the stabilogram provides a wealth of information on postural sway. For example, the overall area covered by the CoP indicates the ability of the subject to maintain a stable upright posture; the shape of the stabilogram might suggest that sway is greater in either the medial-lateral or the anterior-posterior direction, and long straight segments on the plot reflect sudden perturbations or corrections to balance, whereas shorter segments indicate finer control, even if the overall plot covers a large area [149].

The stabilo-ellipse is a new term employed in this work to refer to the stabilogram described by the low frequency components of the CoM acceleration. For the calculation of the stabilo-ellipse the acceleration data was low pass filtered with a fourth order Butterworth filter ($fc = 2$ Hz).
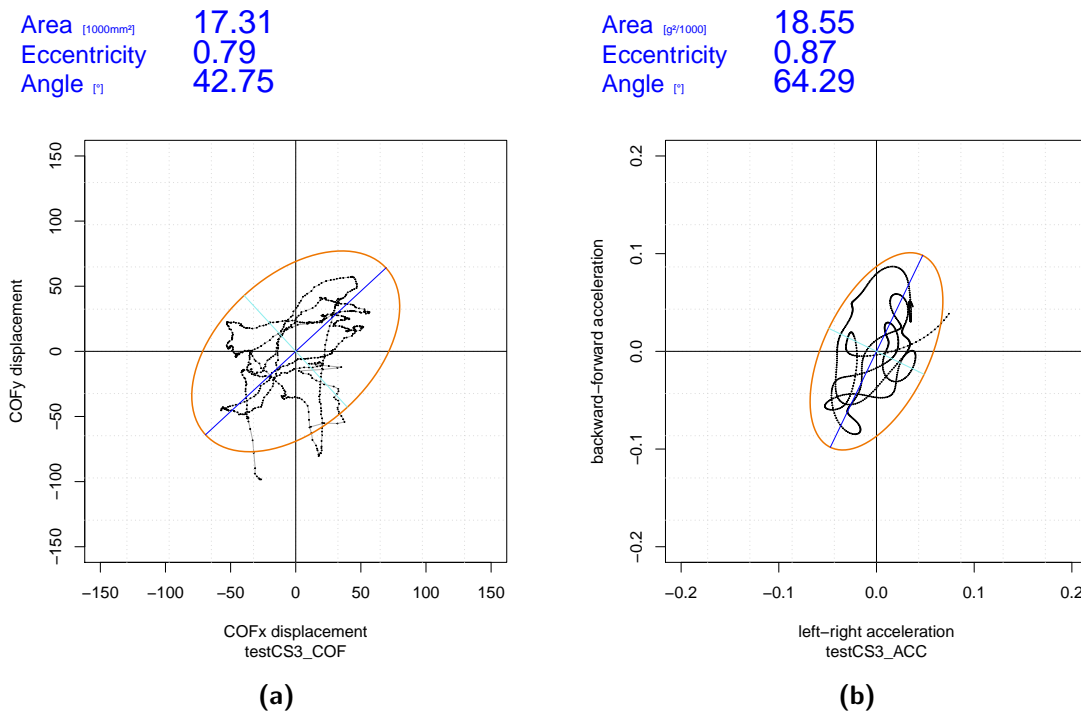

Figure 4.14.: Examples of (a) stabilogram and (b) stabilo-ellipse during the performance of a balance test (level 3) for the same individual. The orange line is the 95% Confidence Interval Ellipse.

Calculation of the 95% Confidence Interval Ellipse   The calculation of the 95% Confidence Interval Ellipse is based on *Principal Component Analysis*. Given two random variables S_1 and S_2 with a bivariate normal distribution and means μ_1 and μ_2, the parametric form of the ellipse's equation can be calculated as [194, 112]

$$x(t) = x_0 + \mathsf{F}(\sqrt{\lambda_1}\cos t \cos\phi - \sqrt{\lambda_2}\sin t \sin\phi) \text{ and}$$
$$y(t) = y_0 + \mathsf{F}(\sqrt{\lambda_1}\cos t \sin\phi + \sqrt{\lambda_2}\sin t \cos\phi), \text{ with} \qquad (4.13)$$
$$\phi = \arctan \mathsf{v}_{21}/\mathsf{v}_{11}.$$

Where (x_0, y_0) are the center of the ellipse, λ_1 and λ_2 are the eigenvalues of the covariance matrix $\Sigma = \mathsf{cov}(\boldsymbol{s_1} - \mu_1, \boldsymbol{s_2} - \mu_2)$, $\phi =$ is the angle between the X-axis and the major axis of the ellipse, F is a scale factor given by a chi-squared distribution with 2 degrees of freedom and p-value equal to 0.95, and v_{21} and v_{11} are eigenvectors of the covariance matrix Σ.

Area of ellipse   The area of the 95% Confidence Interval Ellipse is equal to

$$A = \mathsf{F}^2 \pi \sqrt{\lambda_1} \sqrt{\lambda_2} \; . \tag{4.14}$$

Path Length   The Path Length (sometimes also called Travelled Way) is calculated as

$$L = \sum_{i=1}^{N} \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} \; . \tag{4.15}$$

with x the samples along the X-axis (mediolateral axis) and y the samples along the Y-axis (anterio-posterior axis).

4.3.3. Results

As expected, the area of the ellipse and the Path Length increased with the difficulty level of the test. However the increasing rate was different for each data set, for the acceleration data the evolution was exponential whereas for the CoP data it followed a linear trend, as shown in figure 4.15.

A pearson's correlation analysis showed that a statistically significant correlation exists among the parameters extracted from accelerometer and pressure plate, with a correlation coefficient r = 0.61 (p<0.01) for the 95% Confidence Interval Ellipse and r = 0.73 (p<0.001) for the Path Length. Moreover, the accelerometer and pressure plate have both the capability to distinguish between difficulty levels (Table 4.5).

	Acceleration	Center of Pressure
Area	p<0.001	p<0.01
Path Length	p<0.001	p<0.001

Table 4.5.: p-values of Anova test shows the capacity of accelerometer and pressure plate to distinguish between test difficulty levels.

4.4. Summary of Findings

In the first section of this chapter, I have investigated the accuracy in the estimation of the lateral CoM displacement using a single acceleration sensor attached to the waist and comparing two different methods, the double integration method (Method 1) and the anthropometric filtering (Method 2). Although these methods have been previously used in some studies to estimate the CoM displacement [130, 21, 50, 97, 146], none of the existing publications have reported accuracy values. In the second part of the chapter, I demonstrated the validity of the accelerometer to distinguish among different amounts of body sway and compared it with a force-sensing plate, the standard tool used in laboratory settings to quantify postural sway.

I have found that the error made in the estimation of the trunk sway increases with the amplitude of the trunk movement or in an equivalent manner with the difficulty of the balance test. This is most likely due to the fact that under these circumstances of higher instability the person does not behave as an inverted pendulum anymore. When the amplitude of the body movement is very small, like for example during double or tandem stance, most of the time the body is in equilibrium controlled through an ankle strategy [80] and it can be assumed that the person stays standing fairly straight as an inverted pendulum. However, in those more challenging tests there is a need to adjust rapidly the CoM to a balanced position and this is done through a hip strategy characterized by a large angular trunk acceleration [80]. This rotational movement
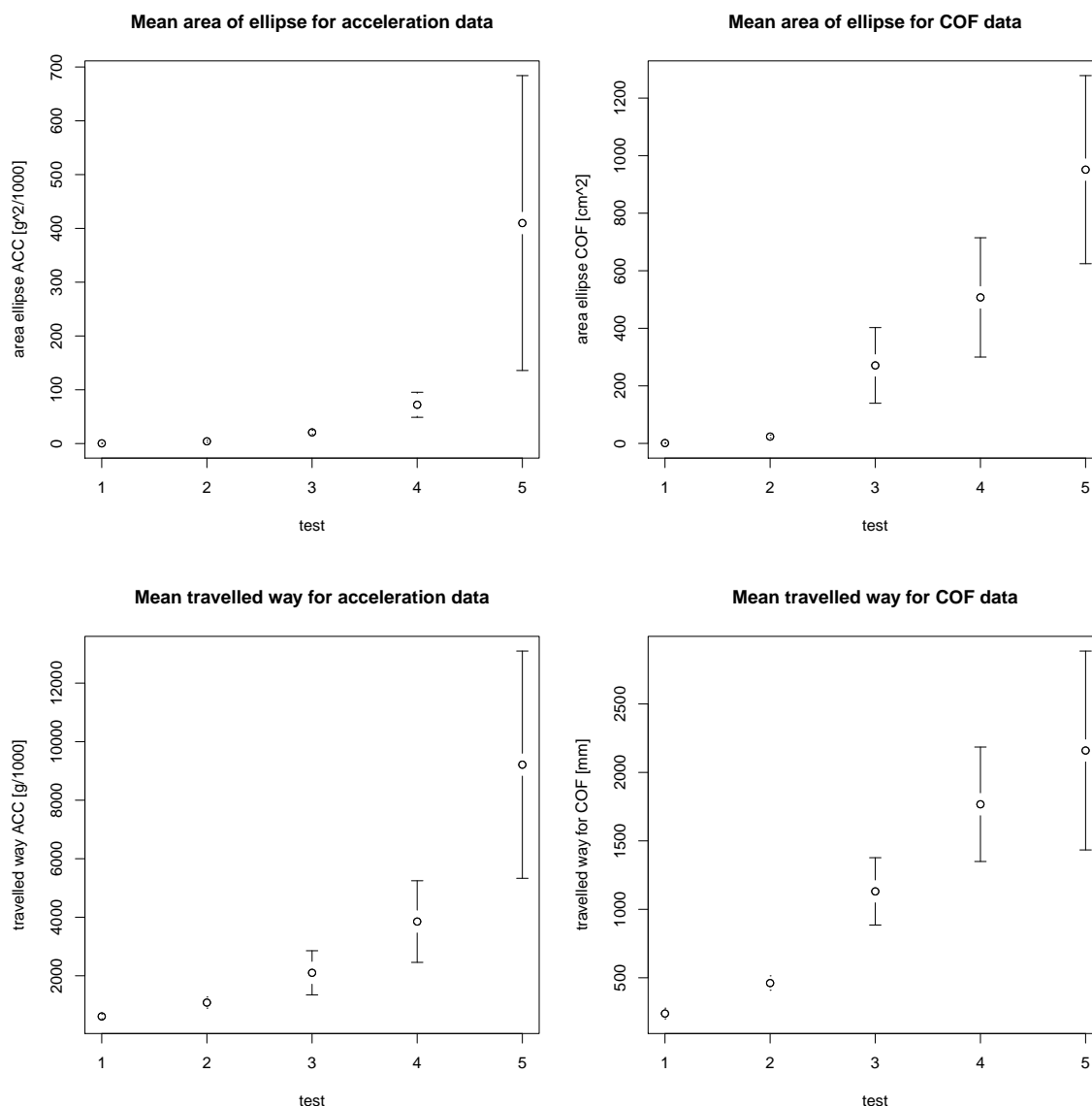
Figure 4.15.: Comparison of pressure plate (right) and accelerometer-derived (left) balance parameters.

of the hip invalidates the inverted pendulum model reducing the accuracy of Method 2 and introduces a strong unsteady gravitational component difficult to remove before integration in Method 1. In quasi static conditions, the body sway estimation, regardless of the used method, is quite accurate especially if we take into account that the trunk displacement is very small (just of the order of 1-2 cm) and so it is the movement-related acceleration in comparison with the noise.

In general, the double integration method estimated the maximum trunk sway better than the method based on the inverted pendulum, but the waveshape similarity with the reference signal using this first method was lower. A limitation of the first study was the small number of participants, thus these results should be further investigated using more and varied data. A possible way to improve the estimation could be to use a gyroscope in parallel with the 3D acceleration sensor.

When comparing the accelerometer and the pressure plate results, I found that the graph described by the accelerometer-derived parameters (area of the 95% Confidence Interval Ellipse

and the Path Length) followed an exponential curve for an increasing body sway, in contrast with the linear tendency described by the same parameters extracted from the pressure plate data. The assumption here is again that the stiffness control of balance in quiet standing [199], the one which governs the inverted pendulum model, is not valid for big perturbations of the body center of mass but a double-link model is needed [38]. Under situations of high instability, the neuromuscular system activates the hip flexor muscles to bring the CoM to a balanced position. The body sway measured by means of the CoP displacement does not account for the joint torques of the hip as much as the accelerometer placed near the body CoM does. The additional amount of sway registered by the accelerometer when the hip rotates as the difficulty of the test increases, depicts the exponential curve.

Another interesting point is that the Path Length may be a better parameter to distinguish among balance ability when the perturbation is very small. The results show how for the first two tests the difference in the Ellipse's area is, although statistically significant, very small, whereas for the Path Length it increases (Fig. 4.15). This fact indicates that, although the body sway covers the same area, the velocity with which the CoP and the CoM moves can be different.

Finally, the standard deviation increases overall with the difficulty of the test. This can be due to a difference in the balance ability of the subjects or rather to a difference in the interpretation of the "1-5 levels of difficulty". Scales, especially number scales with no or less descriptors, like the one employed in this study, are very subjective to the respondents [39].

5. Falls Risk Assessment and Functional Tests

5.1. Introduction

Any well designed fall prevention program requires accurate evidence-based measurements to identify persons at higher risk of falling. A correct screening is a first and crucial step for the effectiveness of the intervention. At the most basic level, evidence-based questionnaires are available but ideally more comprehensive clinical examinations including vision test, cognitive status, medication and functional assessment (examination of an individual's mobility and transfer skills) should be adopted [9]. The selected screening tool will certain depend on several aspects, such as, target population, time required to complete the tool, settings, resources and the expertise of professionals available to execute the protocols. That is the reason why fall risk assessment is usually not standardized within or across settings.

In this chapter we are interested in testing the potential of a single waist-worn accelerometer to enhance the outcome of the current functional mobility tests with the aim to improve the accuracy of the fall risk assessment. For that purpose, different prediction models were trained and validated based on classification and regression analysis on clinical and functional accelerometer data from a group of elderly prospective fallers and non-fallers. Participants were recruited in a multicenter study as part of the VPHOP EU-Project [176].

Throughout the chapter I will refer to "fall prediction" as the actual probability of presenting a fall, estimated using our existing data set, and to "falls risk assessment" as the holistic evaluation of the the risk of falling. The aim here is to improve the fall risk assessment by providing valuable tools, methods and results, among them, new validated functional parameters and the outcome of the fall prediction in a specific cohort of patients.

5.2. Fall Definition and Questionnaires

Since the creation of the first report on prevention of falls (the *Kellogg* report, 1987) until the last *Global Report on Falls Prevention in older Age* of the WHO (2007), many researchers have sought to describe what they mean by fall with the aim of: a) determining which events could be included as a fall and which not and b) classifying different types of falls in order to allow for comparability between research results. As shown in Table 5.1, over almost 20 years no consensus has been reached and there is still not universally accepted definition for a fall. Since researchers often provide definitions for falls that reflect the needs of their particular studies [203] and each fall can be presented as a unique chronological episode of occurrences comprising not only the fall itself but its antecedents and consequences, such universal definition may not be possible or even necessary.

In addition to a comprehensive fall definition, questionnaires should include information about non-injurious falls such as slips, trips and stumbles since these occurrences may constitute a very valuable early warning for rapid intervention [107]. Details about where, when and the circumstances associated with the fall should also be reported. Lamentably, self-reporting of falls is notoriously inaccurate [119], frequently subject to recall and response bias. In particular, the elderly subjects tend to forget falls that occurred during specific periods of time over the preceding 3 to 12 months [41] while others are often afraid to report falls or have home evaluations because they fear being sent to a nursing home [182].

Author	Year	Selected Fall Definitions
Kellogg Group	1987	"A fall is an event which results in a person coming to rest inadvertently on the ground or other lower level and other than as a consequence of the following: Sustaining a violent blow, Loss of consciousness, Sudden onset of paralysis, as in a stroke, An epileptic seizure." (p. 4)
Lach et al.	1991	"... an unexpected loss of balance resulting in coming to rest on the floor, the ground, or an object below knee level." (p. 198)
Buchner et al.	1993	"Unintentionally coming to rest on ground, floor, or other lower level; excludes coming to rest against furniture, wall, or other structure." (p. 301)
Means et al.	1996	"... any involuntarily change from a position of bipedal support (standing, walking, bending, reaching, etc.) to a position of no longer being support by both feet, accompanied, by (partial or full) contact with the ground or floor." (p. 1032)
Berg, Alessio, Mills, & Tong	1997	"... losing your balance such that your hands, arms, knees, buttocks or body touch or hit the ground or floor." (p. 262)
Canadian Institute for Health Information	2002	"... an unintentional change in position where the elder ends up on the floor or ground."
Carter et al.	2002	"... inadvertently coming to rest on the ground or other lower level with or without loss of consciousness and other than as the consequence of sudden onset of paralysis, epileptic seizure, excess alcohol intake or overwhelming external force." (p. 999)
Cesari et al.	2002	"... a sudden loss of gait causing the hit of any part of the body to the floor .. ." (p. M723)
Tideiksaar	2002	"... any event in which a person inadvertently or intentionally comes to rest on the ground or another lower level such as a chair, toilet or bed." (p. 15)

Table 5.1.: Most diverse fall definitions from research and prevention literature published between 1987 and 2005. Adapted from Zecevic, Aleksandra A., et al. [203]
.

VPHOP Falls Questionnaire

The falls questionnaire used in the VPHOP project did not include any specific definition for a fall. However, details about date, place or circumstances associated with the fall were reported. This information was used to exclude from the list of fallers those patients whose falls were related to extrinsic or accidental factors, like for example, slipping on an icy surface, and just to take into account those who fell due to functional deficits. The diagram 5.1 presents all the information reported in the questionnaire of fall monitoring.

5.3. Fall Risk Factors

Risk factors for falls have been classified into two main categories: **intrinsic** or patient-related factors and **extrinsic** or environment-related factors [181]. Falls often result from multifactorial events, sometimes from a combination of environmental and patient-related factors, from correlation of intrinsic factors (that in some cases may contribute to confounding, e.g., age and sex) or occasionally from reverse causality, where some elements are consequences as well as risk factors for falls – for example, fear of falling may yield a decrease in physical activity levels that can reduce muscle strength and consequently lead to a fall, arising fear of falling even more.

A fall risk assessment is often based on a weighted combination of the most important fall risk factors (Fig. 5.2) and it typically employs specific screening forms, e.g., Morse Fall Scale [132], STRATIFY [142], Fall Risk Assessment Tool [118] or Hendrich Fall Risk Model [78]. These tools
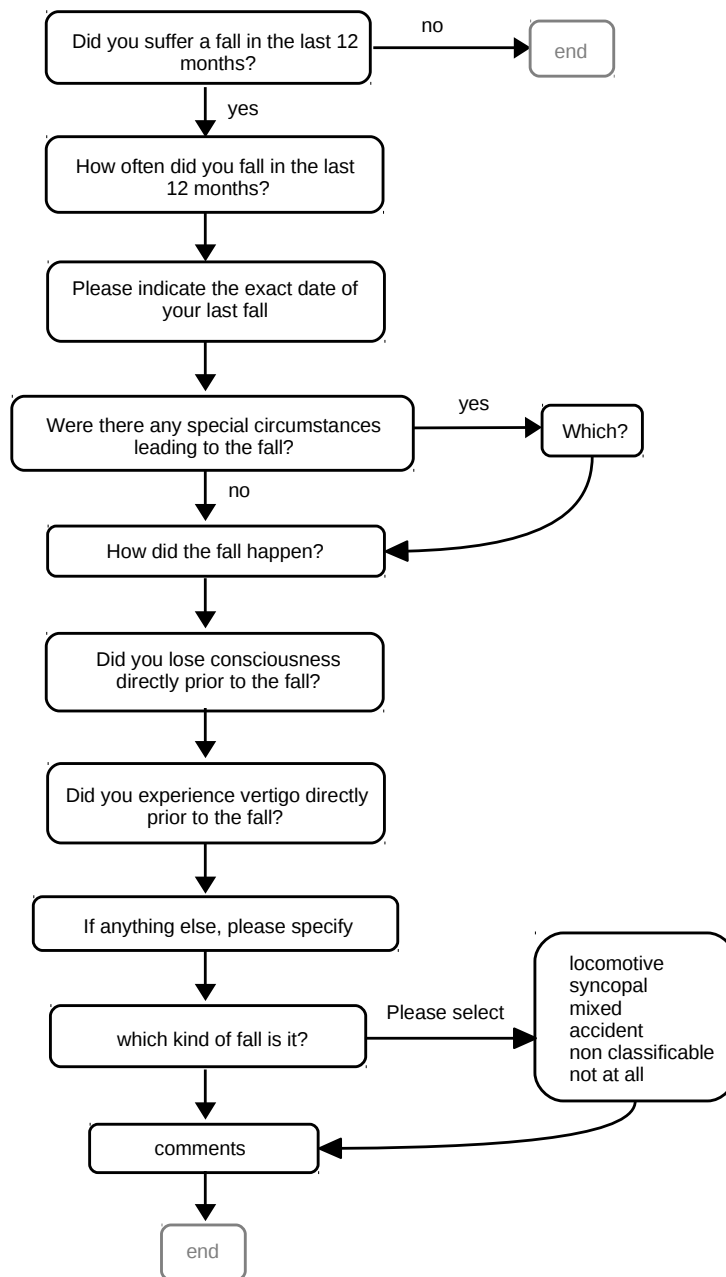
Figure 5.1.: VPHOP questionnaire for retrospective and prospective fall monitoring.

are periodically updated (e.g., per shift, daily or weekly) and take account of intrinsic and medical characteristics of the patients, like psychological and sensory status or mobility dysfunction. Poor scores tend to trigger either further assessment or anticipatory nursing interventions.

5.3.1. Intrinsic Factors

Intrinsic risk factors have been reported to play a dominant role in case of recurrent falls, with the most significant ones those related with mobility problems [171].
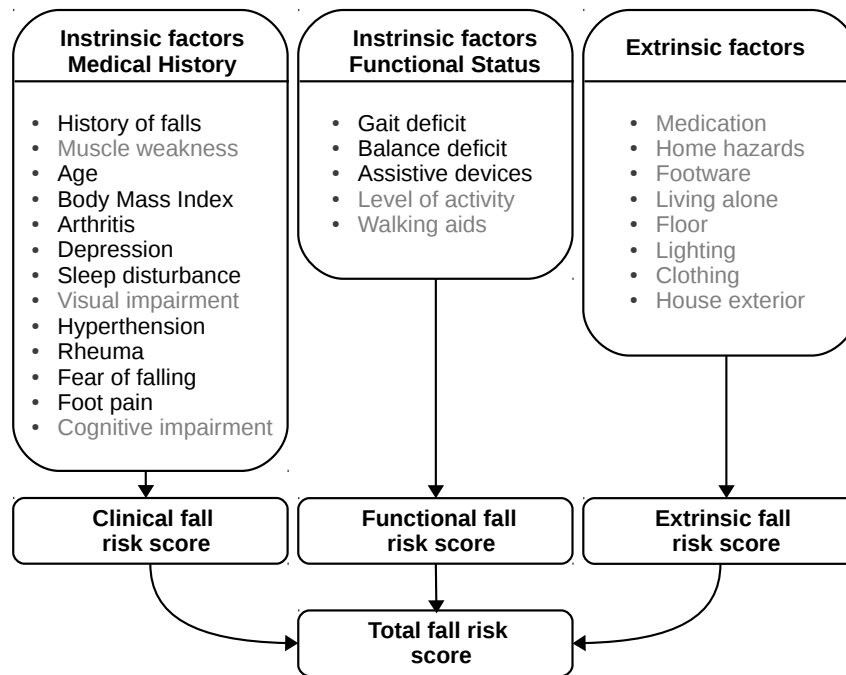
Figure 5.2.: Main fall risk factors reported by the American Geriatrics Society (AGS). Factors included in our fall risk assessment model are in black.

Age   In general, the risk of falling increases with age [181, 46]. However, investigations indicate that superimposed to the decreased postural and gait stability due to the age alone ("presbyastasis") is the increased probability in the elderly of developing specific pathologies which lead to accelerated degeneration in neural and/or musculoskeletal systems that eventually may result in a fall [81, 65, 200]. Age-associated impairments of vision, hearing and memory also tend to increase the number of trips and stumbles [161].

Gender   Most of the reviewed studies found an increased risk for women, both for all fallers and recurrent fallers [46]. Women are also far more likely to incur fractures when they fall [181].

History of falls   History of falls was found to be one of the factors strongest associated with falling, in particular for multiple fallers [170, 32, 46].

Fear of falling   Fear of falling is a consequence as well as a powerful predictor of falls [64, 164, 46]. Up to 85% of recent fallers and up to 50% of those not reporting recent falls acknowledge fear of falling, with the prevalence higher in women than in men. Other consequences of fear of falling include less physical activity, depression, poor postural performance, slower walking speed and muscle weakness [181, 164].

Musculoskeletal condition   Muscle weakness and gait/balance problems are together with history of falls the strongest predictors for falls [32, 161, 170, 46]. Weakness and postural instability stem from age-related degeneration as well as from specific dysfunctions of the nervous, muscular, skeletal, circulatory, and respiratory systems (e.g., Stroke, Multiple Sclerosis, Osteoporosis) or

from simple deconditioning due to inactivity. Plantarflexor and dorsiflexor weakness is a common cause of gait disorders, which affect up to 50% of the elderly population. Case-control studies reported that over two thirds of individuals who have fallen have substantial gait disorders, considerably higher than control subjects who have not experienced a fall [161]. When strength, endurance, muscle power and hence function declines sufficiently, one is unable to prevent a slip, trip or stumble becoming a fall. Difficulty in rising from a chair is also associated with increased risk [181].

Cognitive and dual sensory impairment   Visual acuity, contrast sensitivity and depth perception are the most relevant functions to help maintaining postural control and avoiding a fall [181, 170, 32]. Cataract, glaucoma and macular degeneration are risk factors for falls since they can affect the correct functioning of these visual functions [181] as well as bifocal glasses which impair depth perception and edge-contrast sensitivity hampering the person to focus on their feet and the ground surface to stabilize themselves in case of lost of balance [181, 170]. Hearing impairment negatively affects the ability of a person to orientate in space, hence rising the probability to fall. More than 50% of the elderly have some degree of hearing loss [161, 170]. Dementia and cognitive impairment are also risk factors for falls due to several intrinsic and extrinsic factors including impaired visuospatial perception, neurocardiovascular instability and medication use [161, 170]. Low scores on a short mental status questionnaire is associated with increased risk [181]. Residents in institutional care with diagnosed dementia fall twice as often as those with normal cognition [181]. Fall prevention programs in this population have not shown reduction in falls [89, 167] probably due to the inability to learn and remember new information and/or comply with prolonged exercise regimens.

Medical conditions   Neurological disorders –e.g., Parkison, Multiple Sclerosis –, bone and joint disease like Osteoporosis or Osteoathritis and cerebrovascular accidents, increase the likelihood of a fall through multiple causes including pain, muscle weakness and stiffness, infection, metabolic disorders or decreased proprioception [170]. Diziness and vertigo is common in fallers. However, diziness is a non-specific symptom and may reflect problems as diverse as cardiovascular disorders, hyperventilation, orthostasis, drug side-effect, anxiety or depression. Circulatory disease, chronic obstructive pulmonary disease, depression and arthritis are each associated with an increased risk of 32%. Thyroid dysfunction, diabetes and arthritis leading to loss of peripheral sensation also increases risk. The prevalence of falling increases with rising chronic disease burden. Incontinence is also frequently present in populations of fallers [181, 161].

Sedentary behaviour   Sedentary lifestyle may cause atrophy of muscle and an unstable joint through disuse. Inactive elderly fall more than those who are moderately active or very active in a safe environment [70]. Those subjects cutting back on normal activities because of a health problem in at most 4 weeks previous to fall are at increased risk [143].

Walking aids   Use of walking aids was associated with up to a 3-fold risk of falling [46].

5.3.2. Extrinsic Factors

Previous studies have found that extrinsic precipitating causes are responsible for up to 55% of the falls in the elderly [105, 30]. Extrinsic risks include:

Medication   Use of sedatives, antihypertensives, and, in particular, antiepileptics are directly associated with risk of falling, as well as the number of medications used [46]. The use of four or more medications is associated with a nine-fold increased risk of cognitive impairment and fear
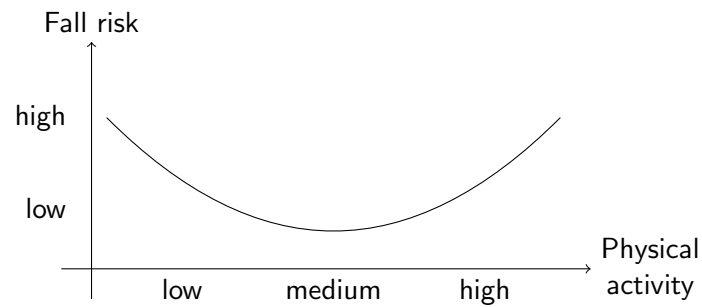
Figure 5.3.: U-shaped graph for activity versus fall risk [14]

of falling [181]. Drugs frequently have side effects that result in impaired mentation, stability, and gait [161].

Home hazards   Environmental hazards such as slippery floors, uneven surfaces and unstable furniture may challenge the subjects's ability to maintain postural stability. Many of those hazards are modifiable through the adoption of some specific environmental improvements like adequate lighting, bathroom grab rails and raised toilet seat, secure stairway banisters, raising or lowering bed and an easily accessible alarm system [161].

Footware and Clothing   Walking indoors barefoot or in socks and walking indoors or outdoors in high-heel shoes have been shown to increase the risk of falls in older people [127]. Baggy clothes which bunch up or drag on the ground can also contribute to risk of falling.

Living alone   Living alone has been shown to be a risk factor for falls, although in the studies which investigated this factor it was reported that when age and gender were controlled for, the association became much weaker [55] and also part of this effect appeared to be related to certain types of housing older people may occupy [181]. An important consequence of living alone is that the result of the injuries can be fateful, in particular if the person is not able to rise from the floor or summon help.

5.3.3. Exposure to Risk

Some studies revealed a U-shaped association between physical activity and risk of fall [143, 131], i.e., those individuals at higher risk of falling are those most inactive and those most active, as shown in Figure 5.3. Individuals who remain inactive fall more than the ones who are moderately active since their balance, flexibility and muscle strength required to counteract postural instability deteriorates. Paradoxically, too high levels of activity increases risk by increasing the exposure to opportunities to fall.

5.4. Functional Clinical Tests

Screening tools for falls risk assessment usually include a battery of functional tests that are typically performed in a clinical environment and administered by qualified care givers. These tests attempt to identify functional limitations in gait and balance and consist of a series of somewhat demanding physical activities, like for example, standing from a chair and sitting down a specific number of times or tandem walking along a line. Examples of functional assessment scales are: Tinetti Performance Oriented Mobility Assessment [178], Berg Balance Test [18],

Functional Reach [54], Dynamic Gait Index [193] or Short Physical Performance Battery (SPPB) [187].

In some cases, these tools can be too time-consuming for the practitioner and burdensome for frail patients. However, when administered in an effective manner and to the right populations they have shown to constitute a very important instrument to recognize those persons at high risk of falling and consequently to trigger tailored interventions for fall prevention (e.g., muscle strengthening, gait/balance training, or aerobic exercise).

5.4.1. Acceleration-Data Collection Protocol

One of the first challenges faced when we introduced the use of the accelerometer in the VPHOP functional tests battery was the creation of a data collection protocol that allowed the identification in the raw acceleration data of each test of the sequence. The protocol had to account also for unexpected events during the test performance, like for example, consecutive repetitions of the same test. In order to accomplish that task, we created a protocol (Appendix B) consisting of a sorted list of eleven gait and balance tests, each of them indicated by tapping on the sensor once at the begining of the test and twice at the end.

Tap detection A tap produces a high peak in the backward-forward axis of the acceleration signal, which is detected by a peak detection algorithm as long as there is a phase before and after the tap where the deviation of the signal is very low. The peak detection algorithm I developed to detect the start and the end of the individual tests was based on the DMW algorithm [42] and it was integrated in a server so that when then clinicians uploaded the acceleration data the algorithm could be run automatically over this data and give immediate feedback to the user about the numbers of tests detected, as shown in Figure 5.4. The peak detection algorithm was tested over 387 taps obtaining the following results: precision=99.59%, recall=86.96%, F-measure=95%. Moreover, I inspected each file to ensure that every test had been identified correctly.

Along with the description of the test performance, the protocol included an evaluation form and a registration form with certain demographic and anthropometric patient data. The protocol was checked by the members of the VPHOP consortium responsible for the execution and instruction of the clinical functional tests in a training session taken place at Charité - Julius Wolff Institut with a test subject (Fig.5.5).

5.4.2. Tests Battery

The tests included in the protocol (Table 5.2) were selected by an expert group of clinicians from a list of standardized functional tests that evaluate different characteristic of the patient's muscular strength, coordination and postural control and that are known to be indicators of risk of falling.

5.4.3. Extracted Features

The outcome measures of the functional test comprise a manual record of time and distances annotated by the clinician with the help of a stopwatch and a tape measure (called "classical outcome" in Table 5.2) and some novel features extracted from the acceleration data ("accelerometer-derived outcome"). For the extraction of features I have created specific algorithms in R language, including the step detection algorithm and the code for the calculation of the stabilo-ellipse described in the previous chapters, as well as R-scripts to automate the process. The following lists contain a detailed description of these accelerometer-derived features for each

No.	Test	Description	Classical outcome	Accelerometer-derived outcome
1	10 meter Walk	patient walks 10 meters at self-selected speed		number of steps, gait asymmetry, mean speed, duration [10ms], step length, cadence
2	10 meter Walk	repetition of test 1	duration [s]	
3	10 meter Walk + cognitive task	10 meter walking at self-selected speed while counting backwards from 100 subtracting 7 and stating the numbers out aloud		
4	Tandem Walk	5 meter walk in one line at self selected speed with one foot leading the other	passed yes/no, distance if<5m	body sway, duration [cs]
5	Timed Up and Go (TUG)	rise from a hard surfaced chair with arm and back supports, walk 3 meters, turn around, walk back and sit down again	duration [s]	duration [cs]
6	Chair Rise	stand up from a sitting position and sit down five times as quickly as possible. Record time from initial movement to final standing position	duration [s]	duration [cs]
7	Romberg stance (quiet-standing)	10 sec standing on feet together (ankles touching) and focussing on point on wall at eye height		
8	Semi-Tandem stance	10 sec standing with one foot ahead of the other as if taking a step and focussing on point on wall at eye height	passed yes/no, duration if<10s	stabilo-ellipse (sway area, eccentricity)
9	Tandem stance	10 sec standing with heel of one foot directly in front of the toes of the other foot and focussing on point on wall at eye height		
10	One-legged stance (on right leg)	10 sec standing on right leg and focussing on point on wall at eye height		
11	One-legged stance (on left leg)	10 sec standing on left leg and focussing on point on wall at eye height		

Table 5.2.: Description of the standardized functional tests included in the VPHOP protocol and their outcomes
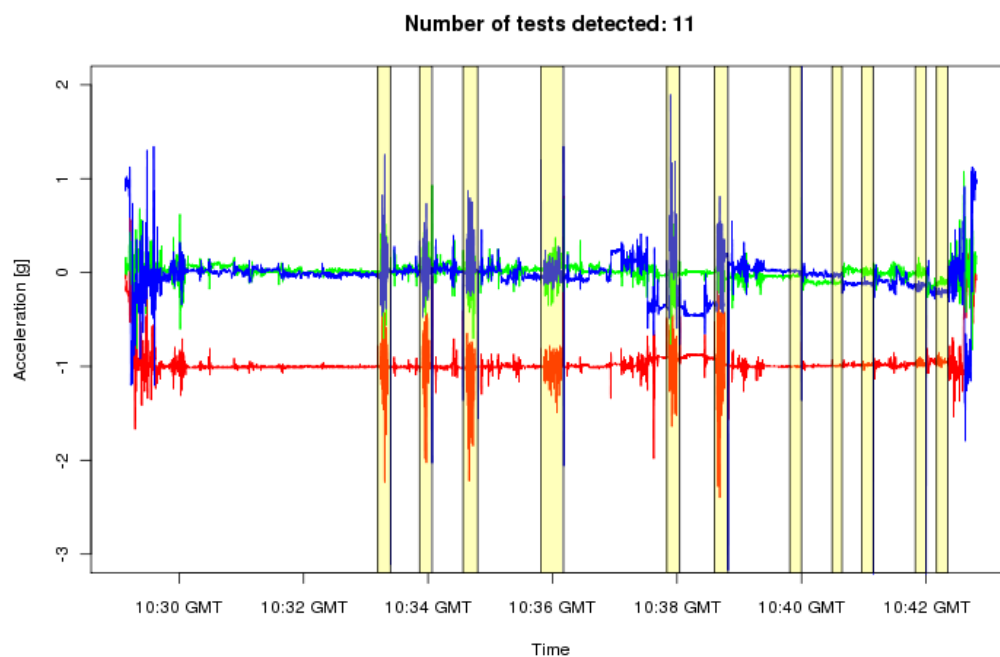
Figure 5.4.: Acceleration signal containing recodings of functional tests. Every test is determined by a single peak in the signal at the beginning of test and two peaks at the end. Tests marked in yellow were detected with a peak detection algorithm.

Figure 5.5.: A clinician instructs the test subject on the performance of the Tandem Walk test at the VPHOP training session for the acquisition of functional tests data with the acceleration sensor.

test or group of tests. For clarity, the X is the axis in the vertical direction, Y is the left-right direction and Z is the direction in the forward-backward axis.

Tests 1, 2 and 3: 10 meter Walk test with/without cognitive task

Number of steps: the number of steps a person took to complete a test computed by the step detection algorithm.

Asymmetry index [%] and standard deviation in X-, Y- and Z-direction : the asymmetry index is calculated based on the mean of the maximum cross-correlation value for pairs of consecutive steps. The minimum number of steps required is 6, and steps detected at the beginning and end of the path are not considered since the irregular walking pattern due to boundary effects could introduce undesired perturbations to the index. The standard deviation of this series of correlation coefficients is also included.

Duration [s]: time needed to complete the test. It is counted from the initiation of movement until the patient stops at the end line by comparison of the body acceleration with an activity threshold.

Mean speed [km/h]: the mean walking speed in Km per hour. It is calculated dividing the distance travelled (10 meters) by the time needed to complete the test.

Step length [m]: average step length, calculated dividing the distance (10 m) by the number of steps.

Cadence [steps/min]: the number of steps per minute, calculated dividing the number of steps by the duration.

Tests 4: Tanden Walk test

Peak to peak difference [g] in acceleration in X-, Y- and Z-direction: difference between the highest negative and positive acceleration value in each of the three dimensions.

Standard deviation [g] of the acceleration in X-, Y- and Z-direction: the standard deviation of the acceleration values over the test time along each axis.

Balance Count (BC) [%] in X-, Y- and Z-direction : the percentage of samples that go over a given threshold. High values indicates difficulty in mantaining the balance during the tandem walk.

Test 5 and 6: Timed Up and Go test and Chair Rise test

Duration [cs]: time needed to complete the test, measured in hundreths of seconds.

Test 7, 8, 9, 10 and 11: Romberg, Semi-tandem, Tandem and One-legged test

Stabilo-ellipse or "acceleration stabilogram": ellipse comprising the 95% of the acceleration values in the transversal plane. For this ellipse two parameters are computed: the **area [BU]** and the **eccentricity**. High values of the ellipse's area indicate posture instability and high values of the eccentricity means that the body tends to sway in a specific direction. For more details about the calculation of the stabilo-ellipse see section 4.3.2. 1 BU (Balance Units) = 0.096 m^2/s^4
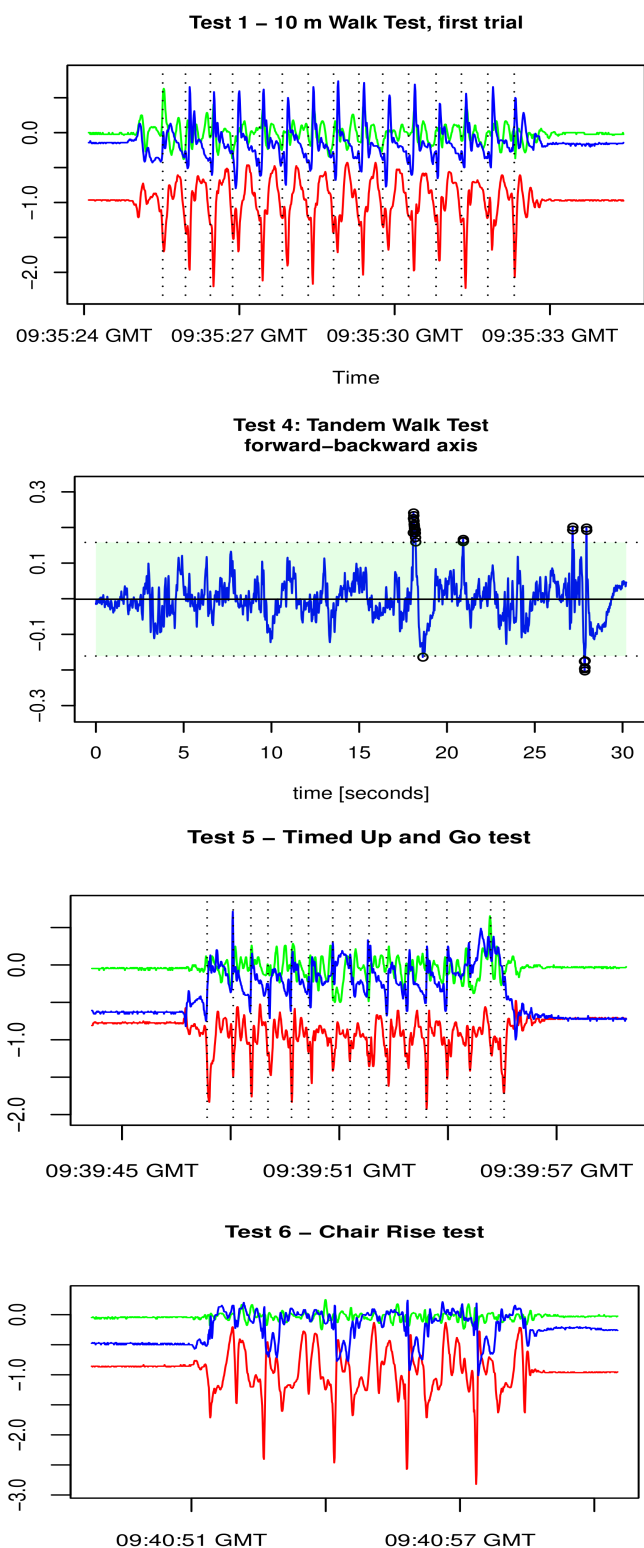
Figure 5.6.: Example of acceleration records of functional tests. Y axis is in *g* units

5.4.4. Android Application for Functional Mobility Tests

We have implemented a prototype of an Android application that allows the user to select among different balance and gait tests (Fig.5.7), record them, and automatically get a visualization

and the values of the extracted parameters. The connection of the smart phone/tablet and the actibelt® is via Bluetooth©. The raw data were recorded in the smart device, and sent automatically through an internet connection to a dedicated server, where they were analyzed. The algorithms that extract the parameters are written in R. The results are sent back to the phone or tablet instantaneously, as long as the internet connection is available.

This prototype is intended to facilitate the functional mobility assessment also outside of the clinic, like for example at home.
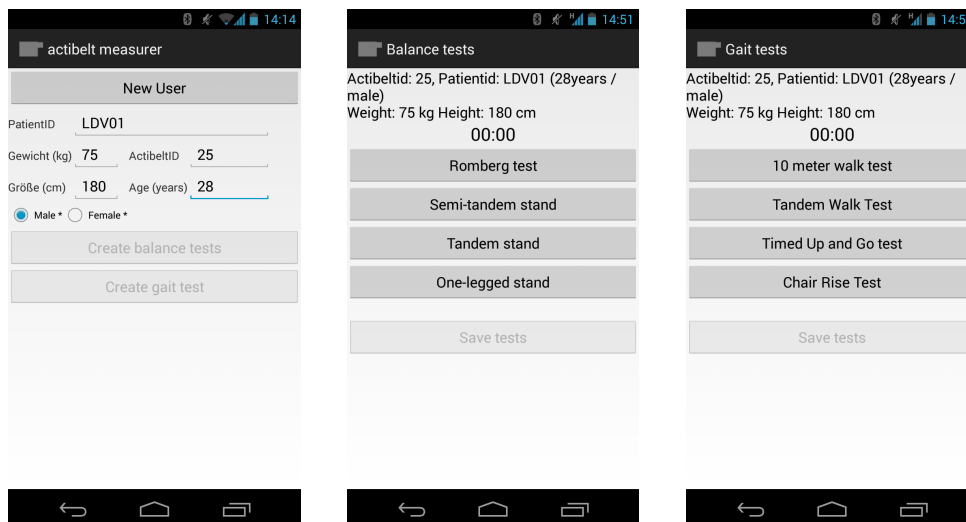

Figure 5.7.: Screenshot of the Android Application Prototype for Functional Mobility Assessment.

5.5. Data Sets

Among all data collected in the study, those variables related to the risk of falling were selected and divided into two data sets: set $\mathcal{A}$ and set $\mathcal{B}$.

Set $\mathcal{A}$ includes only questionnaire information reported by the participants and the classic outcomes of a few functional tests whereas set $\mathcal{B}$ includes also the variables extracted from the functional tests with the use of the acceleration sensor. Figure 5.8 shows the number of participants (including fallers and non fallers) in each data set.

The data were divided into these two groups because we were interested in investigating whether the use of acceleration data in set $\mathcal{B}$ could improve the accuracy of a fall risk assessment model based on questionnare data and a traditional outcomes of functional tests (data set $\mathcal{A}$).

5.5.1. Set $\mathcal{A}$: Clinical Data

This data set contains 17 variables or risk factors for each of the 277 female seniors included in the group. Among those 277 patients, 45 were fallers and 232 non fallers. These variables are specified in Table 5.3 together with the descriptive statistics. Most of them had been identified by the American Geriatrics Society (AGS) [169] as risk factors for falling.

The variables No. 3 to 11 are dichotomous variables, that is, they can take two values, "yes" or "no" ([y/n]). The fall risk factors related to gait and balance deficits are represented by the variables No. 12 - time a subject needs to walk 10 meters -, No. 13 - self-estimation of balance skills -, and variables related to Activities of Daily Living (ADL) - No. 14 to 17. The ADL variables are ternary, they can take the values: "without difficulty", "with some difficulty", or "unable or able only with help" and they are encoded in this table as 0, 5, and 10 respectively. Balance rating can take a value among 10 levels, 10 meaning perfect balance and 0 no balance
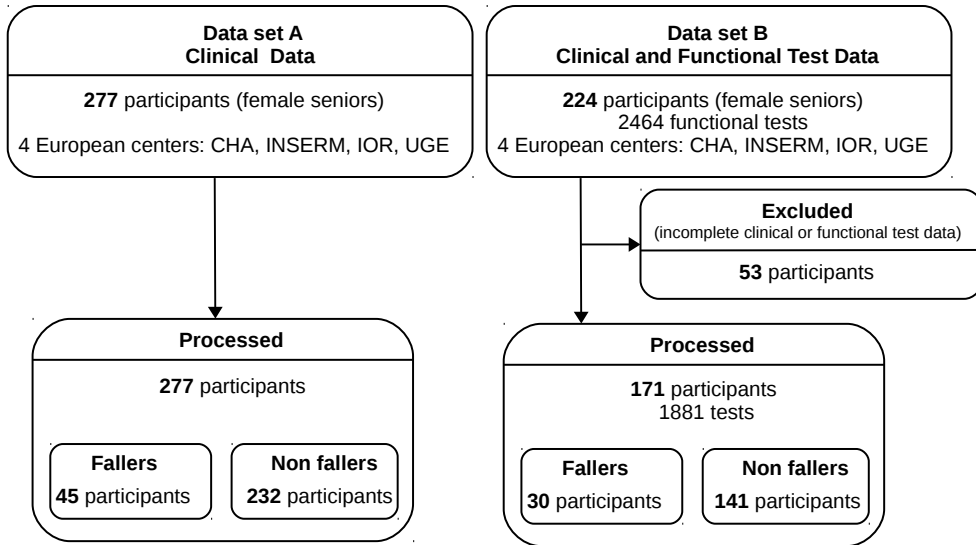
Figure 5.8.: Data sets and number of participants (fallers, non fallers) included in the data base

No.	Variable	Fallers n=45 (16.25%)	Non fallers n=232 (83.75%)	All n=277 (100%)
1	age [years]	69.2 ± 5.5	68.7 ± 5.1	68.8 ± 5.2
2	Body Mass Index [kg/m²]	24.6 ± 3.7	25.4 ± 3.6	25.2 ± 3.6
3	history of falls [y/n]	18 (40.0%)	84 (36.0%)	102 (36.8%)
4	hypertension [y/n]	18 (40.0%)	38 (16.4%)	56 (20.2%)
5	arthritis [y/n]	6 (13.3%)	28 (12.1%)	34 (12.3%)
6	visual deficits (cataracts) [y/n]	13 (28.9%)	34 (14.7%)	47 (17.0%)
7	fear of falling [y/n]	24 (53.3%)	90 (38.8%)	114 (41.2%)
8	foot pain [y/n]	14 (31.1%)	53 (22.8%)	67 (24.2%)
9	sleep disturbances [y/n]	11 (24.4%)	71 (30.6%)	82 (29.6%)
10	depression [y/n]	8 (17.8%)	32 (13.81%)	40 (14.4%)
11	assistive devices [y/n]	2 (4.4%)	10 (4.3%)	12 (4.3%)
12	time for 10 meters walk [sec]	9.8 ± 2.0	9.9 ± 2.4	9.9 ± 2.3
13	balance [0-10]	6.8 ± 2.3	7.1 ± 2.3	7.1 ± 2.3
ADL's self-rating [0,5,10], capacity of:				
14	reaching	1.0 ± 2.3	0.7 ± 1.8	0.8 ± 1.9
15	lifting/carrying	3.1 ± 3.5	2.6 ± 3.4	2.7 ± 3.5
16	washing herself	0.3 ± 1.2	0.2 ± 1.0	0.2 ± 1.0
17	bending	1.1 ± 2.1	0.8 ± 1.9	0.8 ± 2.0

Table 5.3.: Fall risk variables included in the data set $\mathcal{A}$.

at all. For non-dichotomous variables, the mean and standard deviation are given in the specified unit. In the case of dichotomous variables, the number and percentage of "yes" answers is given.

Noise Measurements

The noise present in the data, or equivalently the degree of overlap, was measured with a weighted version, η [14], of the standard measure proposed by Murphey et al.[134], which takes into account the data imbalance. The value calculated on data set $\mathcal{A}$ is $\eta = 78.41\%$, that can described as a high noise level.

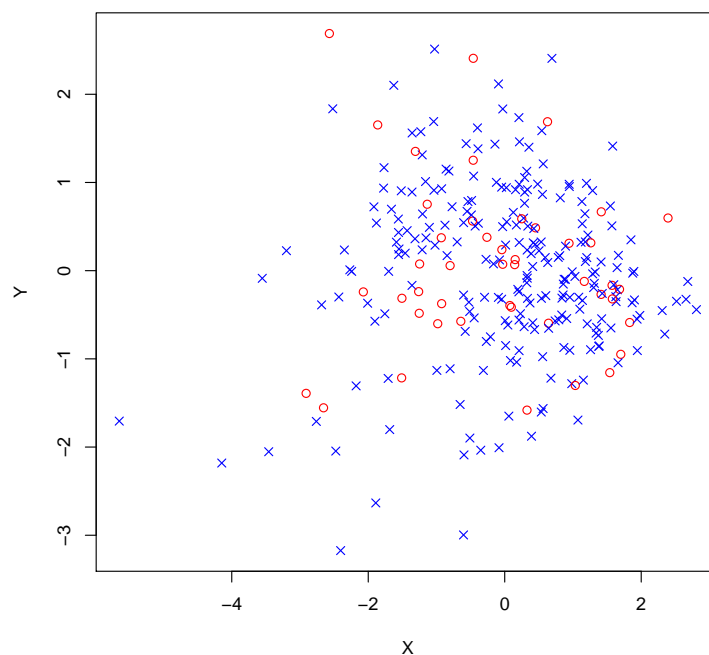To give a graphical idea of the amount of noise present in the data set, we used the Multidi-

Figure 5.9.: MDS map of dataset $\mathcal{A}$

mensional Scaling [103] (MDS). Multidimensional Scaling provides with a visual representation of the noise by projecting the samples onto a low-dimensional space while trying to approximate the distances between the samples as well as possible. The result of the projection onto a two-dimensional space is shown in Figure 5.9, where the blue crosses represent the majority class. As can be seen in the figure, the classes seem to have a large overlap, although this is only an approximation. Actually, the Mardia fit measure [53] is equal to 48.62%, which indicates that the algorithm has omitted more than half of the information during this projection. A value equal to 100% would indicate perfect representation, in which case the samples would lie in a two-dimensional subspace.

5.5.2. Set $\mathcal{B}$: Mixed-Acceleration data

In total, set $\mathcal{B}$ contains 66 variables for each of the 171 patients who have used the accelerometer during the functional assessment. These variables consists of the first 11 parameters included in set $\mathcal{A}$ plus 55 variables extracted from the acceleration data (see section 5.4.3). These acceleration-derived parameters replace the simple functional variables in set $\mathcal{A}$ - duration of 10 meter walk and self-rating of ADLs - and they are expected to have a higher predictive value.

The number of subjects here (171 subjects) are far less than the 277 subjects from data set $\mathcal{A}$. One has to keep in mind that having so few samples (subjects) with so many variables hampers building a proper falls risk model. Due to the large number of variables, Table 5.4 presents only the descriptive statistics for the first 11 clinical variables. More detailed information concerning the acceleration-derived parameters is shown in Table 5.6.

Noise Measurements

The weighted noise measure [14] for data set $\mathcal{B}$ gives again a high noise level with $\eta = 71.33\%$. Thus, both data sets $\mathcal{A}$ and $\mathcal{B}$ are similarly noisy.

No.	Variable	Fallers	Non fallers	All
		n=30 (17.5%)	n=141(82.5%)	n=171 (100%)
1	age [years]	68.7 ± 4.7	67.9 ± 4.6	68.0 ± 4.6
2	Body Mass Index [kg/m^2]	24.8 ± 4.1	25.1 ± 3.6	25.1 ± 3.7
3	history of falls [y/n]	14 (46.7%)	60 (42.6%)	74 (43.3%)
4	hypertension [y/n]	12 (40.0%)	20 (14.2%)	32 (18.7%)
5	arthritis [y/n]	5 (16.7%)	22 (15.6%)	27 (15.8%)
6	visual deficits (cataracts)[y/n]	8 (26.7%)	21 (14.9%)	29 (17.0%)
7	fear of falling [y/n]	14 (46.7%)	49 (34.8%)	63 (36.8%)
8	foot pain [y/n]	11 (36.7%)	33 (23.4%)	44 (25.7%)
9	sleep disturbances [y/n]	7 (23.3%)	51 (36.2%)	58 (33.9%)
10	depression [y/n]	5 (16.7%)	21 (14.9%)	26 (15.2%)
11	assistive devices [y/n]	2 (6.7%)	9 (6.4%)	11 (6.4%)
acceleration data				
12-65	number of steps, asymmetry, test duration, mean walking speed, step length, cadence, peak to peak difference, standard deviation, balance count, stabilo-ellipse			

Table 5.4.: Fall risk variables included in the data set $\mathcal{B}$.

Once more, the Multidimensional Scaling method was used to visually represent the noise, this time on data $\mathcal{B}$. Figure 5.10 confirm the results of a large overlap of the class distributions. The Mardia fit measure is in this case equal to 33.36%.

Criterion Validity and Plausability of Acceleration-Derived Parameters

The criterion validity examines to what extent a measure provides results that are consistent with a gold standard. For a new tool to be accepted, specially in the medical community, the criterion validity must be proved. In our case, the accelerometer should at least provide with the same or comparable results than the current tool used to assess the functional ability of the patients (the stopwatch) and ideally with an added value, like for example, dispense novel predictors for falls risk, reduce the test administration time or automate the data collection. To check the validity of the accelerometer-derived parameters we compared them with the classical outcomes of the functional tests using the Pearson correlation coefficient. To interpret the results we assume that a correlation is strong if the Pearson correlation coefficient r is greater than 0.5, moderate if $0.5 \leq r \leq 0.3$, small for $0.3 < r \leq 0.1$ and insubstantial for $r < 0.1$ [37].

In a strict sense, we can only validate the acceleration parameters by comparing them with a gold standard measure that estimates exactly the same attribute, for example accelerometer-derived gait speed vs. gold standard gait speed or accelerometer-derived time vs. gold standard time. In this case, the only functional attribute we can take as reference or gold standard is the time needed to perform a specific test measured with the stopwatch. However, the comparison was not only done time versus time but the correlation coefficients were calculated between all the gait parameters extracted from the acceleration signal and the gold standard test duration as plausibility check to confirm that the time is inversely proportional to the mean speed as well as to the step length [57, 141]. Moreover, step cadence and number of steps should be inversely and directly correlated to the time due to their obvious dependence upon time and step length. As expected, the results presented in Table 5.5 show a strong correlation between the gold standard time and the time measured with the acceleration sensor, and to a less extent but also in a strong way with the rest of acceleration-derived parameters.

Other plausibility tests were performed to demonstrate comprehensible connections between the acceleration-derived variables and the age of participants since it is well known that functional ability tends to deteriorate with age [74, 91, 117, 141]. Kruskal–Wallis rank sum non-parametric
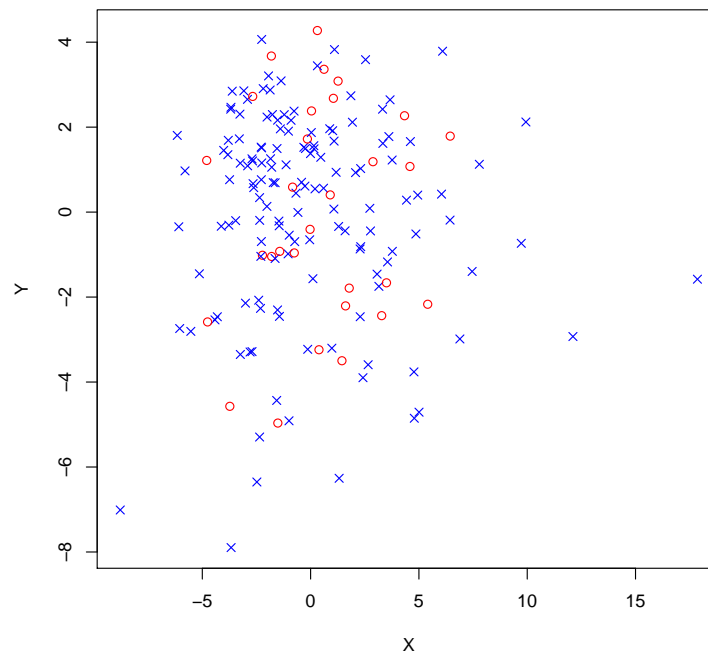
Figure 5.10.: MDS map of dataset $\mathcal{B}$

Gold standard variable	Acceleration-derived variable	Pearson's correlation (r)	p-value
time to walk 10 meters measured with stopwatch	duration of test	0.86	< 0.001
	number of steps	0.75	
	mean speed	-0.79	
	step length	-0.73	
	step cadence	-0.64	

Table 5.5.: Correlation between the outcomes obtained from the gait test with the use of the stopwatch (classical outcome) and those one extracted from the accelerometer.

tests were used to assess the association. Table 5.6 presents the parameters for which the difference is statistically significant (p-value $\leq$ 0.05). The data were divided according to age in three groups; **younger**: $60 \leq$ age < 65 years , **middle** group: $65 \leq$ age ≤ 71 years, and **older**: $71 <$ age ≤ 80 years.

As shown in Table 5.6, age is significantly associated to the number of steps, speed and step length for test 1, 2 and 3. For test 4 (Tandem Walk Test) the Balance Count (BC) across the vertical and lateral axis was higher for the older group compared to the younger and middle group, as well as the test duration in the Timed Up and Go test. All balance test are significantly associated with age, with the area of the stabilo-ellipse bigger as the age of the patient increases. The deviation between subjects in balance tests 8-11 is large, indicating that individuals within the groups may vary considerably across a "stability spectrum" when the difficulty of the test grows. It was also found that the asymmetry index tends to increase in the oldest group compared to the other ones.

Variable	Test	Mean value ± SD			p-value
		Younger n=40	Middle n=90	Older n=41	
number of steps	1	16.0 ± 1.9	16.5 ± 1.7	18.0 ± 3.0	0.0011
	2	15.8 ± 1.6	16.4 ± 1.7	17.8 ± 2.2	0.0001
	3	17.5 ± 2.4	17.5 ± 2.2	19.3 ± 3.9	0.0064
speed [m/s]	1	1.22 ± 0.20	1.19 ± 0.17	1.04 ± 0.23	0.0005
	2	1.25 ± 0.17	1.22 ± 0.18	1.09 ± 0.22	0.0015
	3	1.03 ± 0.18	1.02 ± 0.24	0.88 ± 0.24	0.0016
step length [m]	1	0.64 ± 0.07	0.61 ± 0.06	0.58 ± 0.07	0.0009
	2	0.64 ± 0.06	0.62 ± 0.06	0.57 ± 0.07	0.0001
	3	0.58 ± 0.08	0.58 ± 0.07	0.54± 0.06	0.0010
BC vertical [%]	4	1.35 ± 0.40	1.32 ± 0.45	1.58 ± 0.56	0.0178
BC lateral [%]	4	1.00 ± 0.38	0.99 ± 0.45	1.28 ± 0.55	0.0118
Duration [s]	5	10.05 ± 1.56	10.75 ± 2.14	12.47 ± 4.25	0.0127
Ellipse's area [10×BU]	7	0.04 ± 0.03	0.04 ± 0.03	0.06 ± 0.04	0.0000
	8	0.06 ± 0.04	0.07 ± 0.06	0.16 ± 0.24	0.0000
	9	0.17 ± 0.25	0.22 ± 0.26	0.44 ± 0.47	0.0001
	10	0.34 ± 0.48	0.55 ± 1.30	0.71 ± 0.87	0.0044
	11	0.33 ± 0.35	0.37 ± 0.47	0.87 ± 0.96	0.0003

Table 5.6.: Correlation between accelerometer-derived parameters and patients' age by groups

5.6. Fall Prediction Model

In this chapter the aim is to estimate the risk of presenting a fall based on two different data sets: $\mathcal{A}$ (only clinical data) and $\mathcal{B}$ (clinical and sensor data). Thereby, we want to confirm that the accelerometer data can improve the fall risk assessment, at least in the existing cohort of subjects.

As already mentioned, the available data for our prediction model was retrieved from the EU-Project VPHOP. The process of data collection is depicted in figure 5.11. The collected data contain information about retrospective falls, clinical and questionnaire data (like age, BMI, diseases, etc), actibelt® functional data, and prospective falls. The prospective falls constitute the class label for the prediction model, that is, we try to predict if a subject is a faller or non-faller in the year after the main assessment.

The methodology used to build the model for fall prediction is presented in Figure 5.12. Each of the steps depicted in the diagram are explained in detail in the next sections.

5.6.1. Preprocessing

This section introduces the methods which must be applied prior to use dimension reduction or feature selection techniques on a data set.

Completion of Missing Values

On both data sets, a few entries are missing due to errors in the data collection process. In set $\mathcal{A}$ there is 1 missing value (0.27 ‰ of all entries) and in set $\mathcal{B}$, 53 missing entries (4.72 ‰). One alternative to deal with this problem would be the exclusion of the missing entries, but this would mean either to exclude subjects or variables, and since our data set is already small the decision was to apply low-rank matrix completion instead. The method approximates the given, non-complete matrix with a low-rank matrix assuming normally distributed variables. Afterwards,
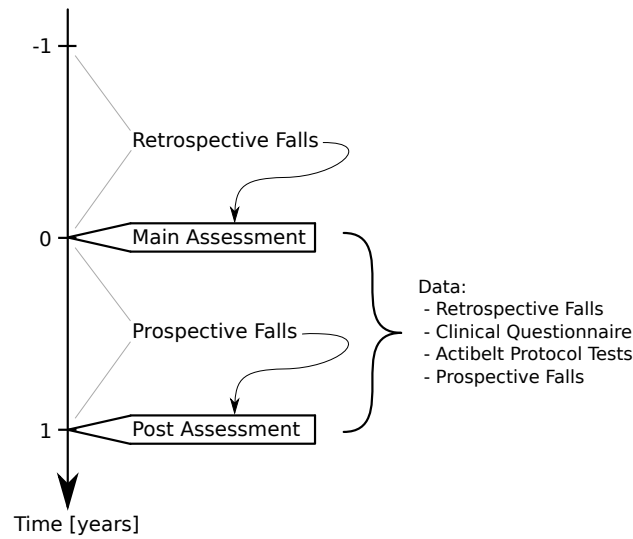
Figure 5.11.: Timeline of data collection [14].

it replaces the empty cells of the original matrix with the corresponding cells from the low-rank matrix.

Standardization

Principal Component Analysis (PCA) and distance-based classifiers (e.g., k-nearest-neighbors or support vector machines), both used for the construction of this prediction model, need the input variables to be rescaled when they have different units so that they do not erroneously place too much importance on a variable with high variance when the magnitude of the variance is just caused by the unit of the measurement. Furthermore, in case of PCA, the mean of each feature must be equal to zero, because non-zero mean variables would distort the singular value decomposition.

There are many different methods of scaling a data set, the most widely used are normalization and standardization. Standardization sets the mean of every variable to zero and the respective standard deviation to one. To conduct this, every feature $\mathbf{x}_{\mathrm{orig,i}}$ is transformed into

$$\mathbf{x}_{\mathrm{std,i}} \leftarrow \frac{\mathbf{x}_{\mathrm{orig,i}} - \mu_{\mathrm{i}}}{\sigma_{\mathrm{i}}}, \tag{5.1}$$

with μ_{i} the mean of $\mathbf{x}_{\mathrm{orig,i}}$ and σ_{i} the standard deviation. The drawback of this method is that it assumes a normal distribution of the variable, which is certainly not the case for the dichotomous features. For normalization, the following formula,

$$\mathbf{x}_{\mathrm{norm,i}} \leftarrow \frac{\mathbf{x}_{\mathrm{orig,i}} - \min(\mathbf{x}_{\mathrm{orig,i}})}{\max(\mathbf{x}_{\mathrm{orig,i}}) - \min(\mathbf{x}_{\mathrm{orig,i}})}, \tag{5.2}$$

is used. Normalization projects the variables to [0,1]. In addition to that, we need to mean-center the data by subtracting the mean from every variable. The problem here is that, after this scaling, the standard deviations are not comparable and this misrepresents the variables when applying PCA. In our case, there exists a difference among the variables with the highest and lowest standard deviation of a factor over 10 after applying normalization and mean subtraction to data set $\mathcal{B}$. This difference cannot be tolerated, therefore, standardization was used, accepting
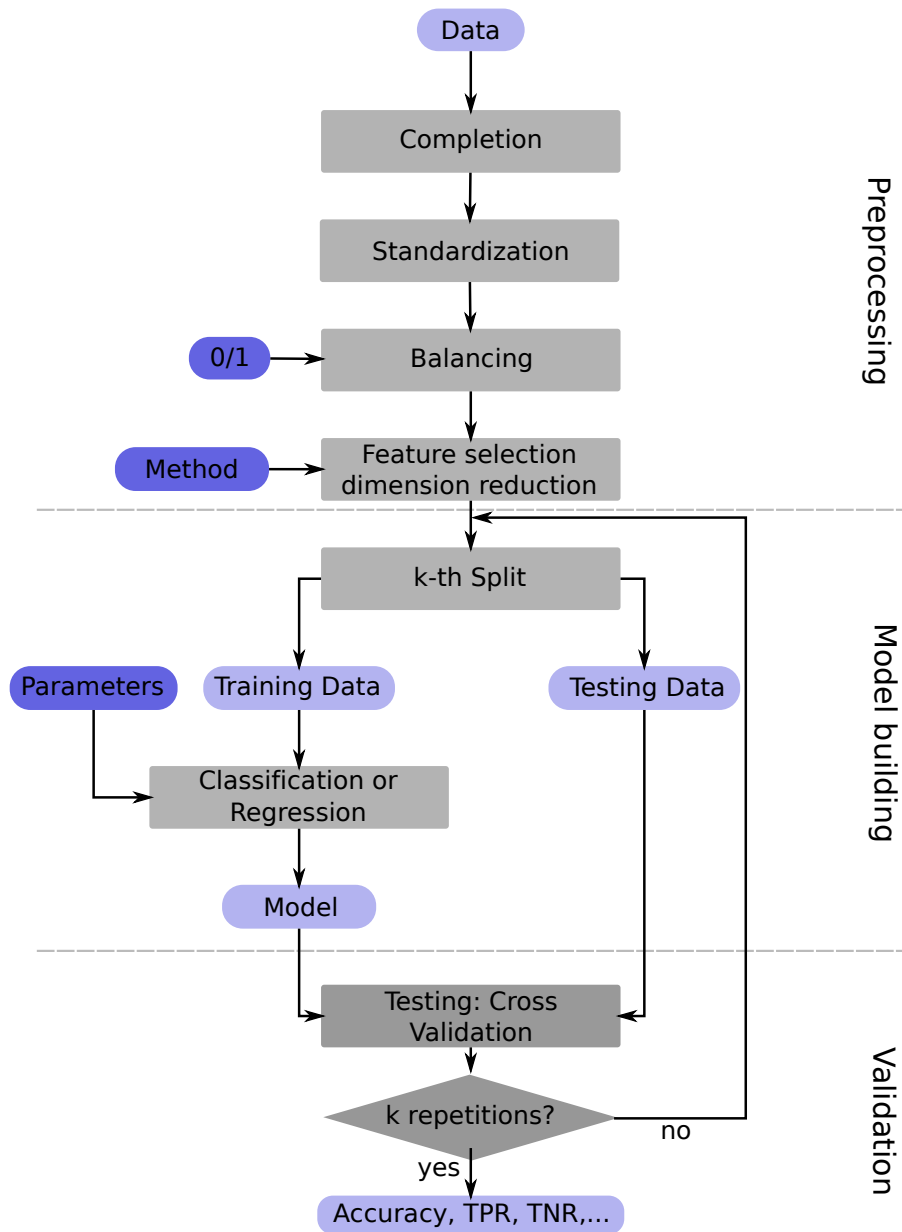
Figure 5.12.: Methodology of the fall risk assessment model [14].

that we violate the normality assumption for binary variables. We assume that this has no considerable impact in the accuracy of the prediction.

Class Imbalance

Most machine learning algorithms work best when the number of instances of each classes are roughly equal. When the number of instances of one class exceeds the other, the classes are unbalanced. In a two-class problem, the class with more samples is called majority class and the smaller class is the minority class. The ratio of the class sizes in our case is 1:5.15 for set $\mathcal{A}$, and 1:4.70 for set $\mathcal{B}$, with the group of non-fallers as the majority class. Building a model on such an unbalanced set works fine as long as the classes are clearly separated by the distributions of their samples, which is not the case for our data as it could be proved in the noise analysis section. If the classes overlap, building good classification models becomes very difficult, since models tend to classify everything as being part of the majority class.

There are several methods to deal with class imbalance [101]. Those mainly investigated are the sampling methods and algorithmic level schemes. As sampling method, SMOTE [34] was used in connection with Logistic Regression. SMOTE performs under- and oversampling simultaneously. The parameters were set to achieve a 2:1 class size ratio in favor of the majority class. This ratio has been chosen because we desire to mitigate the class imbalance, but still having a reliable representation of the original classes. Algorithmic level schemes, which have proven to be much more useful with kNN and SVM, were also used.

5.6.2. Dimentionality Reduction and Feature Selection

When the number of features (variables) increases, or equivalently the dimentionality of the data, the volume of the space increases so fast that the available data become sparse. This sparsity is problematic for any method that requires statistical significance like for the classification and regression algorithms which implicitly estimate the underlying distribution of the samples. This phenomenon is related to the called "curse of dimentionality".

One property derived of this phenomenon, is that with high-dimensional data all samples tend to be equidistant in infinite dimensions. That leads to less meaningful or even meaningless models when using distance-based classification, like kNN or SVN. The question is, from how many dimensions upwards the effect of this phenomenon significantly influences the classification process. We investigated this question with the highest dimensional data set, $\mathcal{B}$, and found that the highest distance is roughly twice as big as the minimum distance [14]. Therefore, although the distances appear to be still useful, they have lost some of their meaning.

In order to get a lower dimensional data set the number of features was reduced using Principal Components Analysis (PCA), sparse Partial Least Squares (sPLS) and Elastic Nets (EN). The following sections contain a short introduction to these methods (for more details, see [14]). As target for our reduction was set to have approximately 5, 10 or 30 times the number of samples as variables, as suggested by Murphey et al. [134].

Principal Component Analysis

Principal components analysis (PCA) looks at the columns of a data set $\mathbf{X}$ (in our case, the features) and tries to find an orthogonal projection $\mathbf{W}$ to a subspace, in which most of the information is kept, so that

$$\min_{\mathbf{W}} ||\mathbf{X} - \mathbf{X}\mathbf{W}||_2^2. \tag{5.3}$$

A way to solve this equation is to apply singular value decomposition (SVD). PCA may be problematic because it considers as information simply the variance in the data set $\mathbf{X}$ without

taking into account the relations to one or more dependent variables $\mathbf{Y}$. Therefore, PCA may extract irrelevant information for the prediction from matrix $\mathbf{X}$ only based on their amount of variance.

Sparse Partial Least Squares

Classical Partial Least Squares (PLS) tries to maximize correlation between the matrix of features $\mathbf{X}$ and the class labels $\mathbf{Y}$ by finding the multidimensional direction in the $\mathbf{S_X}$ space that explains the maximum multidimensional variance direction in the $\mathbf{S_Y}$ space. Thus, every direction vector tries to maximize the covariance as

$$\hat{\mathbf{w}}_k = \arg \max_{\mathbf{w}_k} \mathbf{w}_k^\mathsf{T} \mathbf{X}^\mathsf{T} \mathbf{Y} \mathbf{Y}^\mathsf{T} \mathbf{X} \mathbf{w}_k$$
$$\text{such that } \mathbf{w}_k^\mathsf{T} \mathbf{w}_k = 1 \text{ and } \mathbf{w}_k^\mathsf{T} \mathbf{S_{XX}} \mathbf{w}_j = 0 \text{ for } j = 1, ..., k-1. \tag{5.4}$$

Sparse Partial Least Squares (sPLS), in addition, makes the direction vector to be sparse [36]. In order to get as many components of the direction vector as possible equal to zero, the maximization term for every direction vector $\mathbf{w}$,

$$^\mathsf{T} \mathbf{Y} \mathbf{Y}^\mathsf{T} \mathbf{X} \mathbf{w} \text{ such that, } \mathbf{w}^\mathsf{T} \mathbf{w} = 1, \ ||\mathbf{w}||_1 < t, \tag{5.5}$$

is constrained by an ℓ_1-penalization. By setting the parameter t we control the maximum correlation and consequently the number of $\mathbf{w}$ zero entries. The non-zero entries are the features given aftwerwards to the classifier.

Elastic Nets

Elastic Nets [208] is a regularized regression method that linearly combines the LASSO [177] and the Ridge Regression [79] penalties to obtain sparsity and still work with data sets that contain more variables than samples. Elastic Nets tries to find a regression model $y = \mathbf{X}\boldsymbol{\beta} + \eta$ based on generalized linear regression models and the combined penalty. The solution are the regression coefficientes $\boldsymbol{\beta}$ that minimize the expresion

$$\min_{\boldsymbol{\beta}} ||\mathbf{y} - \boldsymbol{\beta}\mathbf{X}|| \text{ such that } (1-\alpha)||\boldsymbol{\beta}||_1 + \alpha||\boldsymbol{\beta}||_2^2 \leq t. \tag{5.6}$$

The penalty is a weighted sum governed by the parameter $\alpha \in [0, 1[$. By adjusting α we decide the number of non-zero entries or features. EN are particularly useful for situations with few samples, even with less samples than features.

5.6.3. Classification and Regression Models

The models employed for the prediction are classification and regression models. Regression models are traditionally used in the clinical field whereas classification methods have been recently presented as an alternative approach showing good results in falls risk assessment [82].

The following algorithms were chosen to investigate the predictive performance when applied to the data sets $\mathcal{A}$ and $\mathcal{B}$: Logistic Regression (LR), k-nearest-neighbors (kNN), and Support Vector Machines (SVM). Logistic Regression is broadly used in the falls risk prediction field [69, 100, 121, 190], the kNN algorithm is among the simplest of all machine learning algorithms but nontheless considered to be very effective [87], and SVM is viewed as a very powerful classifier. The detailed mathematical description of all those algorithms can be found in [14]. Next, I summarize some properties and important considerations to take into account when using these models.
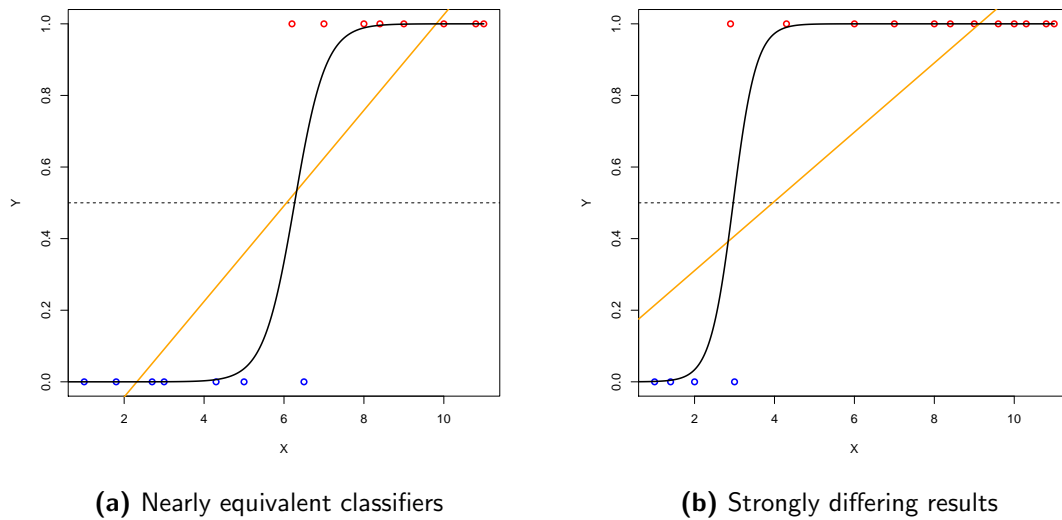
(a) Nearly equivalent classifiers   **(b)** Strongly differing results

Figure 5.13.: Regression curves of linear (orange) and logistic (black) regression for two different data sets [14].

Logistic Regression

Logistic Regression measures the relationship between a categorical dependent variable x and one or more independent variables y, which are usually (but not necessarily) continuous. In the following formula, P gives the logistic function for multidimensional predictive variables,

$$P = \frac{e^{a+\mathbf{bx}}}{1 + e^{a+\mathbf{bx}}} = \frac{1}{1 + e^{-(a+\mathbf{bx})}}. \tag{5.7}$$

Where $\mathbf{b}$ contains the regression coefficients. For example, in the scalar case, the function looks like depicted in figure 5.13. The intercept a allows shifting the entire graph to the left or right, b defines the slope of the linear part. The class label y can be retrieved by rounding P, $y = f(\mathbf{x}) = \text{round}(P)$.

Logistic Regression deals with some pitfalls present in linear regression. First, linear regression assume homoscedasticity. This means that the variance of y, and therefore the variance of the error, are approximately constant over all values of the predicting variable x, which is not the case for a binary y. And second, linear regression assumes linear dependencies between the variables. This is also not the case for a mixture of binary y and continuous variables x and it can lead to severe errors in prediction.

When the classes are unbalanced and do not have similar variances, like in the example of Figure 5.13b, the function of linear regression (in orange) is erroneously shifted along the coordinates axis. In contrast, the Logistic Regression function (in black) still draws a correct decision function. In this example, the classes are defined by the dependent variable $y \in \{0, 1\}$, which stands for blue and red class respectively. The variable $x \in \mathbb{R}$ predicts the class and the horizontal, dashed line at $y = 0.5$ indicates the decision boundary.

k-Nearest-Neighbors

The k-Nearest-Neighbors (kNN) algorithm [153] is a simple non-parametric method used for classification (and also for regression). The general approach of classification is to find relations between features contained in a sample vector $\mathbf{x}$ and its corresponding class y. The data set to

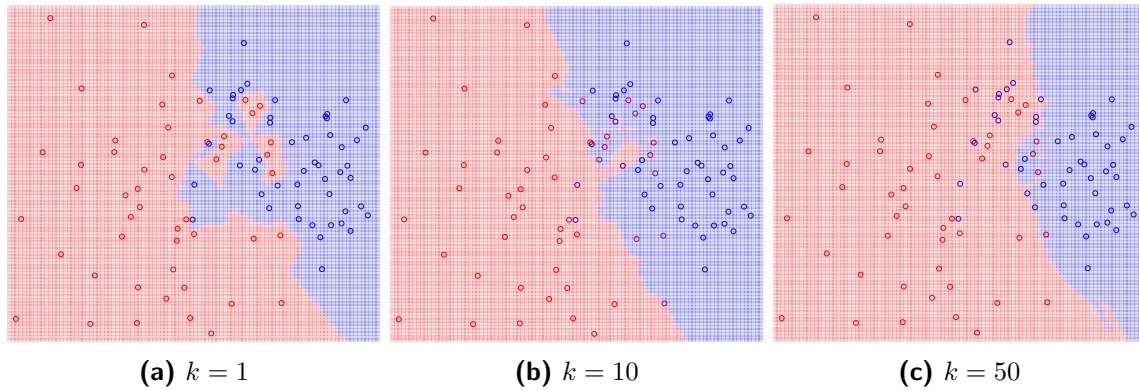**(a)** $k = 1$ **(b)** $k = 10$ **(c)** $k = 50$

Figure 5.14.: Examples for kNN classification areas on one data set [14]

analyze is typically divided into two sets: the training set, used for building the classifier model, and the test set, for validating the model on new, uncorrelated data.

In k-NN classification the input consists of the k nearest neighbors of $\mathbf{x}$, formally described as $N_k(\mathbf{x}) = \{\mathbf{x_i}\}_{i=1,\ldots,k}$. And the output is approximated by the average of the classes of the k nearest neighbors of $\mathbf{x}$ rounded to the closest integer value (which represents a class) such as in

$$\hat{y}(\mathbf{x}) = \text{int}\left(\frac{1}{k}\sum_{\mathbf{x_i} \in N_k(\mathbf{x})} y_i\right). \tag{5.8}$$

This formula is only valid for two-class data sets, for more classes a voting scheme has to be applied, because averaging a sum would then probably give faulty results if there is no linear dependency between the classes. To enable kNN to deal with imbalanced classes, class weighted votes were used.

The k-NN can use different measures to calculate the distance between a new sample and the the k nearest neighbors. Three conventional distances were tested: the Euclidean, the Manhattan, and the Maximum Distance, and found that prediction using the Maximum Distance was worse than the other two whereas the classification with the Euclidean and Manhattan was comparable. The Euclidean distance was chosen for the final analysis because is more widespread.

The parameter k defines the number of nearest neighbors used for choosing a class for a new sample. A small k will produce a complex decision border, whereas a larger k produces a smoother border. Figure 5.14 shows the graphic result after applying kNN over two normal distributed classes marked with blue and red color for different k values. In case that a simple distribution can be assumed for each class (for example, a normal distribution), it is recommended to choose a large k such that the decision boundary will reflect the true simple boundary as well as possible. Using a small k is incorrect for simple distributions because it may cause overfitting. For complex distributions, a small k should be used to avoid overgeneralization. If no assumptions about the distributions can be made, k should be found via cross-validation [14].

Support Vector Machines

The last classification method investigated was Support Vector Machines [173] (SVM). SVM's are a potentially powerful classifier, but the effectiveness depends on the selection of the parameters.

SVM models are based on hyperplanes in a space with many more dimensions than the original feature space. These hyperplanes act as decision borders. They are defined by a transformation function $\Phi(\cdot)$ and its orthonormal vector $\mathbf{w}$ combined with an offset b from the lower dimensional

(a) Class separation via different hyper-planes

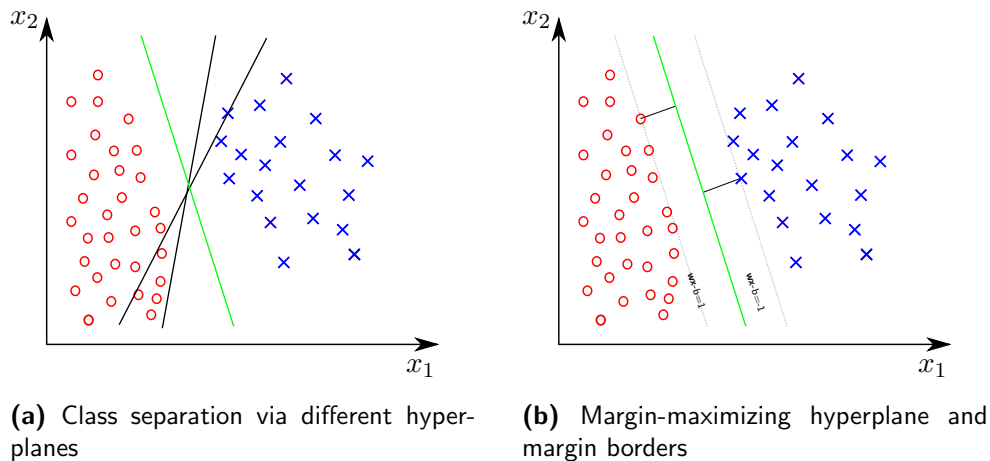**(b)** Margin-maximizing hyperplane and margin borders

Figure 5.15.: Example of hyperplanes in SVM [14].

space. The class $y \in \{-1, 1\}$ of a sample $\mathbf{x}$ is given by

$$y = \text{sgn}(\langle \Phi(\mathbf{w}), \Phi(\mathbf{x}) \rangle + b). \tag{5.9}$$

To illustrate how SVM works, let us assume a simple classification problem with a data set in the 2D-space linearly separable so that no transform to high-dimensional space is needed. In this case, the two classes can be separated by different hyperplanes. The criterion for the optimal hyperplane (in green) is that it has to maximize the distance, also called margin, between itself and the closest point of each class, as seen in Figure 5.15. This is equivalent to minimize the error probability when classifying a new sample. The points closest to the hyperplane are defining $\mathbf{w}$ and b, thus the hyperplane itself. These so-called support vectors, are indicated in dotted lines in figure 5.15b.

A transformation to many dimensions makes the classification problem to likely become linearly separable in the new space. However, this transformation is computationally expensive, this is why in the practice the function $\Phi(\cdot)$ and the scalar product in the high-dimensional space are replaced by a kernel. To what transformation and which inner product the kernel is equivalent, is something difficult to comprehend. Hence, the so called "black box" behaviour of the SVM that makes almost impossible to choose the model parameters intuitively. To overcome this problem a method widely adopted for the search of the optimal parameters is the grid search. More information on types of kernels and their parameters can be found in [14].

5.6.4. Cross-validation

Cross-validation (CV) is a model validation technique for assessing how the results of a statistical analysis will generalize to an independent data set. The different prediction methods were validated using a type of CV called k-fold cross validation that is well-suited for those cases where there are not a big number of samples, like in our data set. In every experiment, a new set of n/k samples from the entire data set was defined as validation set, where n was the number of all samples (Fig. 5.16). The remaining data are the training set. In this way we use every sample as a training and a validation sample, and the size of the training set, $n - (n/k)$, is always relatively high. In particular, $k = 10$ was chosen, being guided by the recommendation of some authors [98].
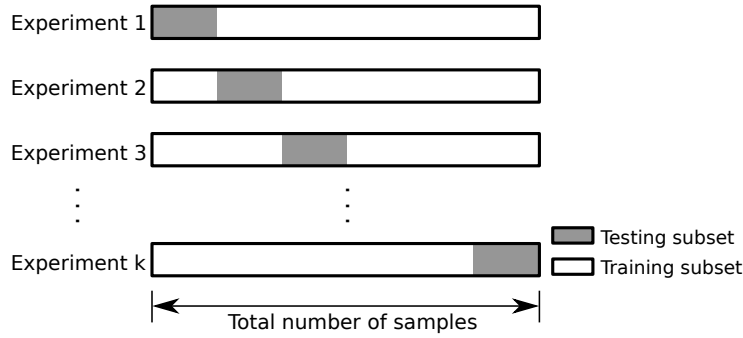
Figure 5.16.: k-fold cross validation

Validation metrics

To quantitatively compare the performance of the different classification methods and our results with those from other authors, it is necessary the use of some metrics. Typically in clinical studies, participants suffering a certain disease or responding to a medication are called positives. In our case, the *fallers* are called *positives* and, in consequence, the *non-fallers* are called *negatives*. Table 5.7 shows the naming of correct and incorrect classifications.

Accuracy   A common metric is the accuracy (Acc), which is calculated by dividing all correct classifications by the number of samples classified. However, in situations where the minority class is more important and with skewed class imbalance, it is better to avoid the accuracy metric.

$$\text{Acc} = \frac{\sum \text{True Positive} + \sum \text{True Negative}}{\sum \text{All Samples}} \in [0,1]. \qquad (5.10)$$

Sensitivity   The sensitivity, also called true positive rate (TPR), is the ratio of individuals predicted as positives divided by the total number of actual positives. Sensitivity is the measure used to report how reliable the prediction model is in identifying individuals with the positive condition, in our case, in identifying fallers.

$$\text{TPR} = \frac{\sum \text{True Positive}}{\sum \text{Actual Positive}} = \frac{\sum \text{True Positive}}{\sum \text{True Positive} + \sum \text{False Negative}} \in [0,1]. \qquad (5.11)$$

Specificity   The specificity or true negative rate (TNR) is the ratio of individuals predicted as negatives divided by the total number of actual negatives. Analogous to the previous metric, the specificity is the measure used to report how reliable the prediction model is in identifying individuals with the negative condition or equivalently, in our study, how good the model is at identifying non-fallers.

$$\text{TNR} = \frac{\sum \text{True Negative}}{\sum \text{Actual Negative}} = \frac{\sum \text{True Negative}}{\sum \text{True Negative} + \sum \text{False Positive}} \in [0,1]. \qquad (5.12)$$

In real-world classification problems, typically a trade-off between TPR and TNR has to be made. Usually when one classifier becomes very sensitive, it looses specificity and vice versa. For us, it is more important to classify accurately the future fallers than the non-fallers because the consequences of wrongly identifying a non-faller as a faller are not so critical as the contrary.

	Actual Positive (Faller)	Actual Negative (Non-faller)
Predicted Positive	True Positive	False Positive
Predicted Negative	False Negative	True Negative

Table 5.7.: Confusion Matrix

Youden Index   Youden's J [202], also called the Youden Index, is the difference between the true positive rate and the false positive rate. Its value ranges from -1 to 1. A value J equal to zero means that the classification gives the same proportion of positive results for groups with and without the condition, i.e the classification is useless. Values of $J < 0$ may indicate a systematic error in the prediction that makes the classification results to be assigned to the opposite class. Finally a J value equal to one indicates that there are no false positives or false negatives, i.e. the classification is perfect. Therefore, a good classifier will present a Youden Index near to one.

$$J = \mathrm{TPR} + \mathrm{TNR} - 1 \quad \in [-1, 1]. \tag{5.13}$$

F_1 *- score*   The F_1 - score is the harmonic mean of precision (Prec) and sensitivity. With the precision as the number of correct positive predictions divided by the number of all the individuals with the positive condition (fallers). An advantage of F_1 over Youden's J is that it takes the relation of the class sizes into account by using the precision.

$$F_1 = 2 \frac{\mathrm{TPR} \cdot \mathrm{Prec}}{\mathrm{TPR} + \mathrm{Prec}} \in [0, 1]. \tag{5.14}$$

Regression Performance

In regression approaches to classification it is common to draw receiver operator characteristic curves (ROC) in order to choose the threshold that optimizes the prediction performance. This is necessary because regression methods do not automatically compute the class of a sample, but a certain value, which is presumably close to the true class value. Therefore, in a two-class problem, all values below this threshold will be assigned to one class and all values above that threshold to the opposite class.

Figure 5.17 is an example of a ROC curve, where the coordinates axis represents sensitivity (TPR) and the horizontal axis the false positive rate (FPR). The FPR is calculated from the specificity (TNR) as $\mathrm{FPR} = 1 - \mathrm{TNR}$. The blue and red lines are examples originated from two different prediction models for a specific range of thresholds. For the lowest threshold, all the samples are classified as belonging to the positive class. This would mean $\mathrm{TPR} = \mathrm{FPR} = 1$ (top right corner). Increasing the threshold will lower the false positive rate and the true positive rate. For a good classifier the TPR would decrease slightly. Thus, in our example the blue curve represents a better classifier than the red one. When the threshold equals the highest value, all samples will be classified as negative, yielding to $\mathrm{TPR} = \mathrm{FPR} = 0$ (lower left corner). If the ROC curve lies near the dotted line, the prediction is useless because the regression values would be distributed similarly. A good threshold is one lying near the top left corner.

An alternative to the visual interpretation is computing the area under the curve (AUC). An AUC value close to 1 corresponds to a good prediction. An AUC close 0.5 means that the curve is close to the dotted line in figure 5.17 and therefore useless.
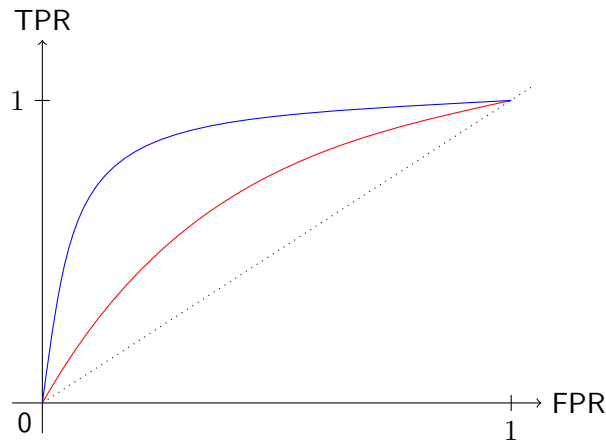
Figure 5.17.: Two receiver operator characteristic curves

5.6.5. Results

This section presents the results of classifiers applied to both data sets, statistical analysis of chosen features, and two alternative approaches via multiple fallers and balanced data sets. The principal metric employed to interpret the performance of the different classification methods is the Youden Index. Eventually, the rest of validation metrics (TPR, TNR, Acc and F_1-score) were included to allow comparability with other studies. For simplicity, to examine the properties of each single classifier only the results obtained with the data set $\mathcal{B}$ were depicted because, as it will be shown in later sections, the prediction using the acceleration-enhanced data set $\mathcal{B}$ is more useful than the classification obtained with data set $\mathcal{A}$.

k-Nearest-Neighbor

The results of applying k-Nearest-Neighbor (kNN) after Elastic Nets (EN) and Principal Component Analysis (PCA) for different numbers of features are shown in Figure 5.18. The Youden Index was also calculated for kNN with sPLS but it was not depicted because the results were very similar to EN, which is always the case for the three investigated classifiers. For every number of chosen features a boxplot was depicted, since the results of feature selection methods are variable (but still quite consistent).

The features obtained after applying PCA are a linear combination of the original variables, which makes it harder to interpret. Other consideration is that EN can sometimes not choose a certain number of features. For example, in Figure 5.18a the results corresponding to a number of features from 12 to 15 is similar because the next higher number of features that EN is able to select is 16.

The results show that Elastic Nets are in comparison to PCA a better feature selection method, in particular if the goal is to get few features. For many features, the curse of dimensionality, begins to influence the results negatively. The effect of this curse makes the predictive value of kNN fade if we use more than 30 variables. The best result achieved by kNN with EN feature selection is for five selected features with $J = 0.43$, representing TPR=0.6 and TNR=0.83. Taking into account that the minority class of fallers is defined as positive, this is a relatively sensitive result. The high TPR is a consequence of the class-specific vote weights, which are set to the reciprocal of the number of samples within the respective class.

The changing trajectory of the Youden Index curve for PCA is due to the fact that PCA does not consider the usefulness of the features, only their variance. Useful features come in by chance
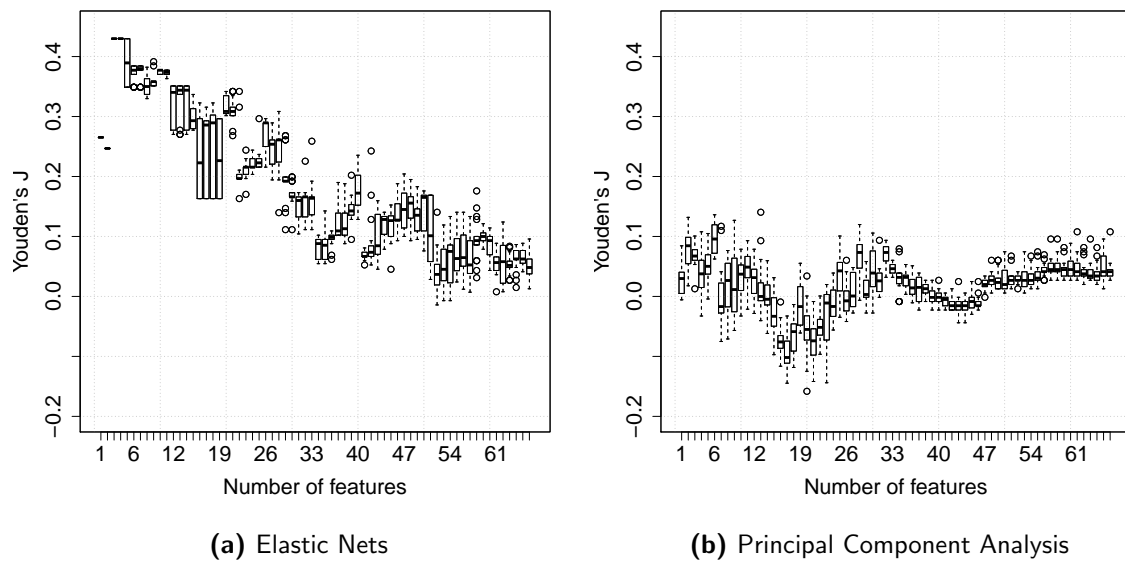
(a) Elastic Nets   **(b)** Principal Component Analysis

Figure 5.18.: Performance of 20-NN Classification on data set $\mathcal{B}$ with 10-fold cross validation [14]

when more dimensions are chosen. The best median performance of kNN after PCA is achieved with J=0.09 corresponding to TPR=0.50 and TNR=0.59.

The parameter k was selected assuming that the classes follow simple distributions, since almost all dependencies between predictive features and the dependent variable "faller/nonfaller" are linear or generalized linear. Therefore, we expect better prediction performance for higher k-values. In effect, the values k=1,2,...,50 were tested and k=20 was found to be the optimal value.

Logistic Regression

The advantage of Logistic Regression (LR) is the absence of tunable parameters except the decision threshold needed for classification. However, the threshold can be simply set to 0.5, given that the regressed values express probabilities.

As can be seen in Figure 5.19, also in combination with Logistic Regression, EN is superior again over PCA. The highest median J=0.45 is in this case obtained for 12 features and corresponds to TPR=0.5 and TNR=0.95. This means that the probability of correctly classifying a faller is only 0.5, but because of the high specificity, the algorithm is much more reliable when classifying a person as non-faller. Another remark is that, Logistic Regression can deal better with an increasing number of features when compared to kNN, since the performance beyond 30 features is notably better.

The results of regression are not directly classes but regression values expressing the probability to belong to one or another class. For the optimum outcome (EN, 12 features), the histogram of these values for fallers and non-fallers are depicted together with the receiver-operator characteristic curve (ROC) in Figure 5.20. From the histograms one can see that non-fallers clearly get lower regression values on average, confirming the plausibility of the method. The ROC showed the usefulness of the prediction. The area under this curve is AUC = 0.80. By inspecting the ROC it would now be possible to adjust the decision boundary for either higher specificity or sensitivity.
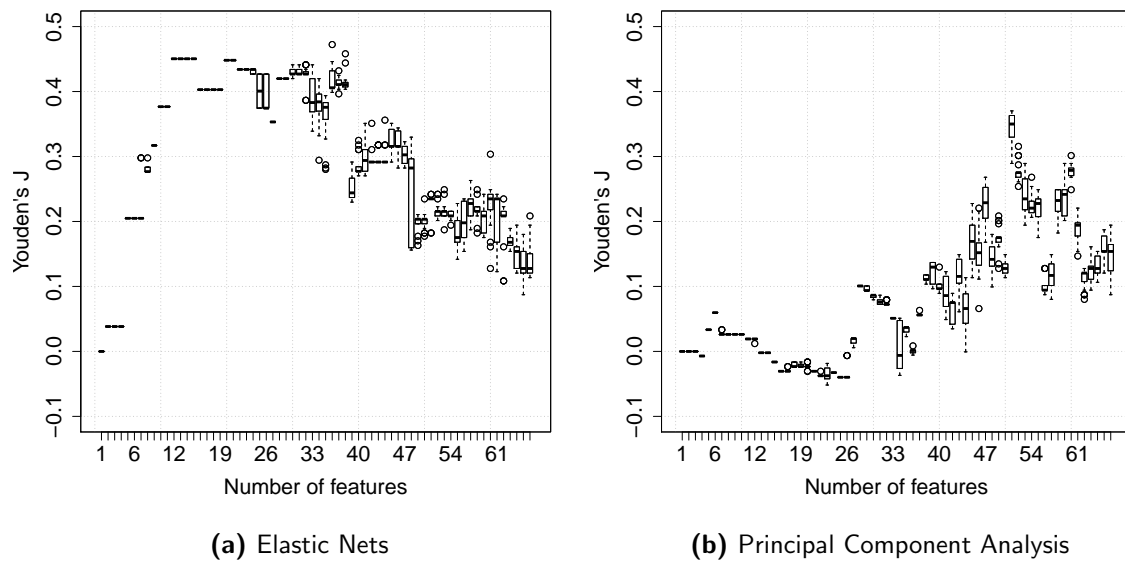
(a) Elastic Nets **(b)** Principal Component Analysis

Figure 5.19.: Performance of classification via Logistic Regression on data set $\mathcal{B}$ with 10-fold cross validation [14].

Support Vector Machines

Again for Support Vector Machines the feature selection method EN shows a better performance than PCA. Figure 5.21 also shows that SVM with EN can handle many features without affecting too negatively the prediction. From 50 onwards, most of the information is passed to the feature selectors and the performances using EN and PCA are approximately equal. Youden's J does not decrease as quickly as in the previous methods for many features. SVM achieves the highest J=0.56 of all our experiments on 36 variables chosen by EN using a linear kernel with cost $C = 10$, obtained through an extensive grid search. The classes were custom weighted with the reciprocal of the number of samples within the respective class [16]. Using the reciprocal of the class size has shown to be a fair choice, with the resulting classifier favoring none of the classes too much. This is reflected in the relation of TPR and TNR. Both are for most tests close to each other, for example Youden's J=0.51 for SVM on 30 features chosen by EN results from TPR=0.73 and TNR=0.78.

Comparison of Classifiers and Feature Selection Methods

The most representative results of the F_1 scores together with the conventional metrics for both data sets are shown in Table 5.8 and Table 5.9. More details can be found in [14]. The first column indicates the type of feature selector, the number of features used (in between brackets), and the classifier. The target numbers of features or dimensions are 5, 15 and 30. In case that EN could not return an specific number of features, the algorithm will look for the closest available number below that value. For example, for a target of 15 features, the EN algorithm provides with 12. In the first three rows of each table, the classifier performance is presented without any feature selection, i.e., all features were given to the classifier. When using Logistic Regression no method to handle imbalance was used, that is why in Table 5.9 the true positive rates (TPR) in few dimensions are low. On the contrary, SVM and kNN can deal with imbalanced classes in the form of per-class defined cost or weight, which explains the much better true positive rates.

Although the accuracy is a widely used metric to present prediction results, specially in the
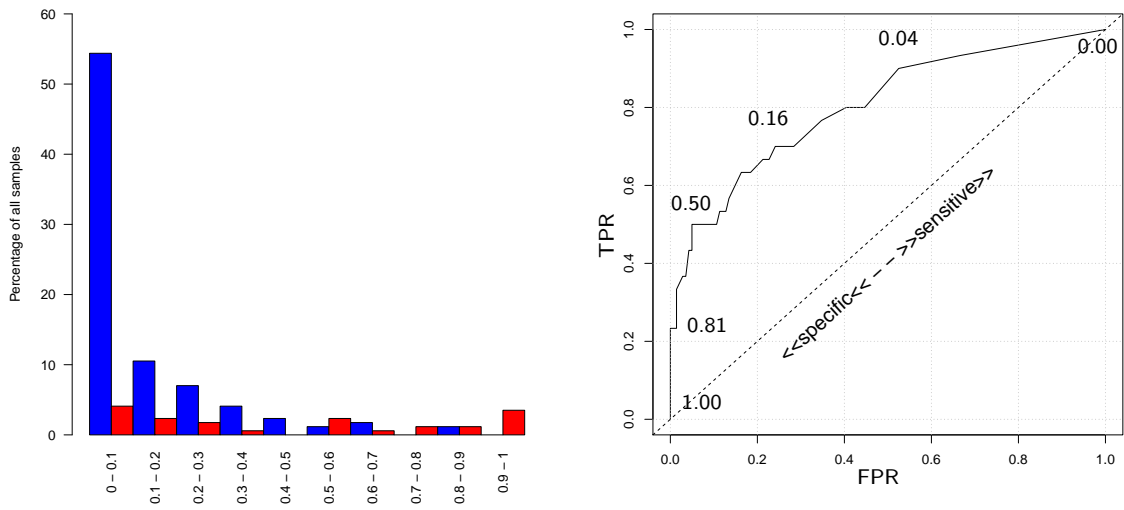
(a) Histogram of the regression results. True non-fallers are blue, true fallers red

(b) Receiver-operator characteristic with threshold values at corresponding locations

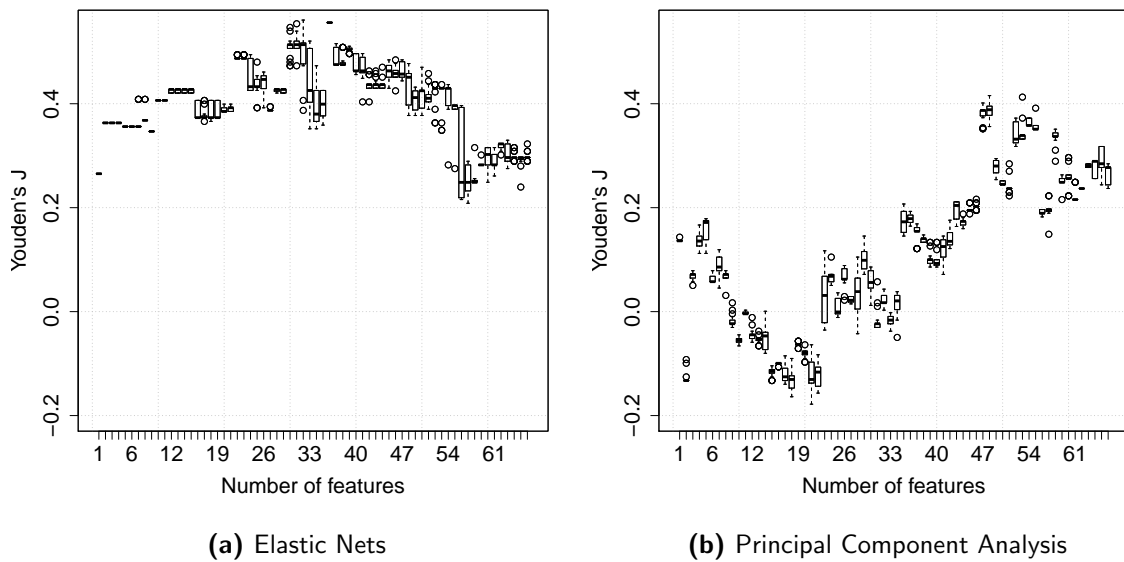**Figure 5.20.:** Results of Logistic Regression on 12 features chosen by Elastic Nets [14].

(a) Elastic Nets

(b) Principal Component Analysis

Figure 5.21.: Performance of Support Vector Machines on data set $\mathcal{B}$ with 10-fold cross validation [14].

89

clinical field, we can see in both tables that the results with the higher accuracy are usually the ones showing the worst performance. For example, in the first row in Table 5.8 we see an accuracy of 0.81, which can lead us to think that the classifier makes a good prediction. However, when we look to the true positive rate, the prediction becomes useless, since the TPR is so small that in most of the cases the classifier predicts a person as non-faller. A low TPR does not affect the accuracy too strongly since the majority of samples are non-fallers. In contrast, the F_1 score reflects this condition, making it a more reliable metric.

	TPR	TNR	Accuracy	F_1 score
NF*(17), LR	0.04	0.96	0.81	0.07
NF*(17), SVM	0.37	0.65	0.61	0.24
NF*(17), 20-NN	0.24	0.73	0.65	0.18
EN(5), SVM	0.42	0.80	0.74	0.35
EN(15), SVM	0.38	0.64	0.60	0.23
PCA(5), SVM	0.40	0.65	0.61	0.25
sPLS(13), LR	0.04	0.97	0.82	0.07
sPLS(13), SVM	0.36	0.66	0.61	0.23
sPLS(13), 20-NN	0.31	0.75	0.68	0.24

*NF: No Feature Selection

Table 5.8.: Metrics for different feature selection methods, classifiers and number of features on set $\mathcal{A}$

Comparing the F_1-scores of equivalent methods in both tables, one can see that the fall prediction obtained with data set $\mathcal{B}$ surpass the prediction of data set $\mathcal{A}$. For example, SVM using data set $\mathcal{A}$ with 15 dimensions selected by EN achieves $F_1 = 0.23$, whereas with the equivalent dimensions using set $\mathcal{B}$ achieves $F_1 = 0.5$, which is a considerable difference. Overall, SVM in combination with EN yields the best prediction.

In addition to the variables automatically chosen by the different feature selection methods, the last 3 rows of Table 5.9 contain the so-called "13-Set". This set comprises 13 variables selected manually by an expert (marked in bold in the Table C.2 of Appendix C) . Thereby, the purpose was to investigate the plausibility of the automatically selected features by comparing their predictive value with the one derived from well-known fall risk factors in the clinical field. As seen in the table, the thirteen "hand-picked" variables with SVM produced a comparable prediction in terms of F_1-scores than the 12 features chosen by EN again in combination with SVM. In this example, the TPR of the "13-Set" was a little worse than for EN, but the TNR was slightly better.

Feature Analysis

The five first features chosen by the EN, the optimal feature selector in our case, and their correlation with the dependent variable (faller/non-faller) are described in Table 5.10. Correlations are only indicated for those statistically significant variables. All p-values are again Bonferroni-corrected.

It is remarkable the few number of statistically significant variables, in particular for data set $\mathcal{A}$, and their low correlation with the class label. This could be a consequence of the multifactorial nature of the risk of falling or also could be caused by the small size of our cohort. Other clinical studies have a couple of hundred or even thousands of test subjects.

A list with the rest of variables selected by EN for a target number equal to 15 (eventually 12) and 30 can be found in the Appendix C.

	TPR	TNR	Accuracy	F_1 score
NF*(66), LR	0.37	0.79	0.72	0.29
NF*(66) SVM	0.50	0.82	0.77	0.40
NF*(66), 20-NN	0.47	0.57	0.55	0.27
EN(5), LR	0.23	0.97	0.84	0.34
EN(5), SVM	0.53	0.82	0.77	0.45
EN(5), 20-NN	0.53	0.82	0.77	0.44
EN(12), LR	0.50	0.92	0.85	0.54
EN(12), SVM	0.67	0.79	0.77	0.50
EN(12), 20-NN	0.43	0.84	0.77	0.40
EN(30), LR	0.53	0.91	0.84	0.54
EN(30), SVM	0.73	0.78	0.77	0.53
EN(30), 20-NN	0.37	0.74	0.67	0.28
PCA(5), SVM	0.50	0.65	0.62	0.32
PCA(30), SVM	0.43	0.65	0.61	0.28
sPLS(30), LR	0.33	0.88	0.78	0.35
sPLS(30), SVM	0.53	0.74	0.70	0.39
sPLS(30), 20-NN	0.37	0.77	0.70	0.30
13-Set, LR	0.30	0.92	0.81	0.36
13-Set, SVM	0.60	0.81	0.77	0.48
13-Set, 20-NN	0.53	0.74	0.70	0.39

*NF: No Feature Selection

Table 5.9.: Metrics for different feature selection methods, classifiers and number of features on set $\mathcal{B}$

Feature No.	Set $\mathcal{A}$	Correlation Coef.	p-value	Set $\mathcal{B}$	Correlation Coef.	p-value
1	**Hypertension**	0.23	<0.05	**Hypertension**	0.26	<0.05
2	Fear of Falling	-	-	**Test9: Area**	0.34	<0.05
3	BMI	-	-	**Test6: Duration**	0.26	<0.05
4	Foot Pain	-	-	Test2: SD Asymmetry in X-Direction	-	-
5	Reaching	-	-	Test4: SD Acceleration in Y-Direction	-	-

Table 5.10.: Order of the first five features, as selected by Elastic Nets on the sets $\mathcal{A}$ and $\mathcal{B}$. Significant features are in bold letters.

Balanced Classes

As already mentioned, distance-based classifiers are able to handle the class imbalance by applying different weights. In contrast, regression methods, like LR, work better when the data which are given to them is balanced. In order to accomplish the balancing of the class sizes in set $\mathcal{B}$, we used two methods. The first method consist in applying SMOTE, and the second method in randomly subsample the majority class.

SMOTE decreases the imbalance via under- and oversampling of the majority and minority class, respectively. The parameters were chosen such that the class ratio on the training set after SMOTE were 2:1 in favor of the original majority class, the non-fallers. With the balanced classes, the LR classifier trained on five features selected by EN, achieved an improvement in the prediction. The Youden Index J increased from 0.20 to 0.31 and the F_1-score from 0.34 to 0.42. The TPR improved from 0.23 to 0.47 at cost of TNR (from 0.97 to 0.84).

The second approach is to randomly sub-sample the majority class such that both classes are of equal size. With this method, the best classification was achieved for SVM with 19 features selected by EN, obtaining TPR=0.77 and TNR=0.73 (Youden's J=0.5). Although SVM was

already tuned to handle the class imbalance, a data balancing prior to classification shows still an improvement.

5.7. Summary of Findings

In this chapter, we demonstrated the feasibility of introducing the use of accelerometers in a big multicenter clinical study by creation of a data collection protocol, a back-end "tap detection" algorithm, and other specific algorithms to automatically extract relevant parameters for functional analysis from the raw data. The plausibility and construct validity of the acceleration-derived parameters was confirmed using statistical analysis. These parameters constitute some of the features used to predict the risk of falling.

A model for fall prediction was set that deals with the main limitations encountered in the exploratory analysis of the real data collected in the study. These limitations include the class imbalance and the high dimensionality. To deal with class imbalance data-level methods like SMOTE can improve the results when used with classifiers that do not include algorithmic methods, however, algorithm-level methods have shown to be superior. To handle the high dimensionality problem, the best feature selection method resulted to be Elastic Nets (EN). Principal Components Analysis (PCA) is not feasible for dimension reduction when not knowing if features are useful for classification.

Youden's J and the F_1-score are more valuable validation metrics than the accuracy when working with unbalanced classes, despite the later metric is widely used in the clinical field.

Support Vector Machines (SVM) are on average the best method for classifying our data. They present less specificity, than for example Logistic Regression, but they are more sensitive. In our case, sensitivity is preferred over specificity because the consequences of classifying a non-faller as faller are not harmful. On the contrary, misclassifying potential fallers may exclude them from taking part in beneficial intervention programs for fall risk prevention.

The best fall prediction result achieved with our model (TPR=0.73, TNR $= 0.78$) is comparable to results from other authors, with the difference being that the presented model takes into account critical aspects, like validation methodology or class imbalance, not considered in previous work (see Table 1.2). Nevertheless, this comparison should be made carefully considering the different characteristics of the subjects databases employed in the different studies. It is also worthy to mention that the achieved prediction is considerably accurate taking into account the high overlap of the class distributions (fallers/non-fallers), the strong imbalance, and the fact that the study group consisted in very young seniors (the average age was 68 years).

From the clinical point of view, the most important finding of this study is that the use of accelerometers has shown to be useful to improve falls risk prediction over conventional methods based in patient clinical history and basic functional measures.

6. Fall Detection and Alarm

6.1. Introduction

Falls are the major source of morbidity and mortality among the older people. According to several studies, about one in three adults aged 65 or more fall each year and one in four of those who fall end up with serious injuries [180, 124]. One major problem is that many of the elderly who fall at home become helpless and require assistance to get up. Tinetti and colleagues [179] conducted a study with 313 noninjured fallers living in the community, aged 72 years and older, of which 47% reported inability to get up after at least one fall. Fleming et al. [60] found that the 80% (53/66) of the elderly fallers in their 90's were unable to get up after a fall, and 30% had lain on the floor for an hour or more. Of those who were alone when they fell, 80% did not activate their alarms. In this context, user-friendly automatic or semiautomatic emergency response and monitoring systems can extend the length of time that seniors are able to live in their own homes.

The focus of this chapter lies in a critical part of such alarm system: the sensitivity and false alarm rate of the fall detection algorithm. Together with the usability aspect of the activity monitor these are known to be the key success factors for fall detection systems. To satisfy the condition that the algorithm runs in real-time for an extended period of time in the microcontroller of the monitoring device and to guarantee a high user acceptance the actibelt$^{®}$, is used as key part of the study.

6.2. Experimental Setup

6.2.1. Acceleration Sensor

The acceleration data was collected with the actibelt$^{®}$ sensor described in section 2.1.

6.2.2. Subject Database

Historical Database

In order to validate the upper limit of the realistic false alarm rate of the algorithm I selected 3 subsamples from the Sylvia Lawry Center database according to age criteria: adolescents, middle-age and elderly people. Despite the fact that the fall detection system is thought to be used mainly among the elderly the study was designed to include subjects of a wide age range to confirm the assumption about the potential decrease of false alarm rate with age.

In total, 2,415.9 hours (100.7 days) of continuous weekly acceleration measurements corresponding to ADLs (Activities of Daily Living) in free living conditions were used for the validation. The subsamples are distributed as follows: the first cohort is comprised of 3 diabetic adolescents (10.5yr), the second cohort is formed by 6 healthy individuals (6 male, 32.3 yrs) and the third by 5 elderly women (70.2 yrs) diagnosed with osteoporosis but without mobility limitations.

Prospective Data Collection

For the development and improvement of the fall detection algorithm, a series of fall-like activities and simulated fall events (Table 6.1) were performed by two healthy individuals (2 male; age:

46 and 79 yrs; height: 1.84 and 1.74 m) in a controlled laboratory setting at the Sylvia Lawry Center for Multiple Sclerosis Research (Munich, Germany). Both subjects were informed about the aims of the activity measurement and had given written informed consent.

Real-World Falls

Real-world falls of two persons (54.5 $\pm$2.1 yrs) who accidentally fell while being monitored with the accelerometer were included in the data base. One of the falls resulted in a fracture of the neck of femur. These unfortunate accidents provided with an extremely valuable data for two main reasons: 1) the difficulty that in itself entails to get a real fall. As an example, to capture 100 real-world falls it would be necessary to record approximately 100,000 days of physical activity (300 person years) [15]. And 2), the detailed documentation about the real-world falls and the physical activity in the previous weeks to the fall; in addition to the self-reported information, there exist 24-hour accelerometer recordings and, for one of the falls, clinical functional test measurements (Timed Up and Go test, 10 meter test and 6 minute walking test).

6.2.3. Trial Protocol

The trial protocol consists of a series of 16 real falls in laboratory conditions and 7 fall-like activities which should not trigger an alarm. The experiments were performed on a sofa and on a crash pad 8 cm thick (see Table 6.1 and Figure 6.1). Falls were performed in three directions: backwards, forwards and sideways, starting from a walking and a standing position and were selected to mimic typical real world situations. E.g., sequence 17 and 18 reflect situations where the subject stumbles before falling down. Sequence 12, 13 and 14 imitate the everyday scenario of a person lying down on the bed and, occasionally, tossing and turning. The last task (number 23) simulates someone who sits down and loses consciousness resulting in a fall.

All participants performed the sequence once and were instructed about how to execute the specific task at the beginning of each activity. All activities start with a pause of five seconds, followed by a tap on the belt and other five-second pause. All activities were video recorded.

6.3. Fall Detection Algorithm

6.3.1. Feature Extraction

The preprocessing of the raw data was performed according the filtering methods described by Karantonis et al. in [95]. First, the acceleration raw data are denoised using a median filter (n = 3). Then, in order to get the gravity acceleration component (GA) the output is low pass filtered, whereby a custom third-order elliptical infinite impulse response (IIR) filter with cut-off frequency at 0.25 Hz. The acceleration forces originated by the body movements (BA) are extracted by subtracting the gravity acceleration to the denoised signal. After that, both components, GA and BA, are transformed into spherical coordinates.

By comparing the magnitude of the acceleration vector produced by the body movement, ρ_{BA}, with a certain threshold I determine a possible impact against the ground. This magnitude can be calculated as following:

$$\rho_{BA}[i] = \sqrt{x^2[i] + y^2[i] + z^2[i]} \tag{6.1}$$

where $x[i]$ is the ith sample along the x axis (likewise for $y[i]$ and $z[i]$).

The θ coordinate of the gravitational component, θ_{GA}, is used to differentiate between standing/sitting position and lying orientation. The value θ_{GA} (equation 6.2) estimates the angle formed by the vertical axis of the upper body and the ground.

$$\theta_{GA[i]} = cos^{-1}(z[i]/\rho_{BA}[i]) \tag{6.2}$$

Table 6.1.: Test protocol

Surface	Starting position	N°	Instructions *	Desired outcome
	3 steps away from the crash pad	1	walk, fall forwards softly, stay in lying position for 10 sec	alarm
		2	walk, fall forwards softly, writhe softly for 10 sec	alarm
		3	walk, fall forwards hardly, stay in lying position for 10 sec	alarm
		4	walk, fall forwards hardly, writhe softly for 10 sec	alarm
	1 foot away from the crash pad	5	fall forwards softly, stay in lying position for 10 sec	alarm
		6	fall forwards softly, turn on your back, stay in lying position for 10 sec	alarm
		7	fall forwards hardly, stay in lying position for 10 sec	alarm
		8	fall forwards hardly, turn on your back, stay in lying position for 10 sec	alarm
Crash pad		9	fall backwards softly, stay in lying position for 10 sec	alarm
		10	fall backwards hardly, turn on your back, stay in lying position for 10 sec	alarm
		11	fall backwards softly, get up after 5 sec	no alarm
	2-3 feet away from the crash pad	12	lie down backwards, stay in lying position for 10 sec	no alarm
		13	lie down backwards, writhe heavily for 10 sec	no alarm
		14	lie down backwards, writhe heavily for 10 sec and tap on the belt	no alarm
		15	fall softly sideways, stay in lying position for 10 sec	alarm
		16	fall hardly sideways, stay in lying position for 10 sec	alarm
	1 foot and 3 steps away from the crash pad	17	walk, stumble after 3 steps and fall forwards softly, stay in lying position for 10 sec	alarm
		18	walk, stumble after 3 steps and fall forwards hardly, stay in lying position for 10 sec	alarm
Sofa	half a foot away from the sofa, facing it backwards	19	fall on the sofa, leave your upper body in an upright position, stay for 10 sec	no alarm
		20	fall on the sofa, leave your upper body in an upright position, writhe softly, stay for 10 sec	no alarm
		21	sit on the sofa, change into a lying position, stay for 10 sec	no alarm
		22	fall on the sofa, change into a lying position, stay for 10 sec	no alarm
Sofa and crash pad		23	sit on the sofa, wait 5 seconds, tap on the belt buckle, wait 10 sec, fall forwards on the crash pad, stay in lying position for 10 sec	alarm

*Explanation: *fall softly*: break the fall with your arms and knees, *writhe softly*: 10°-15° change of position, *stumble*: after 3 steps, hook your one leg behind the other and fall. *writhe heavily*: approx. 90° change of position,

Figure 6.1.: Sequence of photos taken during the performance of the experiment simulating falls in differen positions.

where $z[i]$ is the ith sample along the vertical axis.

To distinguish between rest and activity phases, the sum of the windowed standard deviation, SWSD (equation 6.3), of the ρ_{BA} signal was calculated for a window length $L = 50$ samples.

$$SWSD = \sum_{i=1}^{n} \sigma_i \tag{6.3}$$

where

$$\sigma_i = \sqrt{\frac{\sum_{j=(i-1)L}^{iL-1} \left(\rho_{BA}[j] - \bar{\rho}_{BA_i}\right)^2}{L}} \tag{6.4}$$

with

$$\bar{\rho}_{BA_i} = \frac{\sum_{j=(i-1)L}^{iL-1} \rho_{BA}[j]}{L} \tag{6.5}$$

6.3.2. Threshold Calculation

To find the optimal thresholds for the identification of the fall impacts and the relative body position, the controlled fall data set (30 simulated falls; Table 6.1) was divided randomly in two equal groups, one selected as the training group. The thresholds were calculated using only data from this data set.

The threshold for ρ_{BA} was chosen such that we achieve 100% sensitivity when the algorithm is run over the training data and it was selected as the minimum rounded down value of the

acceleration peaks produced by the fall impacts in the signal ρ_{BA}. This number, th_{mag}, is equal to 1.9 g. The tilt angle threshold, th_{ang}, was selected as the minimum average value of θ_{GA} during the next 13 seconds after the fall, and its value is 49.8 degrees. The three seconds immediately after the fall were ignored due to potential residual movement relating to the knock into the crash pad. To identify periods of inactivity the SWSD of the magnitude vector signal should be under the threshold, $th_{act} = 3.0$ g.

6.3.3. The Algorithm

The fall detection algorithm is based on the detection of an impact by the comparison of the magnitude of the acceleration vector produced by the body movement to a preset threshold and on the body orientation after the impact.

The ρ_{BA} samples which surpass th_{mag} are grouped together in a block provided that the sample difference in between them is not more than 15 samples and the length of the block is maximum 1 second duration (100 samples). Each of these blocks is classified as a potential fall. In a second step, the relative position of the trunk in respect of the perpendicular vector to the ground's plane is calculated along the 10 seconds interval within the 3rd and 13th second after the first sample of the block. It has been empirically found that an interval of 3 seconds after the fall is enough for the tilt signal to stabilize in case a fall occurs. If the trunk tilt given by θ_{GA} within this period goes over the threshold th_{ang}, the person is considered to be in a non-standing position. An alarm is then only activated if the person was in a standing position before the impact and the quantity of movement after it does not go over the threshold th_{act}.

Initially the algorithm was developed and validated in R [174] and later implemented in C language so that it could run in the microcontroller of the actibelt$^\circledR$.

6.3.4. Validation Method

In order to evaluate the performance of the fall detector, the algorithm was run, first over the whole test battery of activities and simulated fall events (Table 6.1); second, over the continuous acceleration data set of ADLs recorded in a free-living environment and then over the real-world falls data.

The output of the first analysis yields an estimation for the sensitivity of the algorithm on the basis that the fall scenarios covered in the experiment protocol are highly representative of a real fall, whereas the free-living activity recordings give an upper limit for the false alarm rate (it is unknown whether some real alarms were within the ones detected as such) in an entirely realistic scenario. Finally, the two real-world falls were useful to test the ecological validity of the test protocol. The fall-like activities included in the protocol were intended to be used for the refinement of the algorithm or in event of troubleshooting in case that the false alarm rate were much higher than the expected one but never to serve as an estimation for the specificity. The ecological validity of the specificity calculated as such is void due to the absence of a prevalence value.

6.4. The Alarm System

A prototype of an Android application was developed such that when a fall is detected by the sensor, an alarm is sent via Bluetooth$^\copyright$ to the smart phone where the application is running and automatically an emergency phone call is activated. The application can be programmed to call different phone numbers, for example, those ones from relatives or caregivers.

6.5. Results

In total, 2,416 hours of uncontrolled ADL activities, 46 simulated falls and fall-like events as well as two real-world falls were analyzed in R retrospectively. The overall sensitivity measures the percentage of positives (alarms) in the prospective data collection which are correctly identified as such. In this study the sensitivity obtained with the validation data reached the 100%.

The algorithm detected two fall-like impacts (see Fig.6.2) when ran over the files which contained the real-world falls at around the same time when it was reported that they happened. 1,835 000 samples per axis × 3 axis (approx. 5 hours recording) were analyzed and the results did not show any false alarm.
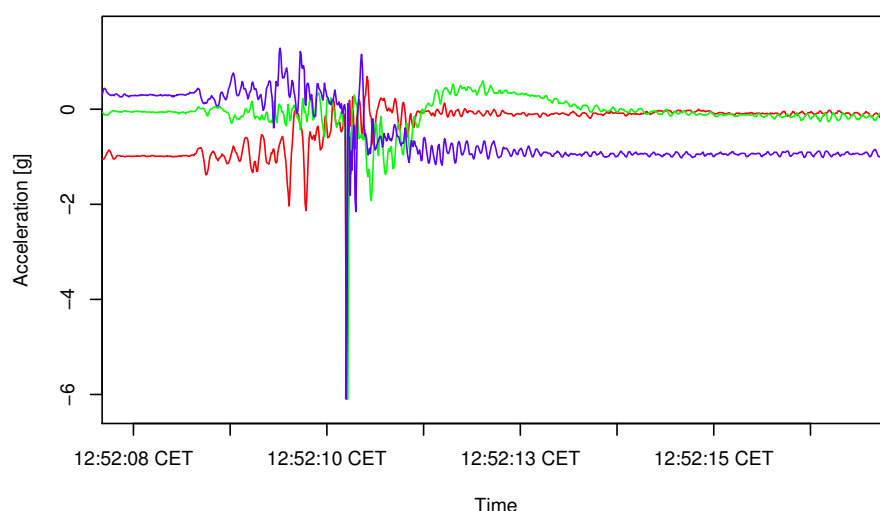
Figure 6.2.: Real fall detected by the algorithm. Red line is the vertical axis, green is the mediolateral axis and blue is the anterioposterior axis (g = 9.81 m/s/s)

The upper limit for the false alarm rate was calculated separately for each cohort in a two-step evaluation. The results are shown in Table 6.2. The UL-cure refers to the results obtained after eliminating the false positives due to the act of removing the device from the belt. The assumption about false alarm rate decreases with age was confirmed.

	Mean age (yrs)	Recorded hours	UL-raw	UL-cure
Cohort 1	10.5	688.1	135	127.7
Cohort 2	32.3	732.1	26.6	11.8
Cohort 3	70.2	995.7	17.4	2.9

Table 6.2.: UL-raw/UL-cure: upper limit false alarm rate without/with data curation (falls per months)

6.6. Summary of Findings

I have developed a promising methodology to detect falls in the elderly using a body worn sensor that has high sensitivity and an acceptable rate of false alarms – 0.56 false alarms per day with the potential of being reduced to 0.09. In comparison with the similar studies cited in the literature

review presented in Section 1.4.3, the algorithm presents the lowest false alarm rate and the highest sensitivity under laboratory conditions. In relation to systems using ambient devices – like video camera systems or floor pressure sensors – the sensitivity, despite the different nature of the data base, is comparable in some cases to the results presented here [148], but the lack of information about the false alarm rate did not allow to fairly check the differences in the number of false positives per day. Although video cameras can provide sufficient information for falls detection, lighting conditions, field-of-view constraints, installation of cameras, coverage range and privacy issues are still major limitations [148].

A particular strength of the current study is to have provided a strict upper limit for the realistic false alarm rate using historical accelerometry data. One should note that the specificity of an algorithm using fall data generated in laboratory settings is meaningless in a real world scenario due to the absence of a prevalence value, although it is still widely used as a measure for algorithm performance.

The success detecting real falls evidence that the simulation protocol chosen for the algorithm development covered at minimum two real falls scenario. The lack of more and various types of real falls restricts the ecologic validity of the result of 100% sensitivity. An open accelerometry database including real world-falls would be of considerable importance for a continuous and independent refinement and validation of fall detection algorithms, as well as for benchmarking.

7. Conclusion

This work has demonstrated the usefulness and added value of a single waist-worn accelerometer to evaluate the functional mobility status of a person, to assess the risk of falling, and to detect falls.

First, I overcame the limitations of the current methods via integration of a customized inertial sensor into the functional assessment routine by:

- procuring objective measurable outcomes with a large resolution that permit the classification of patients lying in the mid-range mobility spectrum and the identification of underlying dysfunctions via analysis of walking pattern,

- offering a portable, continuous, and inexpensive solution for recording functional tests both inside and outside the clinic,

- and potentially reducing the time needed for test administration and data management by using an inertial sensor wirelessly connected to a smart phone/tablet and a medical server.

Second, the feasibility of this technology for fall risk assessment in real clinical practice was proven with a well set-up data collection protocol in a European multicenter study. I developed and validated new algorithms to extract gait and balance features and automated signal processing of the data collected. The statistical analysis of the features was used to develop a new fall prediction model that takes into account important aspects overlooked in previous studies of a similar nature, such as:

- the imbalance between the classes fallers/non-fallers,

- the large overlap of the classes,

- and the high-dimensional data.

The prediction model of this work uses as label classes prospective fallers/non-fallers, unlike most of the studies in this field, which provide a prediction based on retrospective fallers, and are therefore of limited usefulness. From a clinical point of view, the most remarkable finding is that accelerometer-derived features improved fall prediction when compared to the results obtained using conventional features. Moreover, the prediction was considerably accurate taking into account the unbalanced, noisy dataset.

The third and last remark concerns fall detection. In this research, I have developed a fall detection algorithm and a new method to estimate the real upper limit of the false alarm rate, such that:

- to date, this fall detection algorithm, validated using real acceleration data from various subject groups, presents the best performance in terms of false alarm rate when compared with other published algorithms,

- the algorithm achieved 100% sensitivity using data recorded under laboratory conditions as well as the successful detection of real-world falls.

The sensor's capacity to be used as an automatic fall alarm has been proven via exemplary implementation of a fall detection algorithm in the device microcontroller and an android-based application that activates an emergency phone call immediately upon receiving the alarm signal emitted by the device.

A. Literature Review on Step Detection using Waist-Worn Accelerometers

A.1. Search Strategy and Study Selection

To identify eligible publications on human step detection using accelerometers I conducted a systematic search of 3 databases: ACMDL, IEEExplore and Google Scholar, on 25th October 2013 (Table A.1). The search strategy comprised a combination of terms for acceleration sensors and human step analysis.

Name	Discipline	Provider	#Results
ACMDL	Computer Science, Engineering	Association for Computing Machinery	148
IEEE Xplore	Computer Science, Engineering, Electronics	IEEE	64
Google Scholar	Multidisciplinary	Google	67

Table A.1.: Consulted databases and number of results returned

General Rules for Searching. Search restricted to Abstract (or Title in case that restricting to Abstract was not possible).
Search terms for "acceleration sensors":

1. "acceleration sensor"

2. "accelerometer"

3. 1 OR 2

Search terms for "human step analysis":

4. "step"

5. "stride"

6. 4 OR 5

Combination of search terms:

7. 3 AND 6 AND ("human" NOT "animal")

The terms presented above were adapted to include phrases and subject headings specific to each database. Searches covered all type of publications from databases's inception to October 2013 and no language restrictions were imposed. I screened the title and abstract of all identified citations and subsequently assessed full text versions of potentially eligible studies for inclusion according to the selection criteria outlined below. I also screened the references of the eligible studies for identification of additional articles (Figure A.1).

Figure A.1.: Flow-diagram depicting the selection process of studies involving the development of algorithms for step detection using accelerometers.

A.2. Eligibility criteria

To be included in the review, studies must have defined and validated a method for human step detection using a single acceleration sensor attached to the body trunk or head. I did not consider sensors placed on the limbs nor in uncontrolled positions, as for example in pockets, because the accelerations patterns recorded at those positions differ so much than the ones recorded at a near point of the body's COM that a direct comparison among models and algorithm accuracies would be inapplicable.
Only studies published in the English language were considered for inclusion.

A.3. Outcomes of Interest and Assessment of Data Quality

In addition to the primary outcome of interest – performance of the step detector –, other important information extracted from the selected studies comprise the type of study subjects included in the validation (healthy young/adult individuals, elderly, or persons suffering from some motor disorder). To assess the quality of the reported outcome I considered the ecological validity of the test data (i.e., data recorded while walking on treadmill or during overground walking), the validation of the detector at different gait speeds, number of study subjects included, and whether the step detector is able to identify points of interest like for example the heel strike or just count steps based on detection of abrupt changes.

B. VPHOP Data Collection Protocol for Functional Tests

VPHOP Patient Actibelt Clinical Report Form
(To be completed by clinician/experimenter)

Date of testing (dd/mm/yyyy): _____

Initials of observer: _____

Participant ID-Code: _____ (consistent with VPHOP WP8 code)

Age: _____ Years

Weight: _____ kg

Height: _____ cm

Shoe size (European, e.g. 42): _____

Leg length (**Right leg**, from greater trochanter to the floor, with flat shoes, approx):   _____ cm

Time at which Actibelt is turned on: _____
(hh:mm)

Time at which Actibelt is turned off: _____
(hh:mm)

Actibelt Filename: (Patient ID_Actibelt ID number):
_____

Do walking tests require a walking aid? (y/n) _____

Comments (specifically any changes or errors in the data collection e.g.

incorrect   tapping.   Please   include   test   number):

_____

_____

_____

_____

_____

_____

Figure B.1.: Patient clinical form for functional assessment (page 1/2)

1. Normal walking
   ☐   Failed / did not attempt   ☐   Number of tapped events (attempts)
   Record time to walk 10m: _____ s

2. Repeat of test (1). Normal walking
   ☐   Failed / did not attempt   ☐   Number of tapped events (attempts)
   Record time to walk 10m: _____ s

3. Normal walking with patients counting backwards from 100 subtracting 7
   ☐   Failed / did not attempt   ☐   Number of tapped events (attempts)
   Record time to walk 10m: _____ s

4. Tandem walk along a line
   ☐   Failed / did not attempt   ☐   Number of tapped events (attempts)
   Record distance if < 5 m (nearest 0.5 m): _____ m

5. Timed up and go test
   ☐   Failed / did not attempt   ☐   Number of tapped events (attempts)
   Record time: _____ s

6. Chair rise test
   ☐   Failed / did not attempt   ☐   Number of tapped events (attempts)
   Record time to final standing position: _____ s

7. Romberg stance
   ☐   Failed / did not attempt   ☐   Number of tapped events (attempts)
   Record time if < 10 sec: _____ s

8. Semi-Tandem stance
   ☐   Failed / did not attempt   ☐   Number of tapped events (attempts)
   Record time if < 10 sec: _____ s

9. Tandem stance
   ☐   Failed / did not attempt   ☐   Number of tapped events (attempts)
   Record time if < 10 sec: _____ s

10. One-legged stance (standing on right leg)
   ☐   Failed / did not attempt   ☐   Number of tapped events (attempts)
   Record time if < 10 sec: _____ s

11. One-legged stance (standing on left leg)
   ☐   Failed / did not attempt   ☐   Number of tapped events (attempts)
   Record time if < 10 sec: _____ s

*Number of tapped events (attempts) indicates the number of start (single tap) →
stop (double tap) events recorded. The last of these events will be evaluated. In
the case of a non-attempted task, a cross should be indicated in the "failed / did
not attempt box). No task will be evaluated. For all tasks, 0 attempts indicates
that the patient did not attempt the task, 1 attempt is normal for a single task
recording. 2 and above attempts indicates repetitions of the same task.*

Figure B.2.: Patient clinical form for functional assessment (page 2/2)

CLINICAL AND VPHOP RAPID STABILITY TEST CHECKLIST

Protocol for patient testing

1. All patients should wear comfortable flat soled shoes for all tests (or orthopaedic shoes if normally used).
2. Tapping on the belt is used to indicate the start and stop of a test recording. The person in charge of the performance of the test should be the one who taps the belt.
3. For tapping on the belt, a pen or a similar instrument should be used to provide a firm, sharp tap.
4. All tests are automatically recognised by a single tap on the belt to start the test and a double tap to complete the test. If the patient does not or cannot finish a test (e.g. looses their balance before the end of the test), ensure that the test measurement is completed by remaining still, tapping twice on the belt and remaining still once again.
5. An Actibelt and tap test check can be performed by recording a single test trial (pause, single tap, pause, activity, pause, double tap, pause), uploaded to the website, and viewed to ensure that the tap has been automatically recognised.
6. For dynamic (e.g. walking) tests (i.e. test numbers 1-6), a 5 second pause is required (documented in blue) both at the beginning (after the single tap) and at the end of the test activity (before the double tap that indicates the measurement is complete).
7. For static balance tests (i.e. test number 7 onwards), the activity can be commenced directly after the single tap to initiate the test.
8. Once a patient is unable to complete the 10 seconds of balance testing (test number 7 onwards), further testing is not required.
9. All data should be uploaded to the VPHOP website: http://www.vphop.eu/OpenClinica

---

(1) Normal walking – self selected speed, 10 m walk (use space with an additional 2 m tolerance in walk distance; patient start position with toes behind start line. Patients should stop when the leading foot touches or crosses the 10m end line)
    a. Remain still for 5 seconds
    b. Tap once on the buckle to start
    c. Remain still for 5 seconds
    d. Perform 10 m walk, normal speed, stop at end line. Record time.
    e. Remain still for 5 seconds
    f. Tap twice on the belt to stop
    g. Remain still for 5 seconds

(2) Repeat of test (1). Normal walking – self selected speed, 10 m walk (use space with an additional 2 m tolerance in walk distance; patient start position with toes behind start line)
    a. Remain still for 5 seconds
    b. Tap once on the buckle to start
    c. Remain still for 5 seconds
    d. Perform 10 m walk, normal speed, stop at end line. Record time.
    e. Remain still for 5 seconds
    f. Tap twice on the belt to stop
    g. Remain still for 5 seconds

1

Figure B.3.: List of functional tests and protocol for patient testing (page 1/4)

(3) Normal walking – self selected speed, 10 m walk with patients counting backwards from 100 subtracting 7 and stating the numbers out aloud (use space with an additional 2 m tolerance in walk distance; patient start position with toes behind start line)
   a. Remain still for 5 seconds
   b. Tap once on the buckle to start
   c. Remain still for 5 seconds
   d. Perform 10 m walk, normal speed, stop at end line. Record time.
   e. Remain still for 5 seconds
   f. Tap twice on the belt to stop
   g. Remain still for 5 seconds

(4) Tandem walk along a line. Arms folded across chest. 5m walk, self selected speed
   a. Remain still for 5 seconds
   b. Tap once on the buckle to start
   c. Remain still for 5 seconds
   d. 5 m walk in one line with one foot leading the other
   e. Record distance if < 5 m (nearest 0.5 m)
   f. Remain still for 5 seconds
   g. Tap twice on the belt to stop
   h. Remain still for 5 seconds

These dynamic motion tests require the additional 5 second pause before motion begins and after motion ends, to ensure correction recognition of the activity

(5) Timed up and go test (TUG) at self-selected speed (hard surfaced chair with arm and back supports, 42 cm in seating height)
   a. Remain still for 5 seconds
   b. Tap once on the buckle to start
   c. Remain still for 5 seconds
   d. Sit to stand (patients may use arm supports or walking aid)
   e. 3 m walk
   f. Turn around
   g. 3 m walk back
   h. Stand to sit
   i. Record time (start clock when patient moves from back support, stop clock when patient returns to back support)
   j. Remain still for 5 seconds
   k. Tap twice on the belt to stop
   l. Remain still for 5 seconds

(6) Chair rise test. Arms folded across chest (hard surfaced chair with arm and back supports, 42 cm in seating height)
   a. Remain still for 5 seconds
   b. Tap once on the buckle to start
   c. Remain still for 5 seconds
   d. Stand up from a sitting position (no arm support allowed) and sit down five times as quickly as possible. Count repetitions aloud to encourage patient.
   e. Record time from intial movement to final standing position
   f. Remain still for 5 seconds
   g. Tap twice on the belt to stop

(7) Romberg (quiet-standing) stance (10 sec, focussing on point on wall at eye height, patients may move arms to control balance, feet together, ankles touching)
   a. Remain still for 5 seconds

2

Figure B.4.: List of functional tests and protocol for patient testing (page 2/4)

    b. Ask the patient to adopt the Quiet Stance posture (stand with ankles touching each other, holding a support)

    c. Immediately tap once on the buckle to start, patient releases support hold.

    d. Perform 10 seconds of quiet stance. Patient to remain as still as possible

    e. Record time if < 10 sec (start clock when patient releases support, stop clock when patient looses balance/holds support)

    f. Tap twice on the buckle to end (patient still in Quiet Stance posture)

    g. Remain still for 5 seconds (it's not necessary to remain in Quiet Stance posture)

(8) Semi-Tandem stance (max. 10 sec, focussing on point on wall at eye height, patients may move arms to control balance, chosen leg in front)

    a. Remain still for 5 seconds

    b. Ask the patient to adopt the Semi-Tandem posture (heel of one foot alongside the big toe of the other foot, holding a support)

    c. Tap once on the buckle to start, patient releases support hold.

    d. Perform 10 seconds of Semi-Tandem stance. Patient to remain as still as possible

    e. Record time if < 10 sec (start clock when patient releases support, stop clock when patient looses balance/holds support)

    f. Tap twice on the buckle to end (patient still in Semi-Tandem posture)

    g. Remain still for 5 seconds (not necessary for patient to remain in Semi-Tandem posture)

(9) Tandem stance (max. 10 sec, focussing on point on wall at eye height, patients may move arms to control balance, chosen leg in front)

    a. Remain still for 5 seconds

    b. Ask the patient to adopt the Tandem posture (stand with one foot directly in front of the other , holding a support)

    c. Tap once on the buckle to start, patient releases support hold.

    d. Perform 10 sec of tandem stance. Patient to remain as still as possible

    e. Record time if < 10 sec (start clock when patient releases support, stop clock when patient looses balance/holds support)

    f. Tap twice on the buckle to end (patient still in Tandem posture)

    g. Remain still for 5 seconds (it's not necessary to remain in Tandem posture)

(10) One-legged stance (max. 10 sec, focussing on point on wall at eye height, patients may move arms to control balance, standing on **right** leg)

    a. Remain still for 5 seconds

    b. Ask the patient to adopt the one-legged posture, holding a support

    c. Tap once on the buckle to start, patient releases support hold.

    d. Perform 10 seconds of one-legged stance on **right** leg. Patient to remain as still as possible

    e. Record time if < 10 sec (start clock when patient releases support, stop clock when patient looses balance/holds support)

    f. Tap twice on the buckle to end (patient still in one-legged posture)

    g. Remain still for 5 seconds (not necessary to remain in one-legged posture)

3

Figure B.5.: List of functional tests and protocol for patient testing (page 3/4)

(11) One-legged stance (max. 10 sec, focussing on point on wall at eye height, patients may move arms to control balance, standing on **left** leg)

    a. Remain still for 5 seconds
    b. Ask the patient to adopt the one-legged posture, holding a support
    c. Tap once on the buckle to start, patient releases support hold.
    d. Perform 10 seconds of one-legged stance on **left** leg. Patient to remain as still as possible
    e. Record time if < 10 sec (start clock when patient releases support, stop clock when patient looses balance/holds support)
    f. Tap twice on the buckle to end (patient still in one-legged posture)
    g. Remain still for 5 seconds (not necessary to remain in one-legged posture)

Figure B.6.: List of functional tests and protocol for patient testing (page 4/4)

C. Data Sets Variables

C.1. Dataset $\mathcal{A}$

Variable	Unit	Scale	Reference
History of Falls	yes/no	nominal	AGS [169]
Age	years	ratio	AGS [169]
Depression	yes/no	nominal	AGS [169]
Sleep Disturbance	yes/no	nominal	AGS [169]
Assistive Devices	yes/no	nominal	AGS [169]
Arthritis	yes/no	nominal	AGS [169]
Visual Deficits	yes/no	nominal	AGS [169]
BMI	kg/m^2	ratio	Volpato et al. [188]
Hypertension	yes/no	nominal	Gangavati et al. [66]
Fear of Falling	yes/no	nominal	Scheffer et al. [164]
Foot Pain	yes/no	nominal	Todd et al. [181]
Time 10 Meters	seconds	ratio	AGS [169]
Balance	levels 1 - 10	ordinal	AGS [169]
Reaching	levels 1 - 3	ordinal	AGS [169]
Carrying	levels 1 - 3	ordinal	AGS [169]
Washing	levels 1 - 3	ordinal	AGS [169]
Bending	levels 1 - 3	ordinal	AGS [169]

Table C.1.: Variables of data set $\mathcal{A}$ [14]

C.2. Dataset $\mathcal{B}$

Table C.2 presents a list of the variables included in data set $\mathcal{B}$. Information about units and scales can be found in section 5.5.2. Variables in bold letters are variables selected by an expert for the prediction of falls. See section 5.6.5 for more information.

C.3. Variables Chosen by Elastic Nets

This section provides variables from dataset $\mathcal{B}$ chosen by EN for a reduction target to 12 and 30 features.

History of Falls	Test 1: Step cadence	Test 3: Step cadence
Age	Test 2: Number of Steps	Test 4: Peak to Peak in X
Depression	Test 2: X-Asymmetry	Test 4: Peak to Peak in Y
Sleep Disturbance	Test 2: Y-Asymmetry	Test 4: Peak to Peak in Z
Assistive Devices	Test 2: Z-Asymmetry	Test 4: SD of X-Acceleration
Arthritis	Test 2: SD of X-Asymmetry	Test 4: SD of Y-Acceleration
Visual Deficits	Test 2: SD of Y-Asymmetry	Test 4: SD of Z-Acceleration
BMI	Test 2: SD of Z-Asymmetry	Test 4: Balance Count X
Hypertension	Test 2: Mean Speed	Test 4: Balance Count Y
Fear of Falling	Test 2: Duration	**Test 4: Balance Count Z**
Foot Pain	Test 2: Step Length	**Test 5: Duration**
Balance	Test 2: Step cadence	**Test 6: Duration**
Test 1: Number of Steps	Test 3: Number of Steps	Test 7: Eccentricity
Test 1: X-Asymmetry	Test 3: X-Asymmetry	Test 7: Area
Test 1: Y-Asymmetry	Test 3: Y-Asymmetry	Test 8: Eccentricity
Test 1: Z-Asymmetry	Test 3: Z-Asymmetry	Test 8: Area
Test 1: SD of X-Asymmetry	Test 3: SD of X-Asymmetry	Test 9: Eccentricity
Test 1: SD of Y-Asymmetry	Test 3: SD of Y-Asymmetry	**Test 9: Area**
Test 1: SD of Z-Asymmetry	Test 3: SD of Z-Asymmetry	Test 10: Eccentricity
Test 1: Mean Speed	Test 3: Mean Speed	**Test 10: Area**
Test 1: Duration	Test 3: Duration	Test 11: Eccentricity
Test 1: Step Length	Test 3: Step Length	**Test 11: Area**

Table C.2.: Variables of Dataset $\mathcal{B}$[14]

Visual Deficits	Hypertension	Foot Pain
Test 1: SD of Z-Asymmetry	Test 2: SD of X-asymmetry	Test 3: Z-Asymmetry
Test 4: SD of Y-Asymmetry	Test 5: Duration	Test 6: Duration
Test 8: Eccentricity	Test 9: Area	Test 11: Area

Table C.3.: 12 Variables of Dataset $\mathcal{B}$, chosen by EN [14]

Age	Sleep Disturbance	Arthritis
Visual Deficits	Hypertension	Fear of Falling
Foot Pain	Balance	BMI
Test 1: SD of Y-Asymmetry	Test 1: SD of Z-Asymmetry	Test 1: Step Cadence
Test 2: SD of X-Asymmetry	Test 2: Y-Asymmetry	Test 3: Number of Steps
Test 3: Z-Asymmetry	Test 3: Mean Speed	Test 4: SD of Y-Acceleration
Test 4: Balance Count X	Test 4: Balance Count Z	Test 5: Duration
Test 6: Duration	Test 7: Eccentricity	Test 7: Area
Test 8: Eccentricity	Test 8: Area	Test 9: Area
Test 10: Eccentricity	Test 10: Area	Test 11: Area

Table C.4.: 30 Variables of Dataset $\mathcal{B}$, chosen by EN [14]

List of Figures

List of Tables

List of Tables

Nomenclature

ADL   Activities of Daily Living

ASG   Average Step Graph

BMI   Body Mass Index

CHA   Charité - Universitätsmedizin Berlin

CoM   Center of Mass

CoP   Center of Pressure

CSG   Characteristic Step Graph

EN   Elastic Nets

GRF   Ground Reaction Force

INSERM Institut National de la Santé et de la Recherche Médicale

IOR   Istituto Ortopedico Rizzoli

kNN   k Nearest Neighbors

LDV   Fachgebiet für Geometrische Optimierung und Maschinelles Lernen

LDV   Lehrstuhl für Datenverarbeitung

PCA   Principal Component Analysis

SLC   Sylvia Lawry Centre for Multiple Sclerosis Research e.V.

sPLS   Sparse Partial Least Squares

SVM   Support Vector Machines

UGE   University of Geneva

VPHOP The Osteoporotic Virtual Physiological Human Project

Bibliography

1. Functional assessment definition. In *Mosby's Medical Dictionary, 8th edition*. URL `http://medical-dictionary.thefreedictionary.com/functional+assessment`.

2. Illustration of methods for numerical integration. Public Domain CC. (Original photo by Michael Richmond). URL `http://www.cs.utsa.edu/~wagner/CS2073/integral/num_integ2.html#rect`.

3. Photo of Walking Person Silhouette. Public Domain CC0 (Original photo by OCAL). URL `http://www.pixabay.com`.

4. Top causes of death according to the World Health Organization report for calendar year 2001. Ucatlas.ucsc.edu (2004). URL `http://ucatlas.ucsc.edu/cause.php`.

5. Photo of the actibelt sensor on a proband. In *VPHOP meeting*, Berlin, 2009.

6. S. Abbate, M. Avvenuti, G. Cola, P. Corsini, J. Light, and A. Vecchio. Recognition of false alarms in fall detection systems. In *Consumer Communications and Networking Conference (CCNC), 2011 IEEE*, pp. 23–28. IEEE, 2011.

7. A. Abran, A. Khelifi, W. Suryn, and A. Seffah. Usability meanings and interpretations in ISO standards. In *Software Quality Journal*, 11(4), pp. 325–338, 2003.

8. S.G.o.t.E.I.P. on active andhealthy ageing. In Strategic Implementation Plan For The European Innovation Partnership On Active And Healthy Ageing, 2011.

9. W.H.O. Ageing and L.C. Unit. *WHO global report on falls prevention in older age*. World Health Organization, 2008.

10. M. Alwan, P.J. Rajendran, S. Kell, D. Mack, S. Dalal, M. Wolfe, and R. Felder. A smart and passive floor-vibration based fall detector for elderly. In *Information and Communication Technologies, 2006. ICTTA'06. 2nd*, volume 1, pp. 1003–1007. IEEE, 2006.

11. D. Anderson, J.M. Keller, M. Skubic, X. Chen, and Z. He. Recognizing falls from silhouettes. In *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*, pp. 6388–6391. IEEE, 2006.

12. A. Ariani, S.J. Redmond, D. Chang, and N.H. Lovell. Software simulation of unobtrusive falls detection at night-time using passive infrared and pressure mat sensors. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pp. 2115–2118. IEEE, 2010.

13. K. Atkinson, F. Coutts, and A.M. Hassenkamp. *Physiotherapy in orthopaedics: a problem-solving approach*. Elsevier Health Sciences, 2005.

14. C. Bachhuber. Evaluation of Acceleration Data for Falls Risk Prediction using Classification and Regression Analysis. In *Masterarbeit in Elektrotechnik. Technische Universität München*, 2014.

15. F. Bagalà, C. Becker, A. Cappello, L. Chiari, K. Aminian, J.M. Hausdorff, W. Zijlstra, and J. Klenk. Evaluation of accelerometer-based fall detection algorithms on real-world falls. In *PLoS One*, 7(5), p. e37062, 2012.

16. R. Batuwita and V. Palade. Class imbalance learning methods for support vector machines. In *Imbalanced Learning: Foundations, Algorithms, and Applications*, 83, 2013.

17. B.J. Benda, P. Riley, and D. Krebs. Biomechanical relationship between center of gravity and center of pressure during standing. In *Rehabilitation Engineering, IEEE Transactions on*, 2(1), pp. 3–10, 1994.

18. K. Berg. Measuring balance in the elderly: preliminary development of an instrument. In *Physiotherapy Canada*, 41(6), pp. 304–311, 1989.

19. J.E. Bertram and A. Ruina. Multiple walking speed–frequency relations are predicted by constrained optimization. In *Journal of theoretical Biology*, 209(4), pp. 445–453, 2001.

20. R. Blickhan. The spring-mass model for running and hopping. In *Journal of biomechanics*, 22(11), pp. 1217–1227, 1989.

21. A. Bottaro, M. Casadio, P.G. Morasso, and V. Sanguineti. Body sway during quiet standing: is it the residual chattering of an intermittent stabilization process? In *Human movement science*, 24(4), pp. 588–615, 2005.

22. A. Bourke, J. Obrien, and G. Lyons. Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm. In *Gait & posture*, 26(2), pp. 194–199, 2007.

23. A. Bourke, P. Van de Ven, M. Gamble, R. OConnor, K. Murphy, E. Bogan, E. McQuade, P. Finucane, G. Olaighin, and J. Nelson. Evaluation of waist mounted triaxial accelerometer based fall detection algorithms during scripted and continuous unscripted activities. In *Journal of biomechanics*, 43(15), pp. 3051–3057, 2010.

24. A.K. Bourke, P.W. Van de Ven, A.E. Chaya, G.M. ÓLaighin, and J. Nelson. Testing of a long-term fall detection system incorporated into a custom vest for the elderly. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pp. 2844–2847. IEEE, 2008.

25. 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. In *Biomedical Engineering, IEEE Transactions on*, 44(3), pp. 136–147, 1997.

26. J. Boyle and M. Karunanithi. Simulated fall detection via accelerometers. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pp. 1274–1277. IEEE, 2008.

27. J.S. Brach, S.A. Studenski, S. Perera, J.M. VanSwearingen, and A.B. Newman. Gait variability and the risk of incident mobility disability in community-dwelling older adults. In *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 62(9), pp. 983–988, 2007.

28. D. Brown. Tracker video analysis and modeling tool. In *Open source, physics*, 4, 2009.

29. D. Bryant, S. Ravindran, N. Magotra, and S. Northrup. Real-time implementation of a chest-worn accelerometer based heart monitoring system. In *Circuits and Systems (MWSCAS), 2010 53rd IEEE International Midwest Symposium on*, pp. 1057–1060. IEEE, 2010.

30. A. Bueno-Cavanillas, F. Padilla-Ruiz, J. Jimenez-Moleon, C. Peinado-Alonso, and R. Galvez-Vargas. Risk factors in falls among the elderly according to extrinsic and intrinsic precipitating causes. In *European journal of epidemiology*, 16(9), pp. 849–859, 2000.

31. B. Caby, S. Kieffer, M. de Saint Hubert, G. Cremer, and B. Macq. Feature extraction and selection for objective gait analysis and fall risk assessment by accelerometry. In *Biomedical engineering online*, 10(1), p. 1, 2011.

32. A.J. Campbell, M.J. Borrie, and G.F. Spears. Risk factors for falls in a community-based prospective study of people 70 years and older. In *Journal of gerontology*, 44(5), pp. M112–M117, 1989.

33. O. Caron, B. Faure, and Y. Brenière. Estimating the centre of gravity of the body on the basis of the centre of pressure in standing posture. In *Journal of biomechanics*, 30(11), pp. 1169–1171, 1997.

34. N.V. Chawla, K.W. Bowyer, L.O. Hall, and W.P. Kegelmeyer. SMOTE: synthetic minority over-sampling technique. In *arXiv preprint arXiv:1106.1813*, 2011.

35. J. Chen, K. Kwong, D. Chang, J. Luk, and R. Bajcsy. Wearable sensors for reliable fall detection. In *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, pp. 3551–3554. IEEE, 2006.

36. H. Chun and S. Keleş. Sparse partial least squares regression for simultaneous dimension reduction and variable selection. In *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(1), pp. 3–25, 2010.

37. J. Cohen. *Statistical power analysis for the behavioral sciences*. Psychology Press, 1988.

38. B. Colobert, A. Crétual, P. Allard, and P. Delamarche. Force-plate based computation of ankle and hip strategies from double-inverted pendulum model. In *Clinical Biomechanics*, 21(4), pp. 427–434, 2006.

39. E.P. Cox III. The optimal number of response alternatives for a scale: A review. In *Journal of marketing research*, pp. 407–422, 1980.

40. K. Culhane, M. O'Connor, D. Lyons, and G. Lyons. Accelerometers in rehabilitation medicine for older adults. In *Age and ageing*, 34(6), pp. 556–560, 2005.

41. S.R. Cummings, M.C. Nevitt, and S. Kidd. Forgetting falls. The limited accuracy of recall of falls in the elderly. In *Journal of the American Geriatrics Society*, 36(7), p. 613, 1988.

42. M. Daumer and A. Neiss. A new adaptive algorithm to detect shifts, drifts and outliers in biomedical time series. In *Mathematical statistics with applications in biometry. Lohmar: Josef Eul Verlag*, 2001.

43. M. Daumer, K. Thaler, E. Kruis, W. Feneberg, G. Staude, and M. Scholz. Steps towards a miniaturized, robust and autonomous measurement device for the long-term monitoring of patient activity: ActiBelt. In *Biomedizinische Technik*, 52(1), pp. 149–155, 2007.

44. P.J. Davis and P. Rabinowitz. *Methods of numerical integration*. Courier Dover Publications, 2007.

45. B. Day, M. Steiger, P. Thompson, and C. Marsden. Effect of vision and stance width on human body motion when standing: implications for afferent control of lateral sway. In *The Journal of physiology*, 469(1), pp. 479–499, 1993.

46. S. Deandrea, E. Lucenteforte, F. Bravi, R. Foschi, C. La Vecchia, and E. Negri. Risk factors for falls in community-dwelling older people: a systematic review and meta-analysis. In *Epidemiology*, 21(5), pp. 658–668, 2010.

47. U. DESA. World population prospects: the 2008 revision. In *New York: Department for Economic and Social Affairs*, 2009.

48. P.M. Deshmukh, C.M. Russell, L.E. Lucarino, and S.N. Robinovitch. Enhancing clinical measures of postural stability with wearable sensors. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pp. 4521–4524. IEEE, 2012.

49. B. Dijkstra, W. Zijlstra, E. Scherder, and Y. Kamsma. Detection of walking periods and number of steps in older adults and patients with Parkinson's disease: accuracy of a pedometer and an accelerometry-based method. In *Age and ageing*, 37(4), pp. 436–441, 2008.

50. E.P. Doheny, D. McGrath, B.R. Greene, L. Walsh, D. McKeown, C. Cunningham, L. Crosby, R.A. Kenny, and B. Caulfield. Displacement of centre of mass during quiet standing assessed using accelerometry in older fallers and non-fallers. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pp. 3300–3303. IEEE, 2012.

51. E.P. Doheny, D. McGrath, B.R. Greene, L. Walsh, D. McKeown, C. Cunningham, L. Crosby, R.A. Kenny, and B. Caulfield. Displacement of centre of mass during quiet standing assessed using accelerometry in older fallers and non-fallers. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pp. 3300–3303. IEEE, 2012.

52. T. Doi, S. Hirata, R. Ono, K. Tsutsumimoto, S. Misu, and H. Ando. The harmonic ratio of trunk acceleration predicts falling among older people: results of a 1-year prospective study. In *J Neuroeng Rehabil*, 10(7), 2013.

53. I. Dryden and K. Mardia. Multivariate shape analysis. In *Sankhyā: The Indian Journal of Statistics, Series A*, pp. 460–480, 1993.

54. P.W. Duncan, D.K. Weiner, J. Chandler, and S. Studenski. Functional reach: a new clinical measure of balance. In *Journal of gerontology*, 45(6), pp. M192–M197, 1990.

55. S. Elliott, J. Painter, and S. Hudson. Living alone and fall risk factors in community-dwelling middle age and older adults. In *Journal of community health*, 34(4), pp. 301–310, 2009.

56. J. Eng and D. Winter. Estimations of the horizontal displacement of the total body centre of mass: considerations during standing activities. In *Gait & Posture*, 1(3), pp. 141–144, 1993.

57. D.D. Espy, F. Yang, T. Bhatt, and Y.C. Pai. Independent influence of gait speed and step length on stability and fall risk. In *Gait & posture*, 32(3), pp. 378–382, 2010.

58. P. Esser, H. Dawes, J. Collett, and K. Howells. IMU: inertial sensing of vertical CoM movement. In *Journal of biomechanics*, 42(10), pp. 1578–1581, 2009.

59. J.M. Fabre. *Identification of falls risk ractors in community-dwelling older adults: validation of the comprehensive falls risk screening instrument*. Ph.D. thesis, Louisiana State University, May 2009.

60. J. Fleming and C. Brayne. Inability to get up after falling, subsequent time on floor, and summoning help: prospective cohort study in people over 90. In *BMJ: British Medical Journal*, 337, 2008.

61. K.M. Foglyano, J.R. Schnellenberger, and R. Kobetic. Development of a self-contained accelerometry based system for control of functional electrical stimulation in hemiplegia. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pp. 5448–5451. IEEE, 2011.

62. R.C. Foster, L.M. Lanningham-Foster, C. Manohar, S.K. McCrady, L.J. Nysse, K.R. Kaufman, D.J. Padgett, and J.A. Levine. Precision and accuracy of an ankle-worn accelerometer-based pedometer in step counting and energy expenditure. In *Preventive medicine*, 41(3), pp. 778–783, 2005.

63. F. Franzelin. IT-Tools zur On- und Offline Analyse von Einzelschritten die mit dem hochaulösenden 3D-Akzelerometer actibelt aufgezeichnet werden. In *Bachelorarbeit in Informatik. Technische Universität München*, 2009.

64. S.M. Friedman, B. Munoz, S.K. West, G.S. Rubin, and L.P. Fried. Falls and fear of falling: which comes first? A longitudinal prediction model suggests strategies for primary and secondary prevention. In *Journal of the American Geriatrics Society*, 50(8), pp. 1329–1335, 2002.

65. A. Gabell and U. Nayak. The effect of age on variability in gait. In *Journal of Gerontology*, 39(6), pp. 662–666, 1984.

66. A. Gangavati, I. Hajjar, L. Quach, R.N. Jones, D.K. Kiely, P. Gagnon, and L.A. Lipsitz. Hypertension, Orthostatic Hypotension, and the Risk of Falls in a Community-Dwelling Elderly Population: The Maintenance of Balance, Independent Living, Intellect, and Zest in the Elderly of Boston Study. In *Journal of the American Geriatrics Society*, 59(3), pp. 383–389, 2011.

67. D. Giansanti, V. Macellari, and G. Maccioni. New neural network classifier of fall-risk based on the Mahalanobis distance and kinematic parameters assessed by a wearable device. In *Physiological measurement*, 29(3), p. N11, 2008.

68. M. Gietzelt, G. Nemitz, K.H. Wolf, H. Meyer Zu Schwabedissen, R. Haux, and M. Marschollek. A clinical study to assess fall risk using a single waist accelerometer. In *Informatics for health and social care*, 34(4), pp. 181–188, 2009.

69. B.R. Greene, A.O. Donovan, R. Romero-Ortuno, L. Cogan, C.N. Scanaill, and R.A. Kenny. Quantitative falls risk assessment using the timed up and go test. In *Biomedical Engineering, IEEE Transactions on*, 57(12), pp. 2918–2926, 2010.

70. E.W. Gregg, M.A. Pereira, and C.J. Caspersen. Physical activity, falls, and fractures among older adults: a review of the epidemiologic evidence. In *Journal of the American Geriatrics Society*, 48(8), pp. 883–893, 2000.

71. S. Hargittai. Savitzky-Golay least-squares polynomial filters in ECG signal processing. In *Computers in Cardiology, 2005*, pp. 763–766. IEEE, 2005.

72. J.A. Hartigan and M.A. Wong. Algorithm AS 136: A k-means clustering algorithm. In *Applied statistics*, pp. 100–108, 1979.

73. S.S. Hasan, D.W. Robin, D.C. Szurkus, D.H. Ashmead, S.W. Peterson, and R.G. Shiavi. Simultaneous measurement of body center of pressure and center of gravity during upright stance. Part I: Methods. In *Gait & posture*, 4(1), pp. 1–10, 1996.

74. M.R. Haug and S.J. Folmar. Longevity, gender, and life quality. In *Journal of Health and Social Behavior*, pp. 332–345, 1986.

75. J.M. Hausdorff, M.E. Cudkowicz, R. Firtion, J.Y. Wei, and A.L. Goldberger. Gait variability and basal ganglia disorders: Stride-to-stride variations of gait cycle timing in parkinson's disease and Huntington's disease. In *Movement disorders*, 13(3), pp. 428–437, 1998.

76. J.M. Hausdorff, D.A. Rios, and H.K. Edelberg. Gait variability and fall risk in community-living older adults: a 1-year prospective study. In *Archives of physical medicine and rehabilitation*, 82(8), pp. 1050–1056, 2001.

77. V.T. van Hees, L. Gorzelniak, E.C.D. Leon, M. Eder, M. Pias, S. Taherian, U. Ekelund, F. Renström, P.W. Franks, A. Horsch et al.. Separating movement and gravity components in an acceleration signal and implications for the assessment of human daily physical activity. In *PloS one*, 8(4), p. e61691, 2013.

78. A. Hendrich, A. Nyhuis, T. Kippenbrock, and M.E. Soja. Hospital falls: Development of a predictive model for clinical practice. In *Applied Nursing Research*, 8(3), pp. 129–139, 1995.

79. A.E. Hoerl and R.W. Kennard. Ridge regression: Biased estimation for nonorthogonal problems. In *Technometrics*, 12(1), pp. 55–67, 1970.

80. F.B. Horak, S.M. Henry, and A. Shumway-Cook. Postural perturbations: new insights for treatment of balance disorders. In *Physical therapy*, 77(5), pp. 517–533, 1997.

81. F.B. Horak, C.L. Shupert, and A. Mirka. Components of postural dyscontrol in the elderly: a review. In *Neurobiology of aging*, 10(6), pp. 727–738, 1989.

82. J. Howcroft, J. Kofman, and E.D. Lemaire. Review of fall risk assessment in geriatric populations using inertial sensors. In *J Neuroeng Rehabil*, 10, p. 91, 2013.

83. A. Hreljac. Impact and overuse injuries in runners. In *Medicine and Science in Sports and Exercise*, 36(5), pp. 845–849, 2004.

84. B. Huang, G. Tian, and X. Li. A method for fast fall detection. In *Intelligent Control and Automation, 2008. WCICA 2008. 7th World Congress on*, pp. 3619–3623. IEEE, 2008.

85. Y. Huang, H. Zheng, C. Nugent, P. McCullagh, S.M. McDonough, M.A. Tully, and S.O. Connor. Activity monitoring using an intelligent mobile phone: a validation study. In *Proceedings of the 3rd International Conference on PErvasive Technologies Related to Assistive Environments*, p. 10. ACM, 2010.

86. R. Igual, C. Medrano, and I. Plaza. Challenges, issues and trends in fall detection systems. In *Biomedical engineering online*, 12(1), p. 66, 2013.

87. S.R. Jacob Goldberger and R.S. Geoff Hinton. Neighbourhood components analysis. In *NIPS'04*, 2004.

88. W.G. Janssen, J.B. Bussmann, H.L. Horemans, and H.J. Stam. Analysis and decomposition of accelerometric signals of trunk and thigh obtained during the sit-to-stand movement. In *Medical and Biological Engineering and Computing*, 43(2), pp. 265–272, 2005.

89. J. Jensen, L. Nyberg, Y. Gustafson, and L. Lundin-Olsson. Fall and injury prevention in residential care—effects in residents with higher and lower levels of cognition. In *Journal of the American geriatrics society*, 51(5), pp. 627–635, 2003.

90. K. Johnston, K. Grimmer-Somers, and M. Sutherland. Perspectives on use of personal alarms by older fallers. In *International journal of general medicine*, 3, p. 231, 2010.

91. R.L. Kane, M.G. Saslow, and T. Brundage. Using ADLs to establish eligibility for long-term care among the cognitively impaired. In *The Gerontologist*, 31(1), pp. 60–66, 1991.

92. M. Kangas, A. Konttila, P. Lindgren, I. Winblad, and T. Jämsä. Comparison of low-complexity fall detection algorithms for body attached accelerometers. In *Gait & posture*, 28(2), pp. 285–291, 2008.

93. M. Kangas, A. Konttila, I. Winblad, and T. Jamsa. Determination of simple thresholds for accelerometry-based parameters for fall detection. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pp. 1367–1370. IEEE, 2007.

94. P. Kannus, J. Parkkari, S. Koskinen, S. Niemi, M. Palvanen, M. Järvinen, and I. Vuori. Fall-induced injuries and deaths among older adults. In *Jama*, 281(20), pp. 1895–1899, 1999.

95. D.M. Karantonis, M.R. Narayanan, M. Mathie, N.H. Lovell, and B.G. Celler. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. In *Information Technology in Biomedicine, IEEE Transactions on*, 10(1), pp. 156–167, 2006.

96. C. Kirtley, M.W. Whittle, and R. Jefferson. Influence of walking speed on gait parameters. In *Journal of Biomedical Engineering*, 7(4), pp. 282–288, 1985.

97. N. Koceska and S. Koceski. Methodology for calculating the vertical displacement of the body center of mass using mobile phone accelerometer. In *Embedded Computing (MECO), 2012 Mediterranean Conference on*, pp. 136–139. IEEE, 2012.

98. R. Kohavi et al.. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *IJCAI*, volume 2, pp. 1137–1145. 1995.

99. M. Kojima, S. Obuchi, O. Henmi, and N. Ikeda. Comparison of Smoothness during Gait between Community Dwelling Elderly Fallers and Non-Fallers Using Power Spectrum Entropy of Acceleration Time-Series. In *Journal of Physical Therapy Science*, 20(4), pp. 243–248, 2008. doi:10.1589/jpts.20.243.

100. N. König, W. Taylor, G. Armbrecht, R. Dietzel, and N. Singh. Identification of functional parameters for the classification of older female fallers and prediction of 'first-time' fallers. In *Journal of The Royal Society Interface*, 11(97), p. 20140353, 2014.

101. S. Kotsiantis, D. Kanellopoulos, P. Pintelas et al.. Handling imbalanced datasets: A review. In *GESTS International Transactions on Computer Science and Engineering*, 30(1), pp. 25–36, 2006.

102. K. Kowalski, R. Rhodes, P.J. Naylor, H. Tuokko, S. MacDonald et al.. Direct and indirect measurement of physical activity in older adults: a systematic review of the literature. In *International Journal of Behavioral Nutrition and Physical Activity*, 9(1), p. 148, 2012.

103. J.B. Kruskal and M. Wish. *Multidimensional scaling*, volume 11. Sage, 1978.

104. P. Kulkarni and Y. Öztürk. Requirements and design spaces of mobile medical care. In *ACM SIGMOBILE Mobile Computing and Communications Review*, 11(3), pp. 12–30, 2007.

105. H. Lach, A. Reed, C.L. Arfken, J. Miller, G. Paige, S. Birge, and W. Peck. Falls in the elderly: reliability of a classification system. In *Journal of the American Geriatrics Society*, 39(2), pp. 197–202, 1991.

106. D. Lafond, M. Duarte, and F. Prince. Comparison of three methods to estimate the center of mass during balance assessment. In *Journal of biomechanics*, 37(9), pp. 1421–1426, 2004.

107. S.E. Lamb, E.C. Jørstad-Stein, K. Hauer, and C. Becker. Development of a common outcome data set for fall injury prevention trials: the Prevention of Falls Network Europe consensus. In *Journal of the American Geriatrics Society*, 53(9), pp. 1618–1622, 2005.

108. H.K. Lee, J. You, S.P. Cho, S.J. Hwang, D.R. Lee, Y.H. Kim, and K.J. Lee. Computational methods to detect step events for normal and pathological gait evaluation using accelerometer. In *Electronics letters*, 46(17), pp. 1185–1187, 2010.

109. H.K. Lee, S.J. Hwang, S.P. Cho, D.R. Lee, S.H. You, K.J. Lee, Y.H. Kim, and H.S. Choi. Novel algorithm for the hemiplegic gait evaluation using a single 3-axis accelerometer. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pp. 3964–3966. IEEE, 2009.

110. Y.S. Lee and H. Lee. Multiple object tracking for fall detection in real-time surveillance system. In *Advanced Communication Technology, 2009. ICACT 2009. 11th International Conference on*, volume 3, pp. 2308–2312. IEEE, 2009.

111. Y. Lee, J. Kim, M. Son, and J.H. Lee. Implementation of accelerometer sensor module and fall detection monitoring system based on wireless sensor network. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pp. 2315–2318. IEEE, 2007.

112. N. Lemoine. In Drawing a 95% confidence interval in R. URL `http://climateecology.wordpress.com`.

113. B. Lent. Practical considerations of accelerometers noise. In *Endevco Website, March*, 2001.

114. Q. Li, J.A. Stankovic, M.A. Hanson, A.T. Barth, J. Lach, and G. Zhou. Accurate, fast fall detection using gyroscopes and accelerometer-derived posture information. In *Wearable and Implantable Body Sensor Networks, 2009. BSN 2009. Sixth International Workshop on*, pp. 138–143. IEEE, 2009.

115. D. Liang, G. Zhao, Y. Guo, and L. Wang. Pre-impact & impact detection of falls using wireless body sensor network. In *Biomedical and Health Informatics (BHI), 2012 IEEE-EMBS International Conference on*, pp. 763–766. IEEE, 2012.

116. D. Litvak, Y. Zigel, and I. Gannot. Fall detection of elderly through floor vibrations and sound. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pp. 4632–4635. IEEE, 2008.

117. B.J. Logue. Modernization and the status of the frail elderly: perspectives on continuity and change. In *Journal of cross-cultural gerontology*, 5(4), pp. 345–374, 1990.

118. S. MacAvoy, T. Skinner, and M. Hines. Fall risk assessment tool. In *Applied Nursing Research*, 9(4), pp. 213–218, 1996.

119. L. Mackenzie, J. Byles, and C. D'Este. Validation of self-reported fall events in intervention studies. In *Clinical rehabilitation*, 20(4), pp. 331–339, 2006.

120. B.E. Maki, P.J. Holliday, and A.K. Topper. A prospective study of postural balance and risk of falling in an ambulatory and independent elderly population. In *Journal of Gerontology*, 49(2), pp. M72–M84, 1994.

121. M. Marschollek, A. Rehwald, K.H. Wolf, M. Gietzelt, G. Nemitz, H.M. zu Schwabedissen, and M. Schulze. Sensors vs. experts - a performance comparison of sensor-based fall risk assessment vs. conventional assessment in a sample of geriatric patients. In *BMC Med Inform Decis Mak*, 11, p. 48, 2011.

122. M. Marschollek, M. Goevercin, K.H. Wolf, B. Song, M. Gietzelt, R. Haux, and E. Steinhagen-Thiessen. A performance comparison of accelerometry-based step detection algorithms on a large, non-laboratory sample of healthy and mobility-impaired persons. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pp. 1319–1322. IEEE, 2008.

123. A. Martínez-Ramírez, P. Lecumberri, M. Gómez, L. Rodriguez-Mañas, F. García, and M. Izquierdo. Frailty assessment based on wavelet analysis during quiet standing balance test. In *Journal of biomechanics*, 44(12), pp. 2213–2220, 2011.

124. T. Masud and R.O. Morris. Epidemiology of falls. In *Age and ageing*, 30(suppl 4), pp. 3–7, 2001.

125. M.J. Mathie, A.C. Coster, N.H. Lovell, and B.G. Celler. Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement. In *Physiological measurement*, 25(2), p. R1, 2004.

126. R.E. Mayagoitia, J.C. Lötters, P.H. Veltink, and H. Hermens. Standing balance evaluation using a triaxial accelerometer. In *Gait & posture*, 16(1), pp. 55–59, 2002.

127. J.C. Menant, J.R. Steele, H.B. Menz, B.J. Munro, S.R. Lord et al.. Optimizing footwear for older people at risk of falls. In *J Rehabil Res Dev*, 45(8), pp. 1167–1181, 2008.

128. A.E. Minetti. Biomechanics: Walking on other planets. In *Nature*, 409(6819), pp. 467–469, 2001.

129. R. Moe-Nilssen. A new method for evaluating motor control in gait under real-life environmental conditions. Part 1: The instrument. In *Clinical Biomechanics*, 13(4), pp. 320–327, 1998.

130. R. Moe-Nilssen and J.L. Helbostad. Trunk accelerometry as a measure of balance control during quiet standing. In *Gait & posture*, 16(1), pp. 60–68, 2002.

131. J. Moreland, J. Richardson, D.H. Chan, J. O'Neill, A. Bellissimo, R.M. Grum, and L. Shanks. Evidence-based guidelines for the secondary prevention of falls in older adults. In *Gerontology*, 49(2), pp. 93–116, 2003.

132. J.M. Morse, R.M. Morse, and S.J. Tylko. Development of a scale to identify the fall-prone patient. In *Canadian Journal on Aging/La Revue canadienne du vieillissement*, 8(04), pp. 366–377, 1989.

133. R.W. Motl, M. Weikert, Y. Suh, J.J. Sosnoff, J. Pula, C. Soaz, M. Schimpl, C. Lederer, and M. Daumer. Accuracy of the actibelt accelerometer for measuring walking speed in a controlled environment among persons with multiple sclerosis. In *Gait & posture*, 35(2), pp. 192–196, 2012.

134. Y.L. Murphey, H. Guo, and L.A. Feldkamp. Neural learning from unbalanced data. In *Applied Intelligence*, 21(2), pp. 117–128, 2004.

135. M.R. Narayanan, S.J. Redmond, M.E. Scalzi, S.R. Lord, B.G. Celler, and N.H. Lovell. Longitudinal falls-risk estimation using triaxial accelerometry. In *Biomedical Engineering, IEEE Transactions on*, 57(3), pp. 534–541, 2010.

136. M.R. Narayanan, S.R. Lord, M.M. Budge, B.G. Celler, and N.H. Lovell. Falls management: detection and prevention, using a waist-mounted triaxial accelerometer. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pp. 4037–4040. IEEE, 2007.

137. R. Neptune, F. Zajac, and S. Kautz. Muscle force redistributes segmental power for body progression during walking. In *Gait & posture*, 19(2), pp. 194–205, 2004.

138. J. Nilsson and A. Thorstensson. Ground reaction forces at different speeds of human walking and running. In *Acta Physiologica Scandinavica*, 136(2), pp. 217–227, 1989.

139. R.B. Northrop. Signals and systems analysis in biomedical engineering. 2004.

140. N. Noury, A. Fleury, P. Rumeau, A. Bourke, G. Laighin, V. Rialle, and J. Lundy. Fall detection-principles and methods. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pp. 1663–1666. IEEE, 2007.

141. T. Öberg, A. Karsznia, and K. Öberg. Basic gait parameters: reference data for normal subjects, 10-79 years of age. In *Journal of rehabilitation research and development*, 30, pp. 210–210, 1993.

142. D. Oliver, M. Britton, P. Seed, F. Martin, and A. Hopper. Development and evaluation of evidence based risk assessment tool (STRATIFY) to predict which elderly inpatients will fall: case-control and cohort studies. In *Bmj*, 315(7115), pp. 1049–1053, 1997.

143. J.L. O'Loughlin, Y. Robitaille, J.F. Boivin, and S. Suissa. Incidence of and risk factors for falls and injurious falls among the community-dwelling elderly. In *American journal of epidemiology*, 137(3), pp. 342–354, 1993.

144. W.H. Organization. *Preventing Chronic Diseases. A Vital Investment*. World Health Press, 2005.

145. M. O'Sullivan, C. Blake, C. Cunningham, G. Boyle, and C. Finucane. Correlation of accelerometry with clinical balance tests in older fallers and non-fallers. In *Age and ageing*, 38(3), pp. 308–313, 2009.

146. L. Palmerini, L. Rocchi, S. Mellone, F. Valzania, and L. Chiari. Feature selection for accelerometer-based posture analysis in Parkinson's disease. In *Information Technology in Biomedicine, IEEE Transactions on*, 15(3), pp. 481–490, 2011.

147. K. Pandia, S. Ravindran, R. Cole, G.T. Kovacs, and L. Giovangrandi. Motion artifact cancellation to obtain heart sounds from a single chest-worn accelerometer. In *ICASSP*, pp. 590–593. 2010.

148. N. Pannurat, S. Thiemjarus, and E. Nantajeewarawat. Automatic fall monitoring: a review. In *Sensors*, 14(7), pp. 12900–12936, 2014.

149. L. PaulE. The Effect of Ankle Foot Orthoses on Balance. Center of Pressure. In *Online Learning Center. American Academy of Orthotists and Prosthetists*. URL `http://www.oandp.org/olc/lessons/html/SSC_10`.

150. K.L. Perell, A. Nelson, R.L. Goldman, S.L. Luther, N. Prieto-Lewis, and L.Z. Rubenstein. Fall risk assessment measures an analytic review. In *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 56(12), pp. M761–M766, 2001.

151. J. Perry, J.R. Davids et al.. Gait analysis: normal and pathological function. In *Journal of Pediatric Orthopaedics*, 12(6), p. 815, 1992.

152. J. Perttunen. *Foot loading in normal and pathological walking*. University of Jyväskylä, 2002.

153. L.E. Peterson. K-nearest neighbor. In *Scholarpedia*, 4(2), p. 1883, 2009.

154. G.M. Powell and E. Dzendolet. Power spectral density analysis of lateral human standing sway. In *Journal of motor behavior*, 16(4), pp. 424–441, 1984.

155. W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. Numerical recipes in C: the art of scientific computing, 1992. In *Cité en*, p. 92, 1992.

156. T.E. Prieto, J. Myklebust, R. Hoffmann, E. Lovett, and B. Myklebust. Measures of postural steadiness: differences between healthy young and elderly adults. In *Biomedical Engineering, IEEE Transactions on*, 43(9), pp. 956–966, 1996.

157. M. Raîche, R. Hébert, F. Prince, and H. Corriveau. Screening older adults at risk of falling with the Tinetti balance scale. In *The Lancet*, 356(9234), pp. 1001–1002, 2000.

158. J. Raymakers, M. Samson, and H. Verhaar. The assessment of body sway and the choice of the stability parameter (s). In *Gait & posture*, 21(1), pp. 48–58, 2005.

159. G. Rescio, A. Leone, and P. Siciliano. Support Vector Machine for tri-axial accelerometer-based fall detector. In *Advances in Sensors and Interfaces (IWASI), 2013 5th IEEE International Workshop on*, pp. 25–30. IEEE, 2013.

160. A.M. Rose. The subculture of the aging: A topic for sociological research. In *The Gerontologist*, 2(3), pp. 123–127, 1962.

161. L.Z. Rubenstein. Falls in older people: epidemiology, risk factors and strategies for prevention. In *Age and ageing*, 35(suppl 2), pp. ii37–ii41, 2006.

162. A. Savitzky and M.J. Golay. Smoothing and differentiation of data by simplified least squares procedures. In *Analytical chemistry*, 36(8), pp. 1627–1639, 1964.

163. R.W. Schafer. What is a Savitzky-Golay filter?[lecture notes]. In *Signal Processing Magazine, IEEE*, 28(4), pp. 111–117, 2011.

164. A.C. Scheffer, M.J. Schuurmans, N. van Dijk, T. van Der Hooft, and S.E. De Rooij. Fear of falling: measurement strategy, prevalence, risk factors and consequences among older persons. In *Age and Ageing*, 37(1), pp. 19–24, 2008.

165. M. Schimpl, C. Moore, C. Lederer, A. Neuhaus, J. Sambrook, J. Danesh, W. Ouwehand, and M. Daumer. Association between Walking Speed and Age in Healthy, Free-Living Individuals Using Mobile Accelerometry—A Cross-Sectional Study. In *PLoS One*, 6(8), p. e23299, 2011.

166. V. Scott, K. Votova, A. Scanlan, and J. Close. Multifactorial and functional mobility assessment tools for fall risk among older adults in community, home-support, long-term and acute care settings. In *Age and ageing*, 36(2), pp. 130–139, 2007.

167. F.E. Shaw, J. Bond, D.A. Richardson, P. Dawson, I.N. Steen, I.G. McKeith, and R.A. Kenny. Multifactorial intervention after a fall in older people with cognitive impairment and dementia presenting to the accident and emergency department: randomised controlled trial. In *Bmj*, 326(7380), p. 73, 2003.

168. S. Shultz, P. Houglum, and D. Perrin. *Examination of Musculoskeletal Injuries with Web Resource*. Human Kinetics Publishers, 2009.

169. A.G. Society. Guideline for the prevention of falls in older persons. In *Journal of the American Geriatrics Society*, 49(5), pp. 664–672, 2001.

170. T.A. Soriano, L.V. DeCherrie, and D.C. Thomas. Falls in the community-dwelling older adult: a review for primary-care providers. In *Clinical interventions in aging*, 2(4), p. 545, 2007.

171. P. Stalenhoef, J. Diederiks, J. Knottnerus, A. Kester, and H. Crebolder. A risk model for the prediction of recurrent falls in community-dwelling elderly: a prospective cohort study. In *Journal of clinical epidemiology*, 55(11), pp. 1088–1094, 2002.

172. S. Steidl, C. Schneider, and M. Hufnagl. Fall detection by recognizing patterns in direction changes of constraining forces. In *Proceedings of the eHealth2012*, pp. 10–11, 2012.

173. I. Steinwart and A. Christmann. *Support vector machines*. Springer, 2008.

174. R.C. Team et al.. R: A language and environment for statistical computing. In *R foundation for Statistical Computing*, 2005.

175. P.B. Thapa, P. Gideon, K.G. Brockman, R.L. Fought, and W.A. Ray. Clinical and biomechanical measures of balance fall predictors in ambulatory nursing home residents. In *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 51(5), pp. M239–M246, 1996.

176. R. Thiel, K. Stroetmann, V.N. Stroetmann, and M. Viceconti. Designing a socio-economic assessment method for integrative biomedical research: the Osteoporotic Virtual Physiological Human project. In *MIE*, pp. 876–880. 2009.

177. R. Tibshirani. Regression shrinkage and selection via the lasso. In *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.

178. M.E. Tinetti, T. Franklin Williams, and R. Mayewski. Fall risk index for elderly patients based on number of chronic disabilities. In *The American journal of medicine*, 80(3), pp. 429–434, 1986.

179. M.E. Tinetti, W.L. Liu, and E.B. Claus. Predictors and prognosis of inability to get up after falls among elderly persons. In *JAMA: the journal of the American Medical Association*, 269(1), pp. 65–70, 1993.

180. M. Tinetti, M. Speechley, and S. Ginter. Risk factors for falls among elderly persons living in the community. In *New England journal of medicine*, 319(26), pp. 1701–1707, 1988.

181. C. Todd and D. Skelton. *What are the main risk factors for falls amongst older people and what are the most effective interventions to prevent these falls?* World Health Organization, 2004.

182. E. Todd M. Health care providers are the first line of defense against falls.@ONLINE. 2012. URL `http://www.mayoclinic.org/medicalprofs/fall-prevention-elderly-tue0113.html`.

183. A. Turcato and S. Ramat. Predicting losses of balance during upright stance: evaluation of a novel approach based on wearable accelerometers. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pp. 4918–4921. IEEE, 2010.

184. K. Umemura, K. Tachikawa, Y. Kurihara, and K. Watanabe. Analysis of human walking by 3-D acceleration sensor and its application. In *SICE Annual Conference, 2008*, pp. 1723–1726. IEEE, 2008.

185. E.J. Van Someren. Actigraphic monitoring of movement and rest-activity rhythms in aging, Alzheimer's disease, and Parkinson's disease. In *Rehabilitation Engineering, IEEE Transactions on*, 5(4), pp. 394–398, 1997.

186. P.H. Veltink, H.J. Bussmann, W. De Vries, W.L. Martens, and R.C. Van Lummel. Detection of static and dynamic activities using uniaxial accelerometers. In *Rehabilitation Engineering, IEEE Transactions on*, 4(4), pp. 375–385, 1996.

187. S. Volpato, M. Cavalieri, F. Sioulis, G. Guerra, C. Maraldi, G. Zuliani, R. Fellin, and J.M. Guralnik. Predictive value of the Short Physical Performance Battery following hospitalization in older patients. In *J. Gerontol. A Biol. Sci. Med. Sci.*, 66(1), pp. 89–96, Jan 2011.

188. S. Volpato, S.G. Leveille, C. Blaum, L.P. Fried, and J.M. Guralnik. Risk factors for falls in older disabled women with diabetes: the women's health and aging study. In *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 60(12), pp. 1539–1545, 2005.

189. B. Weijters and M. Geuens. Evaluation of age-related labels by senior citizens. In *Psychology & Marketing*, 23(9), pp. 783–798, 2006.

190. A. Weiss, T. Herman, M. Plotnik, M. Brozgol, N. Giladi, and J. Hausdorff. An instrumented timed up and go: the added value of an accelerometer for identifying fall risk in idiopathic fallers. In *Physiological measurement*, 32(12), p. 2003, 2011.

191. A. Weiss, M. Brozgol, M. Dorfman, T. Herman, S. Shema, N. Giladi, and J.M. Hausdorff. Does the evaluation of gait quality during daily life provide insight into fall risk? A novel approach using 3-day accelerometer recordings. In *Neurorehabilitation and neural repair*, 27(8), pp. 742–752, 2013.

192. G.J. Welk et al.. Principles of design and analyses for the calibration of accelerometry-based activity monitors. In *Medicine and Science in Sports and Exercise*, 37(11), p. S501, 2005.

193. S. Whitney, M. Hudak, and G. Marchetti. The dynamic gait index relates to self-reported fall history in individuals with vestibular dysfunction. In *Journal of Vestibular Research*, 10(2), pp. 99–105, 2000.

194. R. Wicklin. In Computing prediction ellipses from a covariance matrix. URL `http://saslist.com`.

195. J.M. Wilken, K.M. Rodriguez, M. Brawner, and B.J. Darter. Reliability and minimal detectible change values for gait kinematics and kinetics in healthy adults. In *Gait & posture*, 35(2), pp. 301–307, 2012.

196. D. Winter. *A.B.C. (anatomy, Biomechanics and Control) of Balance During Standing and Walking*. Waterloo Biomechanics, 1995. ISBN 9780969942009. URL `http://books.google.de/books?id=OlSqQgAACAAJ`.

197. D.A. Winter. Human balance and posture control during standing and walking. In *Gait & Posture*, 3(4), pp. 193–214, 1995.

198. D.A. Winter, A.E. Patla, J.S. Frank, and S.E. Walt. Biomechanical walking pattern changes in the fit and healthy elderly. In *Physical therapy*, 70(6), pp. 340–347, 1990.

199. D.A. Winter, A.E. Patla, F. Prince, M. Ishac, and K. Gielo-Perczak. Stiffness control of balance in quiet standing. In *Journal of Neurophysiology*, 80(3), pp. 1211–1221, 1998.

200. L.I. Wolfson, R. Whipple, P. Amerman, J. Kaplan, and A. Kleinberg. Gait and balance in the elderly. Two functional capacities that link sensory and motor ability to falls. In *Clinics in geriatric medicine*, 1(3), pp. 649–659, 1985.

201. C.C. Yang and Y.L. Hsu. A review of accelerometry-based wearable motion detectors for physical activity monitoring. In *Sensors*, 10(8), pp. 7772–7788, 2010.

202. W. Youden. Index for rating diagnostic tests. In *Cancer*, 3(1), pp. 32–35, 1950.

203. A.A. Zecevic, A.W. Salmoni, M. Speechley, and A.A. Vandervoort. Defining a fall and reasons for falling: comparisons among the views of seniors, health care providers, and the research literature. In *The Gerontologist*, 46(3), pp. 367–376, 2006.

204. Y. Zigel, D. Litvak, and I. Gannot. A method for automatic fall detection of elderly people using floor vibrations and sound-Proof of concept on human mimicking doll falls. In *Biomedical Engineering, IEEE Transactions on*, 56(12), pp. 2858–2867, 2009.

205. A. Zijlstra, E.D. de Bruin, N. Bruins, and W. Zijlstra. The step length–frequency relationship in physically active community-dwelling older women. In *European journal of applied physiology*, 104(3), pp. 427–434, 2008.

206. W. Zijlstra and A.L. Hof. Displacement of the pelvis during human walking: experimental data and model predictions. In *Gait & posture*, 6(3), pp. 249–262, 1997.

207. W. Zijlstra and A.L. Hof. Assessment of spatio-temporal gait parameters from trunk accelerations during human walking. In *Gait & posture*, 18(2), pp. 1–10, 2003.

208. H. Zou and T. Hastie. Regularization and variable selection via the elastic net. In *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2), pp. 301–320, 2005.