Heel2Toe: A Biofeedback Device to Assist Training of Heel-to-toe Gait in the Rehabilitation of the Elderly

Abhishek Vadnerkar

Master of Engineering

Biomedical Engineering Department

McGill University

Montreal, Quebec

April 14, 2015

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Engineering

ACKNOWLEDGEMENTS

I would like to express my appreciation to my supervisor, Prof. Robert Kearney, for giving me the opportunity to work under his supervision and giving me significant guidance and feedback over the course of this research. Without his persistant help and support this thesis would not have been possible.

I would also like to acknowledge Prof. Nancy Mayo with the School of Physical & Occupational Therapy, who proposed this project and gave feedback and guidance from a physiotherapist perspective throughout the project. Additionally, I would like to thank Sabrina Figueiredo, also with the School of Physical & Occupational Therapy, for assigning the physiotherapist ratings, recruiting patients, conducting the experiments with me, and giving me feedback over the course of the project.

I would like to thank all the REKLAB members, Kian, Chris, Mahsa, Diego, Pouya, Ehsan, Hamed, and Ross for allowing me to bounce ideas off them and assisting me with my research. Nancy Abate and Pina Sorrini with the Biomedical Engineering Department were very helpful and made the process a lot easier. I would also like to thank Claire Murray for translating the abstract for me.

Finally I would like to express my gratitude to my partner Samantha, and my family, Varsha, Sunil, and Arohi, whose care and support made this all possible.

This work was supported by funding provided by the Natural Science and Engineering Research Council of Canada (NSERC) and the Helen McCall Hutchinson Award.

ABSTRACT

A feature of healthy gait is a clearly defined heel strike in which the foot makes contact with the ground, and then rotates towards forefoot contact, known as heel-to-toe gait. However, a common consequence of ageing is the deterioration of heel first gait towards a shuffling or toe first gait. Physiotherapy can be useful for correcting this abnormality, however, therapist time is limited and training is limited to a clinical environment. A hypothesis is that gait rehabilitation could be expedited with the use of a device that distinguishes between heel first and shuffling gait and provides feedback to the user. This thesis presents the development and validation of this device.

The device comprises a sensing module and a biofeedback module. The sensing module is a six axis inertial measurement unit that is strapped to the patient's foot. The biofeedback module is a smartphone. Raw data is streamed wirelessly to the smartphone which runs an algorithm that detects step quality on the basis of angular velocity of the foot, and provides binary feedback on step quality to the user. Results from a validation study on the target population demonstrate that the algorithm performs very well, with an overall accuracy of 85.8% when compared with physiotherapist labels. A design choice was made to maximise sensitivity, with a minimum acceptable specificity of 75%. The sensitivity is 96.2% at an operating point of 75% specificity, and the area under the ROC curve is 0.956. With the classification ability of the device established, future work will assess the effectiveness of the device at improving gait.

ABRÉGÉ

Une caractéristique d'une démarche saine est une attaque du talon dans laquelle le pied rentre en contact avec le sol, et ensuite se tourne dans le sens de l'avant du pied. Cette démarche est aussi appelée démarche talon-pointe. Cependant, une conséquence commune du vieillissement est la détérioration de la démarche talon-avant vers une démarche traînante ou orteils-avant. La physiothérapie peut être utile pour corriger cette anomalie, en revanche, le thérapeute est limité en temps et à l'environnement clinique. Une hypothèse de cette thèse est que la réadaptation de la démarche pourrait être prise en charge par l'utilisation d'un appareil qui distinguerait la démarche talon-avant de la démarche traînante et donnerait le résultat de ces analyses à l'utilisateur. Cette thèse présente le développement et la validation d'un tel appareil.

L'appareil se compose d'un module de mesure et d'un module de rétroaction biologique. Le module de mesure est une unité de mesure inertielle à 6 axes qui est attachée au pied du patient. Le module de rétroaction biologique est un téléphone intelligent. Le flux des données brutes est envoyé par une technique sans fil au téléphone sur lequel tourne un algorithme détectant la qualité du pas, en se basant sur la vitesse angulaire du pied. Le téléphone envoie alors son retour binaire sur la qualité du pas à l'utilisateur. Les résultats obtenus grâce à une étude de validation sur population ciblée démontrent que l'algorithme implémenté fonctionne avec une très bonne précision globale de 85.8% par rapport à la classification effectuée par le physiothérapeute. La conception de l'appareil a été choisie afin de maximiser la

sensibilité, avec un minimum acceptable de 75%. La sensibilité est de 96.2% à un point de fonctionnement de spécificité de 75%, et l'aire sous la courbe ROC (fonction d'efficacité du récepteur) est de 0.956. Avec l'aptitude de classification de l'appareil établie, de prochains travaux permettront d'établir son efficacité pour améliorer la démarche.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
Introduction

The ability to walk safely, quickly, and efficiently is a key aspect of daily life that contributes greatly to quality of life. While the primary function of walking is transportation, there are numerous other health benefits including increasing cardiovascular fitness, reducing the risk of heart disease, and increasing bone density. Walking is such an intrinsic feature of life that it is often taken for granted in those with healthy gait.

Gait is highly variable from person to person, but a common feature of healthy gait is a clearly defined heel contact when the foot first touches the ground. The moment when the foot touches the ground is known as initial contact. Normally, it is the heel that touches the ground first, and then the toes rotate downwards leading to forefoot contact. This pattern is known as heel-to-toe gait. A common consequence of ageing is the loss of heel-to-gait in favour of a flat-foot initial contact or 'shuffling' gait. This results in a shortened stride length, slower gait speed, and an increased risk of falling. Individuals with a shuffling gait are often unable to walk for long periods of time, so the health benefits associated with walking are diminished.

Shuffling gait may be treated by physical therapy, during which, the physiotherapist works with the patient to increase muscle strength through repetitive training, that focuses on restoring heel-to-toe gait. Usually the therapist will teach the patient the importance of having a clearly defined heel strike by walking with the patient and providing verbal feedback of ‘good’ and ‘bad’ steps. While physiotherapy can be effective, it is costly, labour intensive, and limited to a therapist’s clinic. Gait rehabilitation can be accelerated with home exercise, but patients often walk with incorrect technique away from therapist supervision. With this in mind, there may be a benefit in developing a way to ‘supervise’ gait training in an unsupervised environment, e.g. in the patient’s home.

This provides the rationale for this project: the development of a wearable device that can detect individual steps, and distinguish between a heel-to-toe step (good) and a shuffling step (bad). This device will provide feedback about the step quality in real time. This will allow the user to correct incorrect gait technique while exercising at home. The device should be cheap, unobtrusive, and portable. The goal for this device is to supplement physical therapy to accelerate rehabilitation.

1.1 Thesis outline

This thesis describes the development and validation of this device. Chapter 2 presents the relevant background material, including a description of healthy gait, causes and effects of its degeneration, and methods of rehabilitation. Prior use of wearable devices for gait applications is also described.

Chapter 3 (Methods) describes both the hardware and software used for data acquisition. The device comprises a sensing module and a user interface. The sensing module is an inertial measurement unit attached to the foot, which streams raw data wirelessly to a PC running the user interface. The PC runs an algorithm to detect and classify each step, while giving feedback about the step quality to the user in real time. Chapter 3 also presents the classification methods used to distinguish between good and bad steps (feature extraction, classifier type, and various performance metrics use to analyse the classification result). Finally the chapter explains the preliminary experimental procedure.

Chapter 4 describes the development of the algorithm to distinguish between good and bad steps. The physiotherapist rating procedure is described, as well as the classification performance results after applying the algorithm. An analysis of the calibration procedure for the sensing module is also included.

Chapter 5 presents a pilot clinical study to validate the algorithm, including a description of the methods used, the clinical experiments, the main findings, and an analysis of the algorithm's accuracy.

Finally, Chapter 6 discusses the entire project, focusing on potential improvements to the classifier, and future work.

CHAPTER 2
Background

The aim of this project is to develop a biofeedback device that distinguishes between good and bad steps for the purposes of gait rehabilitation in the elderly. This chapter reviews the relevant literature. First we explore healthy gait, focusing on the quantification of healthy gait, including its kinematics and spatio-temporal parameters. We then examine abnormal gait, first looking at causes, and then how these causes manifest in terms of the kinematics and spatio-temporal parameters. This is followed by a review of gait rehabilitation. Finally we look at prior use of wearable devices for gait, splitting the literature into sensing applications and feedback applications, since this project involves a combination of sensing and feedback.

2.1 Gait overview

Locomotion is the process of transporting one's body from one place to another. A subset of locomotion, and by far the most common form in humans, is gait, which includes both walking and running. The purpose of gait is to allow efficient and safe transport across a variety of terrains. Winter [1] proposed that five functions were critical to human gait:

1. Provision of support of the upper body
2. Maintenance of upright posture and balance
3. Control of foot swing to achieve adequate ground clearance and allow a gentle foot landing

4. Provision of forward propulsion
5. Provision of sufficient shock absorption throughout the gait cycle

Gait is primarily a learned activity [2] with large variability from individual to individual. Indeed, the kinematic variability of gait is large enough that an individual can usually be recognised from their particular gait pattern alone [1]. Despite this variation, there are aspects of the gait cycle that are common to most of the population that can be used to subdivide the gait cycle into a standard set of phases [3]. Figure 2–1 shows one commonly used set of phases of the gait cycle for one leg. Starting with initial heel contact the important events include:

- Stance phase: the entire period when the foot is contact with the ground.
  - Initial contact: The instant the heel first hits the ground, represented by 0% cycle in Figure 2–1.
  - Loading response (or weight acceptance): The period from initial contact to maximum knee flexion, during which the energy from the impact of landing is absorbed, predominantly by the knee joint.
  - Mid stance: Lasts until the beginning of the push off phase.
  - Push off: Starts with initial plantarflexion of the ankle and ends with toe off. Generates forward momentum for the swing phase.
  - Toe off (or lift off): The instant the foot leaves the ground
- Swing phase: the entire period when the foot is moving through the air.
  - Pre-swing: Begins at toe off. Weight is rapidly transferred to the contralateral limb
  - Initial swing: After toe off, advancement of the thigh commences.

– Late-swing (reach): Prior to initial contact, characterized by knee extension.

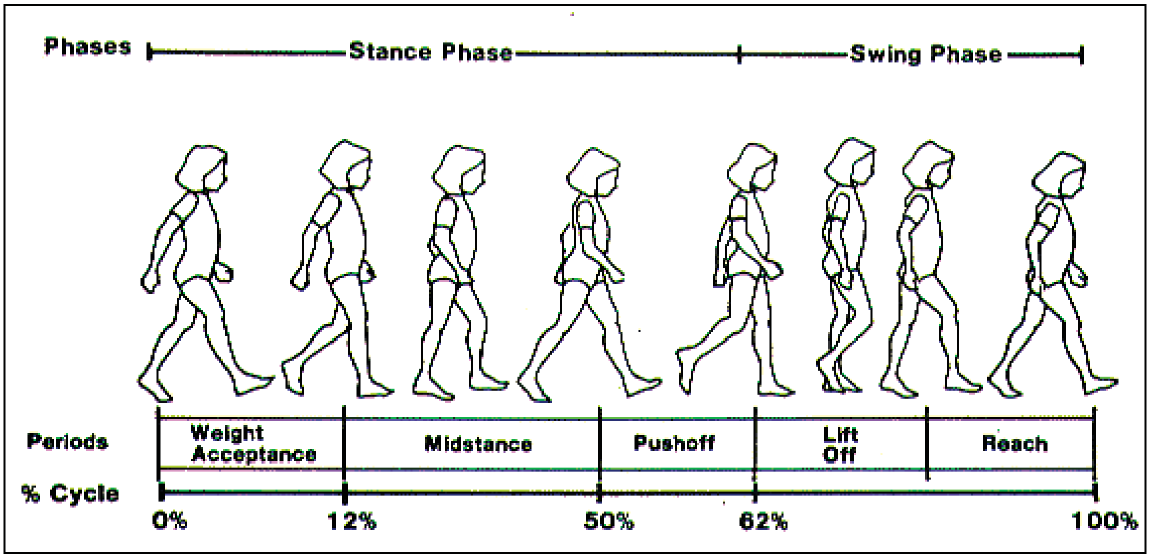

Figure 2–1: Graphical representation of gait phases, as described by Winter [4]. Image taken from Kirtley's Clinical Gait Analysis [5]

2.2 Kinematics of healthy gait

Quantitative measurements of gait can be divided into three types, kinetics, kinematics, and electromyography. Kinetic analysis involves the measurement and estimation of forces, joint moments and energy. Kinematics involves the measurement of position (linear and angular) of joints and segments. Electromyography (EMG) uses the electrical activity of muscles as an indication of their level of activation. This review will focus on the kinematics of gait since it is most relevant to this application, since the distinction between heel-to-toe and shuffling gait is primarily kinematic in nature.

Sensing systems have been used extensively to describe the kinematics of gait. While humans have demonstrated a keen interest in dissecting the patterns of gait for hundreds of years [6], it was not until the 1870s that a full kinematic study was possible due to the advent of the still camera. Photography as a method of gait deconstruction was pioneered by Muybridge and Marey [7] amongst others. This method was advanced further with the development of video recording equipment and major studies on gait kinematics were conducted in the 1940s and 50s (by [2]), with limb positions and angles digitized manually from film. Integrated motion capture systems in the 1970s and 80s allowed much faster and more convenient video analysis [6], which spawned many studies of gait kinematics (including Aptekar et al. [8] and Grieve [9]).

Winter [1] used a motion capture system to quantify gait cycle kinematics for normal, elderly and pathological gait. Measurements included the absolute position of 15 body segments including the feet, legs, thighs, upper arms, forearms, hands, head, trunk, and pelvis. Each segment required 6 variables to describe its linear motion in the plane of progression (i.e. position, velocity and acceleration in the vertical and horizontal directions), and 3 rotational variables (i.e. angle, angular velocity, and angular acceleration). This resulted in a total of 135 gait variables. Some of Winter's results that are relevant to our study are presented next.

Figures 2–2, 2–3, and 2–4 show the linear displacement and velocity for the average gait cycle of 14 healthy subjects. Initial contact is at 0% stride. Positive displacements and velocities are upward and forward. Joint angles are positive for hip flexion, knee flexion, and ankle dorsiflexion. Important features to note include:

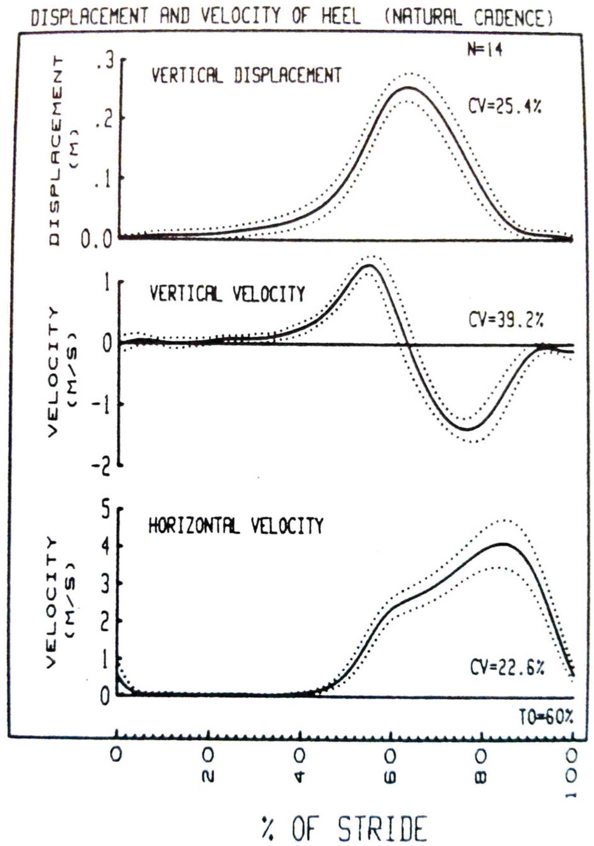

Figure 3.11(b)

Figure 2–2: Displacement and velocity of heel for 14 natural cadence subjects, showing the mean as the solid line and 1 standard deviation from the mean as the dashed lines. Figure taken from Winter [1], pg 20.

- Most importantly for this research is the position of the foot at initial contact (0% stride). Figure 2–2 shows that there is zero vertical heel displacement at initial contact, indicating that the heel makes the first contact. There is significant vertical displacement of the toes at initial contact (Figure 2–4) which then falls to zero. The vertical velocity of the toes is negative at 0%, indicating

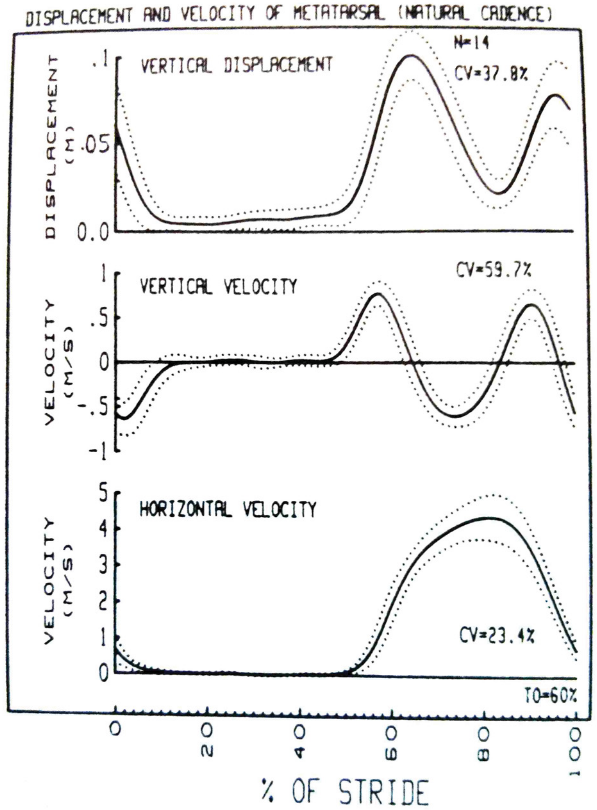

Figure 3.11(c)

Figure 2–3: Displacement and velocity of metatarsal for 14 natural cadence subjects, showing the mean as the solid line and 1 standard deviation from the mean as the dashed lines. Figure taken from Winter [1], pg 20.

rotation downwards towards the ground at heel contact. Thus normal gait is characterized by a *heel-to-toe* transition.

- Figure 2–3 shows that the horizontal heel velocity falls rapidly to near zero at initial contact, mitigating the risk of slip related falls by reducing the need for friction to decelerate the foot.

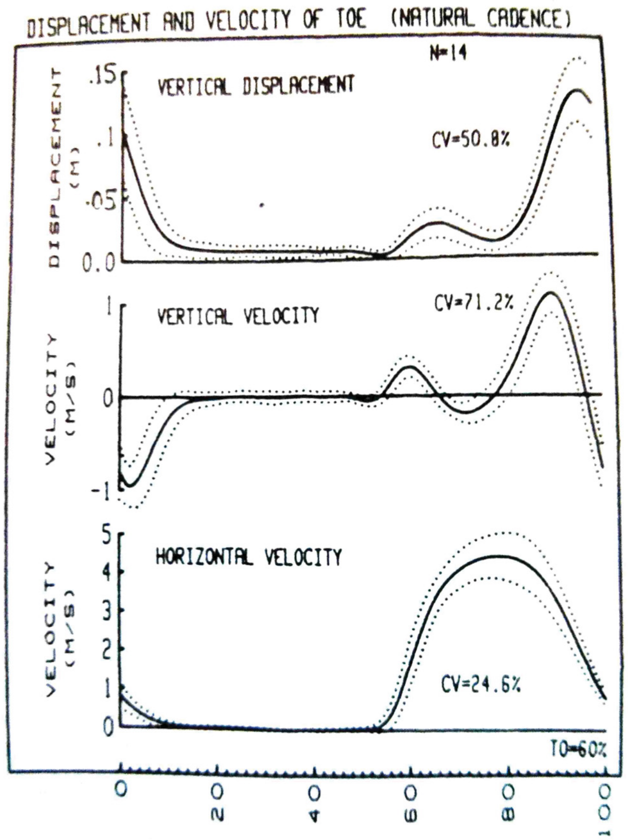

Figure 3.11(d)

Figure 2–4: Displacement and velocity of toe for 14 natural cadence subjects, showing the mean as the solid line and 1 standard deviation from the mean as the dashed lines. Figure taken from Winter [1], pg 21.

- Figure 2–4 shows that the toe clearance (vertical displacement) during the swing phase (55% to 100% of stride) is at a local minimum of 1.3cm the same time as when the horizontal velocity is maximum. This highlights the fine control required to avoid trips; the foot clearance is dependent on hip, knee, and ankle movement so controlling the swinging foot to achieve such a small clearance while avoiding tripping is demanding. Section 2.3 will show that the toe clearance changes with degenerative gait.

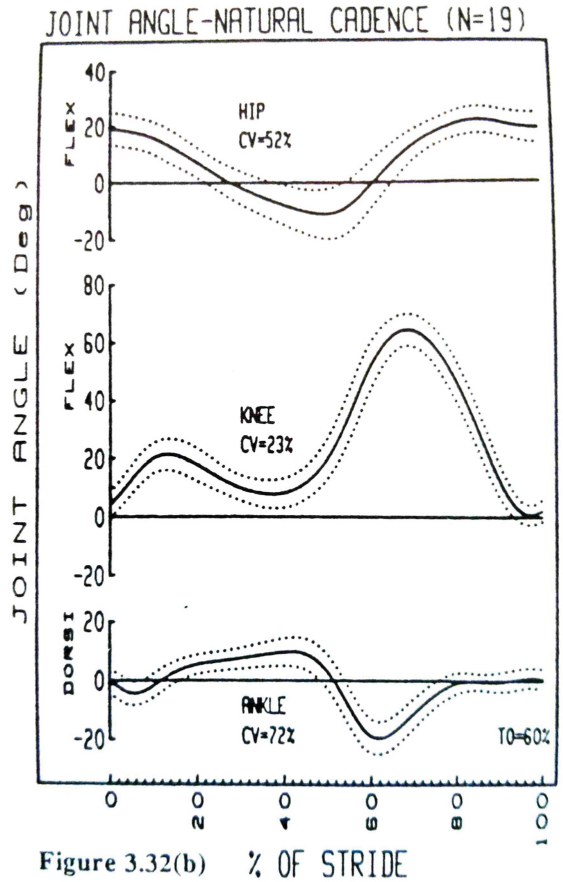

Figure 2–5: Joint angles for 19 subjects walking at a natural cadence. Dashed line shows inter-subject variation. Figure taken from Winter [1], pg 28.

Figure 2–5 shows average hip, knee, and ankle angles in the plane of progression through the gait cycle. The ankle is at the neutral position at heel contact, plantarflexes slightly as the forefoot makes contact with the ground, and then dorsiflexes as the shank rotates forward over the foot. Prior to toe off, the ankle begins to plantarflex to generate power for the stride phase. Following toe off, the ankle dorsiflexes to prepare for initial contact. Winter showed that this profile is roughly independent of walking speed [10]. Inter-subject amplitude variation is high, with coefficient of variations of 52%, 23%, and 72% for the hip, knee, and ankle respectively, which explains the individual recognisability of a person's gait.

2.2.1 Spatio-temporal parameters

A common approach to quantifying gait (especially when assessing the effects of treatments) is to measure spatio-temporal gait parameters, such as walking speed, stride length, and cadence. The advantage of these measures (as opposed to kinematics/kinetics) is that they are easy and inexpensive to measure and have practical significance. A study by Schmid et al. [11] found that outcome assessment schemes rooted in gait velocity produced "meaningful indicators of clinical benefit".

Stride length is defined as the distance travelled between consecutive stance periods of the same foot; cadence is the step rate, or number of steps of one foot per minute. Walking speed is the product of stride length and cadence. Cadence can be calculated simply with a stopwatch and counter. It can also be extracted easily from most measurement systems (instrumented walkways, inertial units, video recording devices etc.). Stride length is easily extracted from instrumented walkways and camera systems.

Average spatio-temporal gait parameters for various groups have been reported. The average cadence for adults ranges from 101 - 122 steps per minute [1]. Winter [12] found that while previous studies had reported a reduced cadence for elderly

[13, 14], when elderly subjects were screened for fitness and gait abnormalities, there were no significant differences in cadence between elderly and non-elderly adults. A corollary to this finding is that elderly subjects with gait abnormalities have a lower cadence than healthy subjects.

A study by Lamoreux [15] found that cadence and stride length were interdependent over the natural walking speed range. From 80 steps per minute to 120, stride length varied linearly with cadence, after which it plateaued. Thus changes in gait speed result from increases in both stride length and cadence [16]. This correlation between cadence and stride length can be used to estimate stride length (and hence gait speed) from the easily measured cadence.

2.3 Gait degeneration and causes

Sections 2.2 and 2.1 discussed the segment kinematics and spatio-temporal parameters of healthy gait. While these profiles are applicable to the majority of the population, there are a number of pathologies that change these profiles, and have a negative impact on the gait pattern and quality of life.

Gait deterioration is commonly associated with ageing. One study reported that 35 percent of older adults have abnormal gait [17]. Specifically, gait disorders were found in 25% of adults aged 70 to 74, and in 60% of adults over 84. However, while abnormal gait is common in older individuals, it is not inevitable since many elderly individuals have healthy gait [18].

The origins of some gait pathologies are well understood, but often the causes are multifactorial. Hough et al. [19] found that 75% of gait abnormalities in elderly patients stemmed from multiple sources. In a study [20] of 153 elderly patients, 40%

of those reporting gait abnormalities cited ‘more than one cause’. The most common self-reported causes of gait impairment were joint pain, stroke, visual impairment, and back or neck pain.

This multifactorial nature of gait disorders can make it difficult to diagnose gait disorders. Gait disorders often present as a ‘slow, broad-based, shuffling, and cautious walking pattern’ [18]. Traditionally, this has been given the name ‘senile gait disorder’ or ‘idiopathic senile gait’. However, this diagnosis often masks the underlying cause of gait which can lead to slower rehabilitation times, since the optimal treatment depends on the root cause [18].

Table 2–1 summarizes the characteristics and causes of some common gait disorders based on the work of Alexander and Goldberg [21].

In the study presented in Section 2.2, Winter also [1] examined the kinematics of gait in healthy elderly individuals. Figures 2–6, 2–7, and 2–8 summarize the kinematics observed in this group. Although these subjects were assessed as being healthy (no gait disorders), there were some key differences between elderly (mean age 68.9 years) and young (mean age 24.9 years).

- Elderly subjects had a significantly ($p<0.01$) lower stride length than young adults. However cadence was not significantly different.
- Heel velocity at initial contact was higher for the elderly despite a lower gait velocity, resulting in a greater slip-induced fall risk (Figure 2–6).
- Elderly subjects had less ankle dorsiflexion at initial contact, indicating that heel strike was less pronounced (Figure 2–8).

Table 2–1: Common causes of gait disorders

Type	Characteristics	Cause
Sensory ataxia	Unsteady, tentative and uncoordinated gait.	Sensory deficits, especially to the visual and vestibular systems.
Peripheral motor deficit	Arthritic - Avoidance of weight bearing on affected side (asymmetric gait), excess flexion of the affected knee. Slower and shorter steps. Myopathic - Excess trunk flexion and a 'waddling' gait. 'Foot drop' (uncontrolled heel strike to foot flat phase) is also present.	Pain while walking. Muscle weakness, particularly of dorsiflexors.
Spasticity	Circumduction gait (one leg for hemiplegia, both for paraplegia). Excessive plantarflexion may be present too.	Vascular (cerebral haemorrhage, stroke), traumatic, or congenital (cerebral palsy) origins, resulting in muscle tightness.
Parkinsonism	Small, fast, shuffling steps, with postural instability (forward and backward leaning).	Degeneration of the central nervous system.
Cerebellar ataxia	Unsteady gait, with increased trunk sway. Steps are often staggered, and with a wide base.	Brain injury, substance abuse, and/or genetic disease.
Fear of falling	Widened gait base, with a shorter stride length and slower cadence.	Usually follows from prior falls, adaption towards steadier gait.

2.4 Rehabilitation

The gait changes observed in Winter's study [1] were minor, and would not normally warrant intervention. However, many individuals with gait disorders have more significant changes in gait kinematics. Since each disorder has different gait characteristics, an 'average' gait profile is not informative. However, many gait disorders lack a clearly defined heel strike [21]. This results in a shortened stride

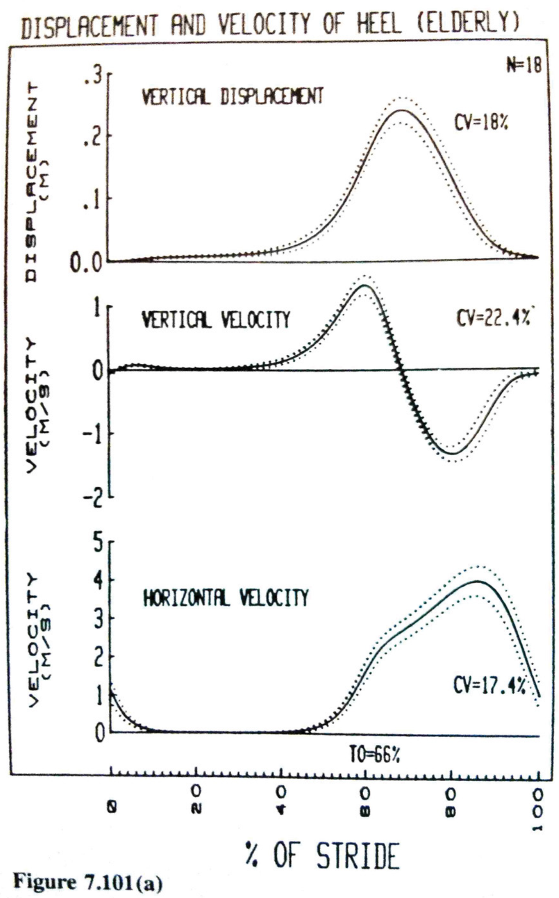

Figure 2–6: Displacement and velocity of heel for healthy elderly subjects. Figure taken from Winter [1], pg 90.

length, and slower walking speed [22]. It also inhibits use of the dorsiflexors, which leads to weakness of these muscles, and hence an increased risk of falling, since the dorsiflexors are critical for stability [23]. In addition, loss of heel strike can lead to cardiovascular deconditioning, since affected individuals are unable to walk at a moderate-to-vigorous intensity for the recommended period of time (150 minutes per week) needed to maintain cardiovascular fitness [24].

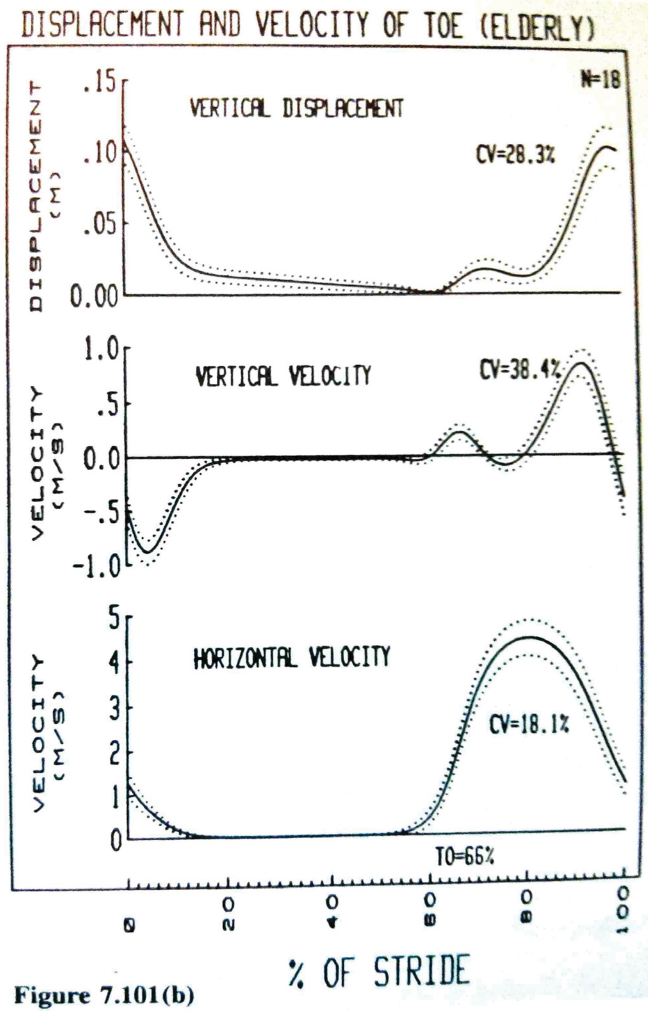

Figure 7.101(b)

Figure 2–7: Displacement and velocity of toe for healthy elderly subjects. Figure taken from Winter [1], pg 91.

Rehabilitation programs aim to increase the tolerance to walking, increase gait speed, and improve balance. These objectives are accomplished by a variety of methods including gait training, walking training, muscle strengthening, stretching, and proprioceptive training [25]. It is important to distinguish between gait training

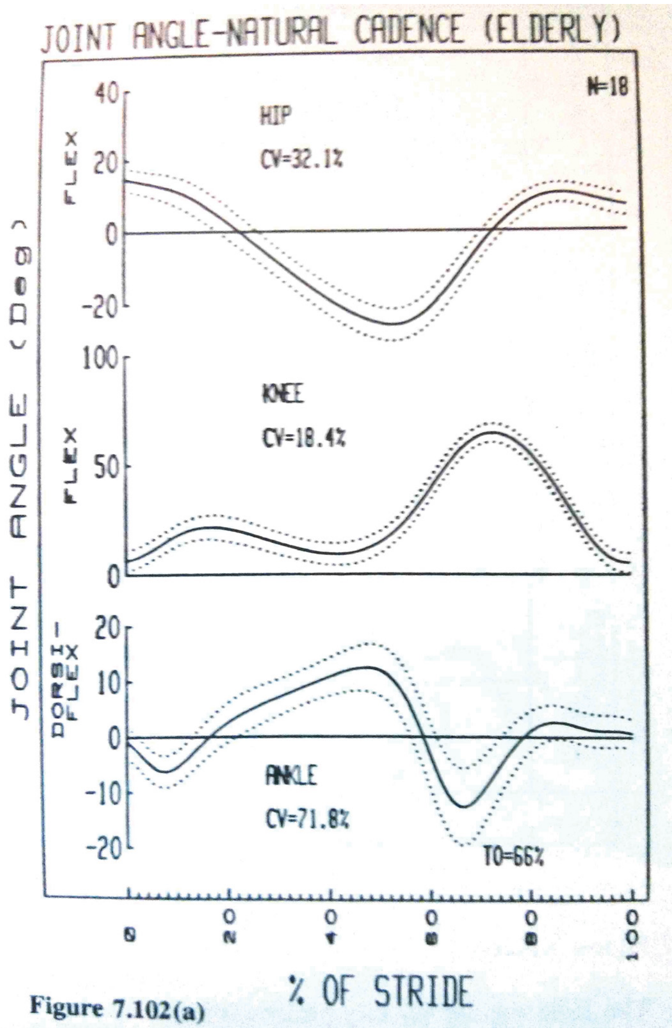

Figure 2–8: Joint angles for healthy elderly subjects. Figure taken from Winter [1], pg 91.

and walking training. Gait training uses repetitive activities to improve walking technique through practice [26]. Due to the space limitations common in a clinical setting, patients are usually asked to walk back and forth in a corridor for a pre-determined period of time [25]. The importance of heel-to-toe gait is stressed, since this helps improve trunk posture, stride length and stability. External cues are often used to improve outcomes, these include visual cues (e.g. marked target lines for steps), and auditory cues (e.g. metronome) for cadence [27]. Other strategies for gait training include encouraging the visualisation of proper gait, segmentation of the gait cycle (training each phase separately), and reciting phrases that promote heel-to-toe gait [26]. The effector of heel-to-toe training has not been explored objectively.

Walking training, on the other hand, focuses on quantity, with an attempt to walk for longer distances and durations. Since duration is the focus, training often takes place outside the clinic. For example, mall walking has been proven popular amongst the elderly [28]. The effectiveness of walking training has been documented in several studies that used spatio-temporal parameters as the study outcomes due to their ease of measurement. Studies by Salbach et al. [29], Moffet et al. [30], and Mangione et al. [31] all reported significant changes in walking distance for elderly subjects with abnormal gait (due to stroke, knee arthoplasty, and hip fracture respectively) over fixed time periods. The training used was overground walking at a comfortable pace and low intensity, and while the increase in distance walked by the end of the training was significant, the effect size was small. Park et al. [32] measured walking speed for healthy elderly, and found a significant increase after 48 weeks of training (20 minutes, thrice a week), but again, the effect size was small (0.4) [25].

It is difficult to combine gait training and walking training since gait training requires feedback from a therapist, while walking training requires more time than can be allotted by a therapist. Patients with abnormal gait are usually instructed by physiotherapists to walk with a heel first gait but there is often a regression of gait away from heel-to-toe when patients are not followed closely. Patients are usually given exercises to complete at home, but without the constant instruction of the physiotherapist, patients often forget the importance of heel strike. Training is more effective when guided, and one way to do so without a physiotherapist is to use portable electronics to measure gait kinematics. The next section describes literature relating to wearable devices for gait sensing applications.

2.5 Prior use of wearable devices

2.5.1 Sensing

Camera systems have been used extensively to make quantitative measurements of gait. These systems, particularly with recent advances in computer vision, can provide accurate kinematic measurements [33], but are normally restricted to laboratory environments, and are expensive. They also require post-processing, and so cannot be used in real time.

An alternative, and traditionally the gold standard for clinical gait analysis, is an instrumented walkway. These devices generally consist of a roll-up walkway, typically less than 10 metres long, embedded with pressure sensors that measure temporal and spatial gait parameters. Instrumented walkways give very good temporal accuracy when compared to video based systems, however spatial parameters (stride length, stride velocity) derived from these systems are significantly different from those of video based systems [34]. While these instrumented walkways have the advantage of being easy to set up and reliable, they can only measure a few steps of walking consecutively due to their limited length, and are currently too expensive for widespread home use.

However, the miniaturization of electronic sensors has opened up a market for wearable sensing; wireless sensors that are small and light enough to be worn without hindering movement. The unobtrusiveness and portability of these sensors allow the widespread analysis of gait and other movements without the need for a specialised laboratory environment. Thus miniature sensors can be transported between clinics easily, and given to the patient for home use. Computational devices capable of communicating with these sensors are also commonplace in the form of smartphones. Hence there has been much development of wearable systems for both scientific and commercial purposes. The majority of wearable systems for gait use a combination of accelerometers to measure linear acceleration and gyroscopes to measure angular velocity.

Temporal gait parameters can be calculated easily from the measurements using wearable sensors by extracting periodic features of gait kinematics. The time lag between consecutive features gives an estimation of cadence. Spatio-temporal parameters (walking speed, stride length) have been estimated using these sensors in a number of studies. For example, Aminian et al. [35] used gyroscopes on the shank and thigh to estimate cadence by detecting heel strike. Their method used a wavelet transform of the shank gyroscope signal to determine time and frequency characteristics. Heel strike and toe off events were detected by searching for local minima in the time domain within certain frequency bands. The angular velocity signals were integrated to give angular position, which was then used to find the linear position based on a double inverted pendulum model. Hundza et al. [36] described a similar approach, but used the zero crossing in shank angular velocity rather than its local minimum to define initial contact. They argued that this was a more general approach since the angular velocity of shuffling gait, associated with pathologies, does not have a local minimum. They also modified Aminian's double pendulum model to provide for a more accurate calculation of stride length. Zijlstra et al. [37] described a method, using a single trunk mounted accelerometer, of estimating stride length and walking speed using an inverted pendulum model.

Stride length and walking speed models often rely on the accurate detection of gait events, in particular heel strike and toe off. Several studies detail methods to detect these events. Aung et al. [38] detected both heel strike and toe off events on flat, inclined, smooth and rough terrains using tri-axial accelerometers on various locations (foot, ankle, shank, waist) for healthy subjects. They used a wavelet transform and a Gaussian mixture model to classify time points as either heel strike, toe off, or no event, achieving 86% accuracy. Since this method relies on a wavelet transform, it cannot be used in real-time. Aminian [35] also calculated heel strike and toe off in the process of finding stride length, however their algorithm also involved a wavelet transform.

For healthy subjects, heel strike is synonymous with initial contact (IC). Most analyses of gait events attempt to find initial contact in healthy subjects, while relatively few studies have attempted to detect initial contact in abnormal gait. Studies by O'Connor et al. [39], Desailly et al. [40], and Sharenkov et al. [41] used maximum metatarsal acceleration to detect initial contact in patients with abnormal gait (children with spastic diplegia, children with cerebral palsy, and total hip arthroplasty respectively). These studies detected initial contact with high temporal accuracy, however the kinematics were measured with motion capture rather than inertial measurement units.

Other studies have used inertial measurement units to reconstruct the full gait kinematics of lower body segments. Findlow et al. [42] collected data from healthy subjects and estimated hip, knee, and ankle angles from foot and shank inertial measurement units (accelerometer and gyroscope). This was carried out using a generalised regression neural network model to characterise the relationship between the IMU signals and the motion capture measurements. Rouhani et al. [43] developed a multi-segment joint angle measurement system using inertial measurement units located on the metatarsals, forefoot, hind foot and shank. They calculated lower limb segment orientation by integrating gyroscope signals. Drift error was accounted for using a three step method: 1) detecting heel strike and toe off using Aminian's wavelet transform method [35] (described above). 2) Detecting orientation of the segments during the quasi-stationary mid-stance (between heel strike and toe off) using the gravity vector. 3) Using this orientation as an initial condition for integration at every step.

Wearable inertial measurement units have also been used to distinguish gait from other activities. For example, Yoneyama [44] exploited three properties of gait (high intensity relative to other movements, periodicity, and biphasicity of movement in the superior-inferior, anterior-posterior directions) to detect periods of gait using a trunk mounted accelerometer. However, their algorithm is not applicable in real time since their frequency analysis method required access to future time points.

One major potential application of wearable sensing devices is the characterisation of movement disorders. Dalton et al. [45] were able to distinguish between healthy contols, presymptomatic Huntington's disease, and symptomatic Huntington's disease patients using a trunk mounted tri-axial accelerometer. They used an inverted pendulum model of trunk acceleration to calculate cadence, velocity, and step length and then classified the subject groups based on these variables. Guo et al. [46] presented an algorithm that distinguished between hemiplegic subjects and healthy controls, using 9 degree of freedom sensors mounted on the foot, shank, and thigh. Their algorithm extracted features (angle of foot at IC, angle at toe off, and maximum knee flexion), and used averages over each trial to classify subjects. Adkin et al. [47] used a single, two axis gyroscope to classify subjects as healthy or having Parkinson's disease using trunk sway magnitude.

The research described in this section shows that wearable devices have been used extensively for gait sensing applications. Algorithms have been developed for inertial measurement units attached to the trunk, thigh, shank, and/or foot to measure gait kinematics. Several studies have been able to accurately measure stride length using geometric models of the lower limbs. Many methods are also able to detect gait events (i.e. heel strike, toe off). However, the major drawback of the existing methods is that they are restricted to post hoc analysis. Frequency representations using transform methods of the IMU signals are generally needed, which means that real time calculation is not possible. Hence, these methods cannot be used to generate real time feedback. The next section explores literature pertaining to the use of feedback with wearable sensors for gait training.

2.5.2 Feedback

The research described so far deals with sensing applications of wearable measurement systems for gait. While this represents the majority of applications for wearable systems, another category of devices provides feedback to the user. Shull et al. [48] performed a detailed literature review of wearable devices for sensing and feedback of gait, and found 127 articles related to sensing applications, but only 14 articles pertained to feedback applications. Of these 14 articles, the majority (including [49, 50, 51, 52]) were focused on reducing trunk sway due to its strong correlation with balance. Only the few studies where feedback was applied to the lower limbs are reviewed here.

Shull [53] examined the use of haptic feedback applied to the trunk, knee, and foot to retrain gait in healthy subjects to reduce the knee adduction moment (the conventional physical therapy treatment for reducing symptoms of osteoarthritis). Gait kinematics were measured in real time with a motion capture system, and feedback was given to the subjects to encourage them to produce the desired changes. Three kinematic features were targeted for retraining: trunk sway, tibia angle in the frontal plane, and foot progression angle, i.e. the angle about the vertical axis. A vibration pad attached to the trunk would vibrate when the trunk sway was outside the desired range. Similarly, vibration pads on the knee and foot would vibrate when the tibia angle and foot angle were outside the desired range. The amplitude of the feedback was proportional to the error signal. The study reported a reduction in the knee adduction moment by at least 29% in all nine subjects over a 5 minute trial.

They also found that gait retraining based on a single kinematic feedback parameter was more effective than using three feedback parameters together.

Basaglia et al. [54] used auditory feedback to minimise the risk of knee hyperextension in patients with neurological disorders. Knee angle was measured using an electrogoniometer, and feedback was relayed to the patient at each full knee extension. Patients showed significantly more improvement with this technique than for controls with no feedback. Koheil and Mandel [55] performed a similar study using an electrogoniometer on a single stroke patient for gait rehabilitation purposes, applying haptic feedback to the knee when full extension and full flexion were reached. They concluded that "biofeedback techniques are useful as an adjunct to physical therapy with stroke patients".

2.5.3 Summary

Wearable devices are increasingly being used to assist with gait analysis and rehabilitation. From our literature review we have shown that there are published applications of inertial measurement units (accelerometers and gyroscopes) for gait sensing and feedback. From a sensing perspective, accelerometers and gyroscopes have been used extensively to monitor human gait. Many studies have successfully detected individual steps, and several studies have been able to detect separate phases (heel strike and toe off). However, the majority of these studies rely on time-frequency analyses and so cannot be used in real time. In addition, these methods were designed for healthy gait where heel strike necessarily precedes toe contact. This is not the case for most abnormal gaits. Some studies have looked at the use of inertial measurement units to distinguish between healthy and abnormal gaits, but generally classify patients' disorder in general, not on a step by step basis.

Papers that describe the use of feedback for gait rehabilitation are not common, and the few which have been found relate to either minimising trunk sway or limiting knee extension/flexion. These studies do show success in retraining gait, both over the course of a single trial and long term, which is promising. Shull [48] published a detailed literature review in 2014 that concluded:

> "This review shows that the benefits of real-time biofeedback for gait retraining through wearable feedback devices are underutilized; evidenced by the fact that no articles reported wearable feedback for training the hip, pelvis, thigh, or ankle. Wearable systems have the potential to extend the benefits demonstrated in laboratory biofeedback studies to a broader population."

2.6 Rationale

Conventional physiotherapy to restore heel-to-toe gait is often restricted by the time the physiotherapist can spend with each patient; patients often revert to a shuffling gait away from the clinic. While physiotherapists prescribe exercises, patients can forget the importance of heel strike, highlighting a need for a way to instruct patients on gait technique in a non-clinical environment. This leads to the following hypothesis: a device for home use that distinguishes between heel-to-toe and shuffling gait and provides feedback to the user can supplement physiotherapist instruction and improve rehabilitation outcome.

This device should detect the difference between a heel-first step and a flat foot or toe-first step, and provide feedback to the user on a step by step case. To be effective, feedback must be provided for every step, in real time, and the device must be portable, relatively cheap, and unobtrusive.

Based on our literature review, we can believe that such a biofeedback device to improve heel-to-toe gait is indeed novel, and has the potential to produce long term benefits for elderly patients.

CHAPTER 3
Methods

This chapter details the methods used to collect and process data for gait classification, including: the data acquisition hardware, the PC user interface, the classification tools, and the preliminary experiments conducted to gather data for algorithm development.

3.1 Hardware and software

The data acquisition equipment consists of two major components: the sensing module and the user interface.

3.1.1 Sensing module

The main requirements for the sensing hardware were that it should be inexpensive, portable, unobtrusive, and capable of accurately detecting heel strike. Wireless communication was also desirable. These requirements ruled out large, expensive gait analysis hardware such as video motion capture devices or instrumented walkways. Pressure and strain sensors fulfilled these requirements, and have been used previously to detect gait parameters accurately [56]. However, they are usually custom built, and building a custom sensing module was beyond the scope for this work. Inertial measurement units (IMUs) have been used extensively for gait applications and are available off the shelf, hence an IMU was chosen as the sensing module.

The Shimmer 2r wireless development kit, produced by Shimmer Sensing [57], was selected as the sensing module. Specifically, the motion development kit was purchased; this included 3 sensor units, 2 USB charging docks, and 2 foot straps. Figure 3–3 shows a block diagram of the sensor unit, which comprises:

- 3 axis linear accelerometer (Freescale MMA7361)
  - Range $\pm 6g$
  - Sensitivity $200mV/g$
- 3 axis gyroscope (InvenSense 500 series MEMS Gyros)
  - Range $\pm 500deg/s$
  - Sensitivity $2mV/\deg/sec$
- Microcontroller (MSP430) with 8 channels of 12 bit A/D
- Bluetooth connectivity - raw data can be transmitted to a PC or Android device using the Bluetooth wireless technology standard
- Size – 53mm x 32mm x 15mm

This device was chosen for the sensing module because of its relatively low cost (282CAD per unit), ability to interface with both MATLAB and the Android operating system, wireless connectivity, and set up ease. Development of software for the unit was straightforward since drivers were supplied for both MATLAB and Android. The MATLAB driver made it possible to develop the software in the environment used in our laboratory and utilize MATLAB's extensive range range of toolboxes. The Android driver provided the tools needed to communicate with the biofeedback module (used in the clinical pilot study, Chapter 5). Figure 3–2 shows a picture of the sensing device with the foot strap attached. Figure 3–1 shows the sensing module attached to a subject's shoe, along with the measurement coordinate system. The 3 acceleration channels were labelled as xAccel, yAccel, and zAccel, and the 3 gyroscope channels labelled as xGyro, yGyro, and zGyro. The device was positioned so that the z axis was parallel with the subject's medial-lateral axis, and held in place using an elastic strap around the shoe.

Figure 3–1: Photo of Shimmer device strapped to subject's shoe. The 3 acceleration channels are labelled xAccel, yAccel, and zAccel, while the 3 gyroscope channels are xGyro, yGyro, and zGyro. The measurement coordinate system is shown.

The Shimmer sensing devices were calibrated using the manufacturer's calibration software (see the instruction manual [58] for details).

3.2 User interface

The preliminary experiments used a PC running Matlab for data logging. A graphic user interface for MATLAB was built to handle the data acquisition using MATLAB's GUIDE function.

The objective of the MATLAB GUI was to log and display data. Biofeedback was not necessary for the preliminary experiments, however classification based on

Figure 3–2: Shimmer 2r device with foot strap attached

angular velocity of the foot was still performed (see Section 4.1). Figure 3–4 shows a screen shot of this GUI.

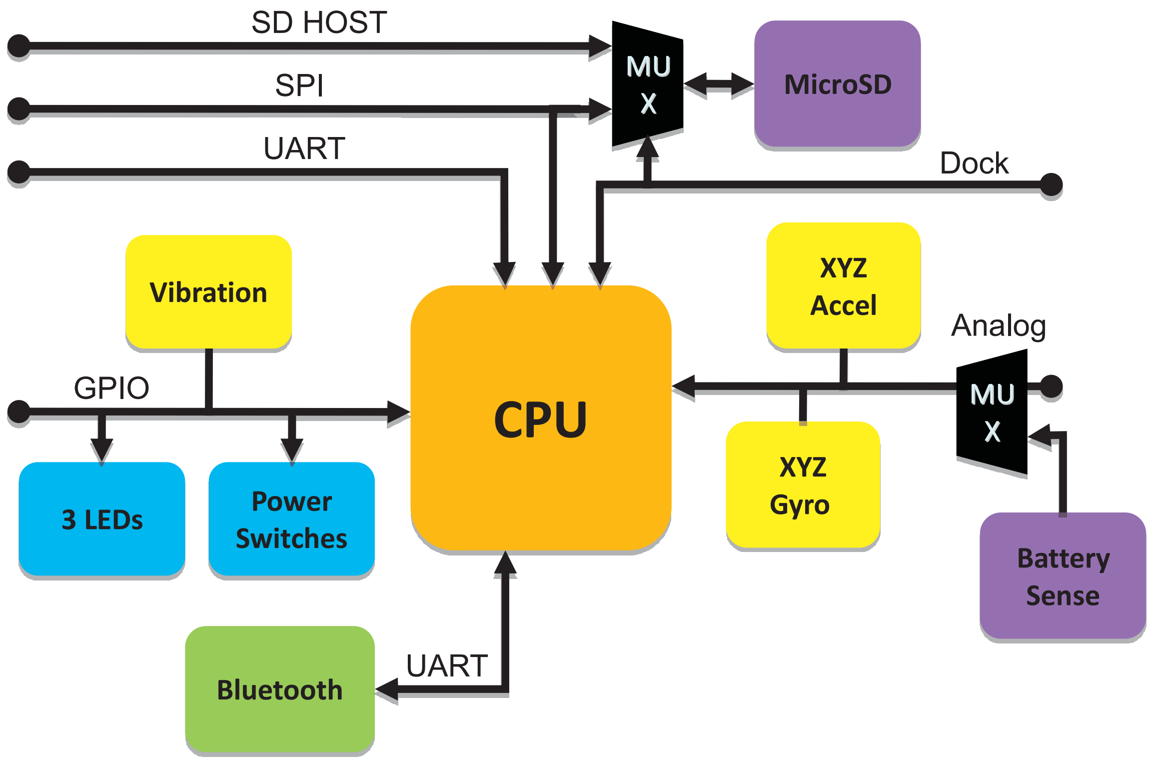

Figure 3–3: Shimmer 2r sensor block diagram. Adapted from Shimmer Research [57]

The MATLAB GUI allows the operator to select the COM port associated with the Shimmer device connected via bluetooth, and set the duration of a trial. When the operator pressed 'start', data was received in packets of 0.2 seconds and then plotted. The top panel shows the zAccel channel (useful for visualising the time of heel strike which manifests as spikes in this channel), while the bottom panel of Figure 3–4 shows the zGyro channel. In addition, the cirles show the estimated angular velocity at heel strike, colour coded as green (good step) or red (bad step). The angular velocity at heel strike was determined using the algorithm to be described in Section 4.1.

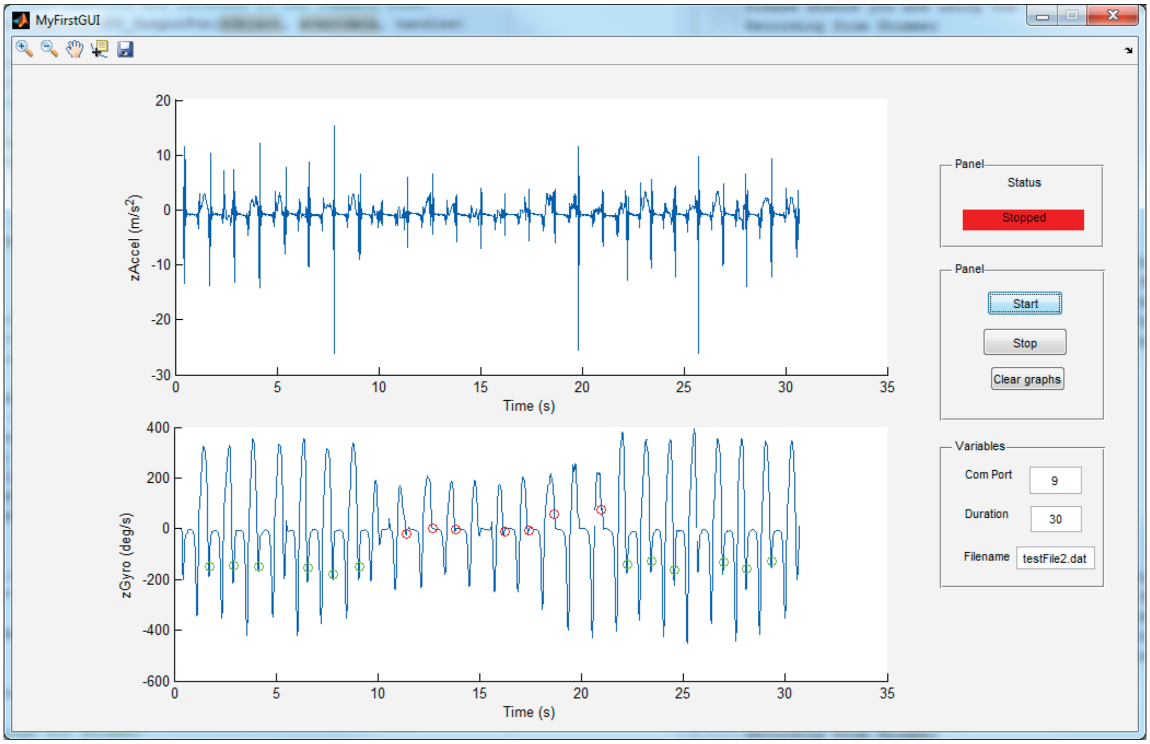

Figure 3–4: MATLAB Graphic user interface for Shimmer data collection. The GUI displays the zAccel and zGyro channels. The circles show the angular velocity at initial contact.

The raw six axis data was logged at 256 Hz to a `.dat` file, a generic data file containing tab separated values, while information about each detected step (time of step, quality of step, and angular velocity of foot at initial contact) was logged to a MATLAB `.mat` file.

3.3 Classification

The aim of the device was to categorize each step as good or bad in real time. This was achieved by first extracting features from the data that help distinguish between the two classes (feature extraction), and then classifying the step as "good" or “bad” based on these features. This section details the feature extraction and classification methods used.

3.3.1 Feature extraction

The most important task in this project was to identify features that separate good steps from bad. A brute force approach has been used frequently in gait studies (by Aung [38], Aminian [35], and Pendharkar [59], among others). Here, the signal (or signals, when using biaxial or triaxial sensors) is decomposed into many features that completely describe the signal at each time. Commonly these features are in the frequency domain, computed using Fourier or wavelet transforms, and comprise a set of magnitude values over a range of frequencies for every time. This produces many features for each gait cycle. For example, Aung et al. [38] used 30 features, evaluated every 50ms, to classify events as heel strike or toe off. The advantage of using large numbers of features is that classification performance is usually very good. However, it is very computationally intensive and since frequency domain transfers are used, cannot be carried out in real time.

This project required the analysis to be performed in real time and so we opted for a more analytical approach to feature extraction. Specifically, we used *a priori* knowledge of the gait cycle to select features likely to be useful for classification. Only small feature sets (less than 5) were considered to minimize computation cost and facilitate real time implementation.

3.3.2 Classifier

Once the relevant features are available they must be used with a classifier to distinguish between the two classes. Classification may be defined formally as:

> "Typically, we will be presented with a set of N training objects, $x_1..x_N$. Each is a vector with dimension D. For each object we are also provided with a label t_n that will describe which class object n belongs to. ... if, there are C classes, $t_n = \{1, 2, ..., C\}$. Our task is to predict the class t_{new} for an unseen object x_{new}." - Rogers and Girolami [60]

Here the dimension D refers to the number of features used in the classification process. There are many different types of classifiers, which can be subdivided into two main categories; probabilistic or non-probabilistic. These differ in the output they produce. Probabilistic classifiers calculate the probability that a new object belongs to each class; non-probabilistic classifiers simply assign a class to each new object. A non-probabilistic support vector machines classifier was used for this project. The following description of the classifier is adapted from Rogers and Girolami [60].

Support vector machines

A Support vector machine is a non-probabilistic classifier that only operates on two classes (binary classifier). The decision boundary between the two classes is generally linear, and is described by:

$$t_{new} - sign(\boldsymbol{w}^T \boldsymbol{x}_{new} + b) \tag{3.1}$$

where $\boldsymbol{t}_{new}$ is the class of the unseen object, $\boldsymbol{x}_{new}$ is the set of features for the unseen object, and $\boldsymbol{w}$ and b are parameters that define the boundary. The learning task involves using a set of training points to calculate the optimal values for these parameters. This is achieved by maximising the *margin* between the boundary and the nearest training points (called the support vectors). Figure 3–5 illustrates this concept for a 2 dimensionsional feature set.

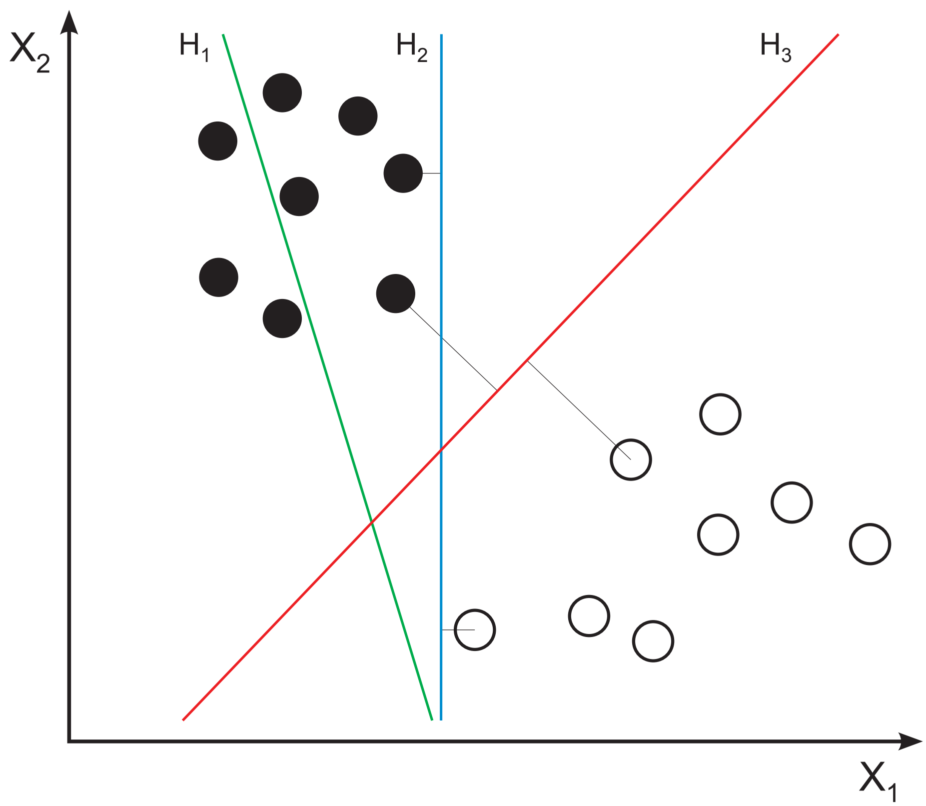

Figure 3–5: Support vector machine illustration. H1 does not separate the classes. H2 does, however only with a small margin. H3 maximises this margin. Source [61].

For two dimensions, the boundary that maximises the margin (H3 in Figure 3–5) best separates the classes. This is applicable to higher dimensions too. MATLAB has in built functionality for calculation of parameters $\boldsymbol{w}$ and b (the function `svmtrain` in the Statistics toolbox). The algorithm for calculating these parameters is described in detail by Rogers and Girolami [60].

3.3.3 Performance metrics

Once a classifier has been built it is necessary to assess its performance. This section presents the various tools and graphs used to quantify performance.

Statistical measures

With a two class classification problem such as ours, we can rename the two classes 'positive' and 'negative'. Generally, in medical applications, the diseased state is considered the positive class, and the healthy state is the negative class. Hence for this application we assign the positive label to bad steps and the negative label to good steps. The classifier uses the feature set extracted for each step to assign a positive or negative label to each step. This may be correct or incorrect, which gives rise to four possible categories:

- True positive (TP) - Correctly classified positive step
- True negative (TN) - Correctly classified negative step
- False positive (FP) - Incorrectly classified negative step
- False negative (FN) - Incorrectly classified positive step

With these categories of classification results defined, we can explore the most common statistical measures for classification performance, which are accuracy, sensitivity, and specificity.

Accuracy is the most intuitive measure, and represents the percentage of correct classifications made. It is formally defined as:

$$\text{Accuracy} = \frac{n_{TP} + n_{TN}}{n_{TP} + n_{TN} + n_{FP} + n_{FN}} \tag{3.2}$$

where n_{xx} is the total number of steps in each category.

This measure is useful, but not very informative when the training set is unbalanced. For example, if 90% of the training data were negative steps, a classifier that called everything negative would have 90% accuracy but be of no use. Hence, measures such as sensitivity and specificity must be used.

Sensitivity, also known as the true positive rate, measures the percentage of positive steps that are correctly identified:

$$\text{Sensitivity} = \frac{n_{TP}}{n_{TP} + n_{FN}} \tag{3.3}$$

Specificity, also known as the true negative rate, gives the proportion of negative steps correctly identified:

$$\text{Specificity} = \frac{n_{TN}}{n_{TN} + n_{FP}} \tag{3.4}$$

Sensitivity and specificity together give a more complete picture of the classification accuracy.

3.3.4 ROC curves

Accuracy, sensitivity, and specificity characterize the performance of the classifier for a fixed boundary. Usually the boundary that gives the maximum accuracy is used when reporting these values, however, in some applications it is desirable to have a different boundary, weighted towards higher sensitivity or specificity. Higher sensitivity usually comes at a cost of lower specificity and vice versa. Receiver operating characteristic (ROC) curves visualise this trade-off as the boundary is varied. Figure 3–6 shows a sample ROC curve. Classification performance improves as the curve moves from Curve C towards curve A. An overall measure of the performance of a classifier from an ROC curve is the area under the curve (AUC), with $AUC = 1$ being perfect, and $AUC = 0.5$ being classification by random chance.

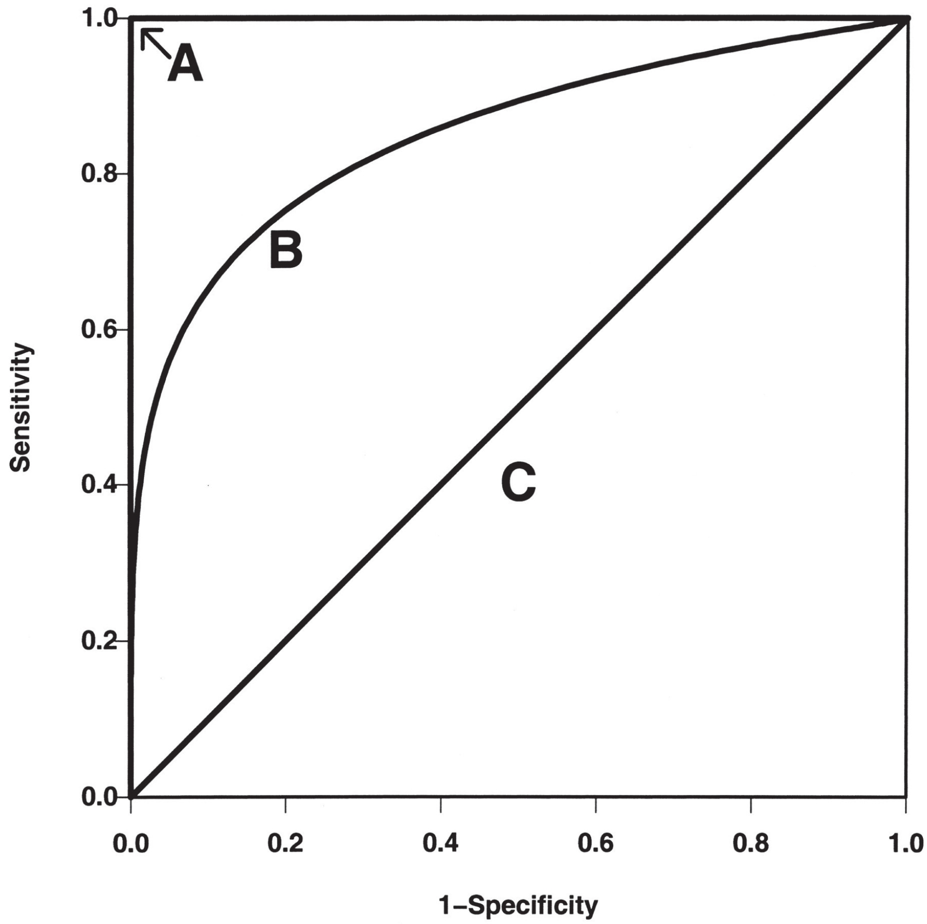

Figure 3–6: Example ROC curves. Curve A represents perfect classification performance, Curve B is a typical curve, and Curve C corresponds to random chance classification. Source [62].

3.3.5 Cross validation

When developing a classifier, it is generally preferable to have two separate data sets, one for training and one for validation. When the sample size is limited, however, k-fold cross validation can be used. This involves separating the total number of data points (or steps in this case) into k partitions. A single partition is used for validation, while the remaining partitions are used for training. This is repeated using each partition for validation exactly once, and mean values for the threshold, accuracy, sensitivity, and specificity can be calculated. 2 fold Cross validation was used for all accuracy values reported in this thesis.

3.4 Preliminary experiments

A set of preliminary experiments was carried out to gather data for use in developing the feature extraction algorithm.

Subjects were instructed to walk on a treadmill at a self-selected, comfortable speed, while data was streamed from the Shimmer device attached to their foot at 256 samples/second. Each trial lasted 60 seconds. A video was recorded simultaneously using a web cam (Microsoft Lifecam HD-3000), with a resolution of 720x480 at a frame rate of 30 fps.

Subjects were recruited from McGill University's School of Occupational and Physical Therapy. Most subjects had a healthy gait, but several subjects were trained physiotherapists who mimicked various types of abnormal gait to increase the number of bad steps for algorithm development and classification. Three sets of preliminary experiments were as follows:

3.4.1 Preliminary experiments 1 (P1)

There were two goals for this set. Firstly, to acquire the data needed to develop the feature extraction algorithm, and secondly to explore the effectiveness of the different features in classifying good and bad steps. Since algorithm validation was not planned for the P1 experiments, trials were not rated on a step by step case. Instead a high level rating was given to each trial (good, bad, or marginal).

The trial list is shown in Table 3–1, which shows the subject and walk type for the 13 trials acquired in P1.

Table 3–1: P1 subject list

Trial	Subject	Walk type
1	VB	Good
2	VB	Bad (mimic)
3	VB	Bad (mimic)
4	SF	Good
5	SF	Bad (mimic)
6	NM	Good
7	NM	Bad (mimic)
8	IOC	Bad
9	LM	Bad
10	LN	Marginal
11	SS	Marginal
12	PP	Marginal
13	SH	Good

The type of walk in each trial was assigned at the time of the experiment by a physiotherapist who monitored each trial.

3.4.2 Preliminary experiments 2 (P2)

The P2 experiments were conducted to validate the algorithm developed using the data from the P1 experiments. The full 60 seconds of video was recorded in these experiments. This permitted the physiotherapist to rate each step using the GUI described in Section 4.2. In addition, the walking speed, as displayed by the treadmill, was recorded. Table 3–2 shows the trials acquired.

Table 3–2: P2 subject list

Trial number	Subject	Speed (m/s)	Walk type
1	IOC-1	0.8	Diplegic
2	VB-1	1.6	Healthy
3	VB-2	1.6	Abnormal (mimic)
4	SH	1.8	Healthy
5	LN	1.9	Healthy
6	SS-1	1.6	Healthy
7	PP	1.9	Healthy
8	SF-1	2.3	Healthy
9	SF-2	1.0	Abnormal (mimic)
10	SF-3	1.0	Abnormal (mimic)

3.4.3 Preliminary experiments 3 (P3)

The P3 experiments were conducted to increase the sample size since the data gathered in the P2 set was highly unbalanced with only 20% of the steps rated bad. P3 experiments were similar to P2, except that the calibration routine to be described in Section 4.3 was executed prior to each trial to correct for any misalignment between the axes of the device and the ankle. The subject list is shown in Table 3–3.

Table 3–3: P3 subject list

Trial number	Subject	Speed (m/s)	Walk type
1	AH	2.1	Healthy
2	AR	2.2	Healthy
3	AV	3.3	Healthy
4	BL	1.6	Healthy
5	IOC	0.9	Diplegic
6	KJ	2.0	Healthy
7	SF	2.3	Healthy
8	SF	1.8	Abnormal (mimic)
9	SF	2.0	Abnormal (mimic)
10	VB	1.4	Healthy
11	VB	1.4	Abnormal (mimic)

Chapter 4 shows the results from the preliminary experiments.

CHAPTER 4
Algorithm development and validation

This chapter describes the development and validation of the algorithm for feature extraction. This includes a description of the algorithm with justifications, and the classification results from the preliminary experiments using this algorithm. In addition, the orientation calibration procedure is described, along with an assessment of its validity.

4.1 Feature extraction algorithm

The data from the preliminary experiments was used to develop the feature extraction algorithm. This was an iterative process in which the algorithm was repeatedly modified, validated, and refined. This section describes how the final algorithm extracts features from the data.

4.1.1 Hypothesis

The ideal feature to determine whether or not a step is good would be the angle of the foot with respect to the ground at initial contact. A large angle with the ground at heel contact would indicate a good step, while a small angle, corresponding to flat foot or initial contact made with the forefoot/toes, would indicate a bad step.

Unfortunately, angular position is not available directly from a 6 degree of freedom IMU. However, angular velocity is available and when a step has a clearly defined heel strike (i.e. a good step), the toes should rotate towards the ground faster than for a bad step. This lead us to the hypothesis:

The angular velocity of the foot around the medial-lateral axis at initial contact is a distinguishing feature between good and bad steps. The higher the velocity, the higher the quality of step.

Visual analysis of the preliminary data supported this hypothesis. Figure 4–1 shows data for a typical good step, while Figure 4–2 shows data for a bad step. The z axis in the measurement coordinate system corresponds to the medial-lateral axis with negative velocities corresponding to rotation towards the ground. Note that at initial contact, indicated by the high frequency spike in the acceleration signals, there is a local minimum in the zGyro signal for both the good and bad steps. Initial contact was determined using the video synchronised with the data. The amplitude of the minimum is more negative for the good step than for the bad, supporting the hypothesis that a more negative velocity is indicative of a good step.

With the angular velocity of the foot at initial contact established as the feature to be used for classification, it was then necessary to develop an algorithm to accurately extract it from the kinematic data. The next section describes this algorithm:

4.1.2 Algorithm

To extract the value of zGyro at initial contact, the first task is to identify the time at which initial contact occurs. Initial contact is associated with high frequency spikes in acceleration, which might be used as a marker for initial contact. However, it proved difficult to detect these spikes reliably, since they were often masked by unrelated spikes in the acceleration signal.

Instead we focused on the zGyro signal where the local minimum at initial contact seemed to be a promising marker. However, smoothing was necessary to

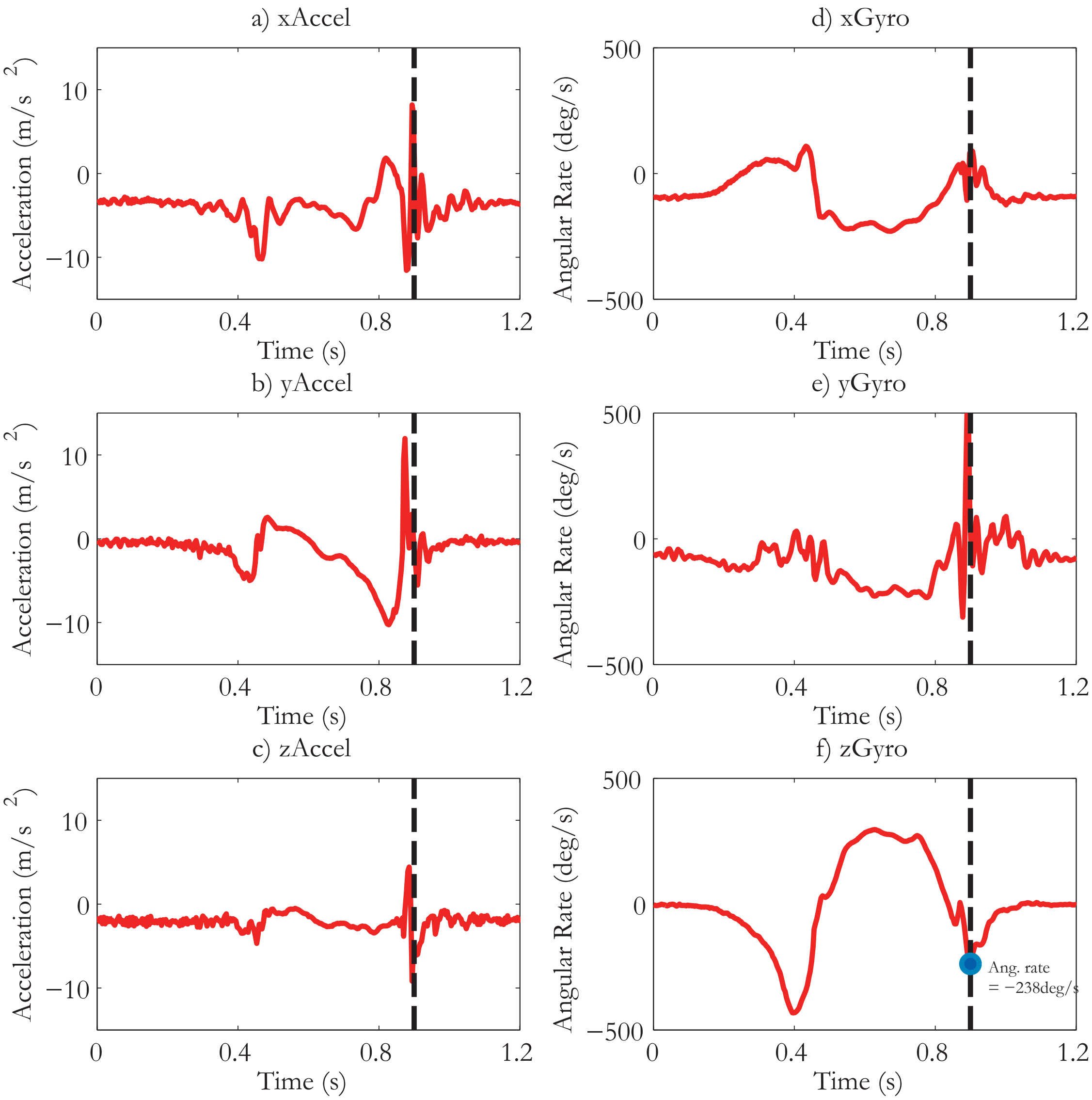

Figure 4–1: Sample gait cycle data for a subject with healthy gait (SF), showing all six channels. Initial contact is marked by the dashed line. Panels a, b, and c show the acceleration in the x, y, and z directions repsectively, and panels d, e, and f show the angular velocity about the x, y, and z axes.

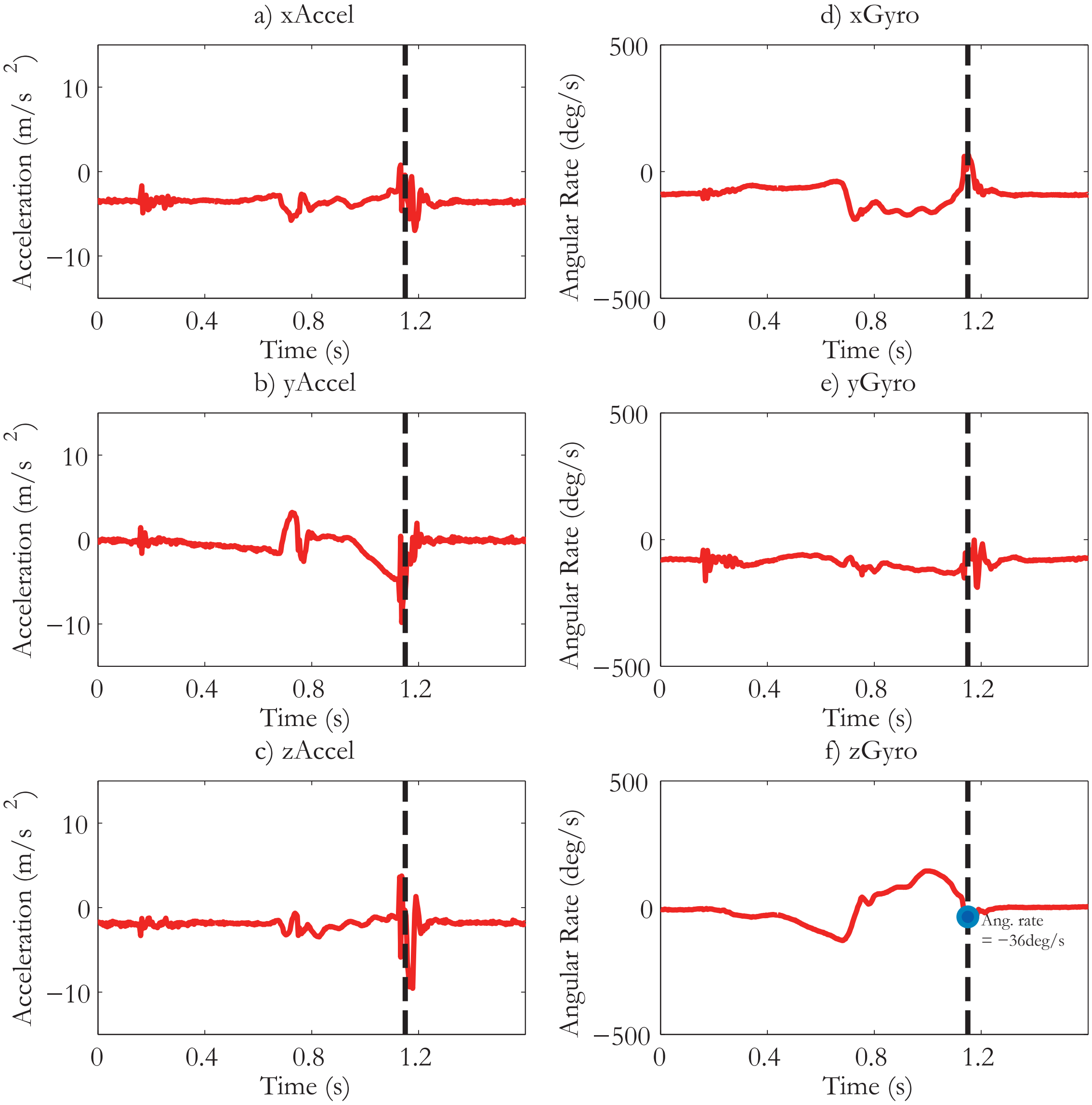

Figure 4–2: Sample gait cycle data for a subject with abnormal gait (IOC), showing all six channels. Initial contact is marked by the dashed line. Panels a, b, and c show the acceleration in the x, y, and z directions repsectively, and panels d, e, and f show the angular velocity about the x, y, and z axes.

prevent high frequency noise in the signal from masking the minimum of interest. A Butterworth IIR filter with an order of 3 and a cutoff frequency of 4Hz was chosen, due to its ability to meet the required cutoff specification with a low order (fast computation). A cutoff frequency of 4Hz was chosen since it was found to eliminate minima caused by noise while preserving the minimum associated with initial contact. The parameters for this filter are:

$$\begin{aligned} b &= \begin{bmatrix} 0.0096 & 0.029 & 0.029 & 0.0096 \end{bmatrix} \\ a &= \begin{bmatrix} 1.00 & -2.03 & 1.47 & -0.37 \end{bmatrix} \end{aligned} \tag{4.1}$$

where b contains the feedforward coefficients and a contains the feedback coefficients.

Figure 4–3 shows the zGyro signal from a typical gait cycle after filtering. There are two local minima present in the gait cycle, one occurring at initial contact and one associated with toe off. Since there is more than one minimum, the time of initial cannot be identified solely on the basis of zGyro being a minimum.

However, we noted that zGyro has only one positive local maximum in the gait cycle, associated with the mid-swing phase. Consequently, this point, which we term the cycle division point, can be used to split the signal into discrete cycles. Then, the local minimum at initial contact can be distinguished from the local minimum at toe off based on the time with respect to the cycle division point. Initial contact occurs immediately after the cycle division point, while toe off occurs immediately before the cycle division point. This leads to the algorithm to determine angular velocity of the foot at initial contact, ω_{IC}, summarized below, with reference to Figure 4–3.

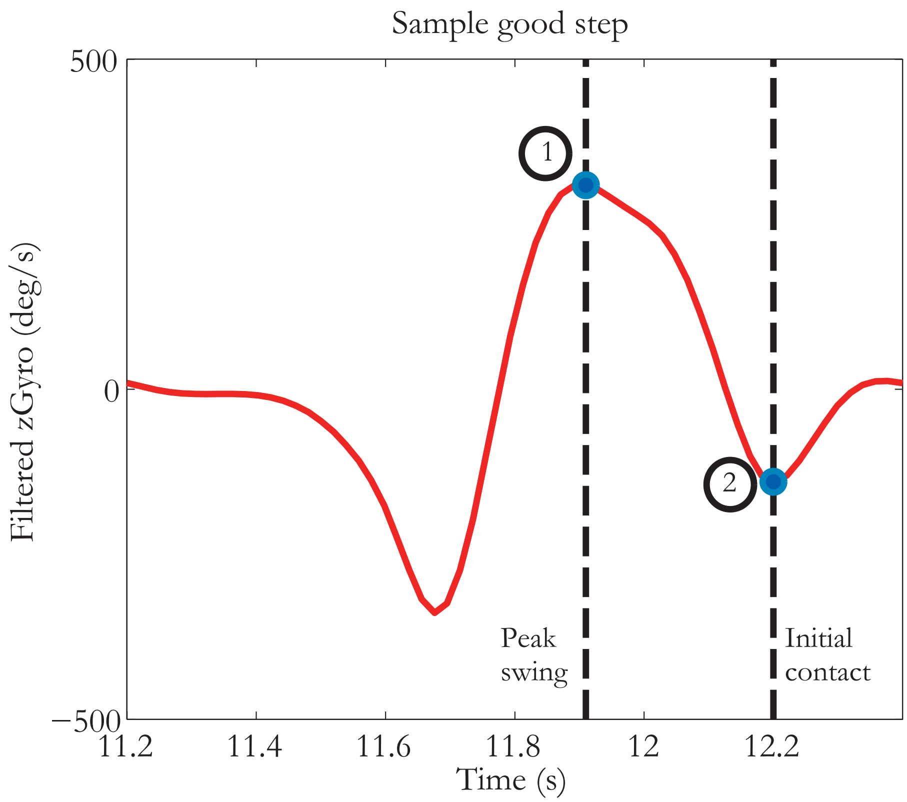

Figure 4–3: Filtered zGyro channel for a sample gait cycle. Some points of interest relating to the feature extraction algorithm are highlighted.

Algorithm description

1. Stream data from Shimmer at 51.2 samples/second.
2. Rotate the 3 dimensional gyro vector using calibration parameters to account for device orientation (see Section 4.3).
3. Lowpass filter the zGyro data (using filter described in Equation 4.1.
4. Search for the cycle division point (point 1), that is:
   - Greater than the preceding point

- Greater than the following point

  AND

- Greater than 50 deg/s

5. Once a cycle division point is found, wait 140ms seconds.
6. Search for initial contact point, that is:
   - Less than the preceding point
   - Less than the following point

     AND

   - Less than 100 deg/s

   The filtered zGyro value at this point is the feature of interest, ω_{IC} (point 2).
7. Begin searching for the next cycle division point 390ms after initial contact
8. If initial contact is not found within 780ms seconds of the cycle division point, reject the current cycle division point and begin searching for a new one.

The decision boundary referred to in item 6 of the algorithm description is determined by the classifier, details are given in Section 4.4.

4.2 Physiotherapist labels

The goal of the algorithm was to determine an accurate good/bad rating for each step. Consequently, to train the algorithm and assess its accuracy it was necessary to have a dataset where the true rating of each step was known. This was achieved by having a physiotherapist watch the videos frame-by-frame and assign a rating to every step. The labels assigned were good, bad, or marginal, based on the angle of the foot at initial contact. A flat foot initial contact or forefoot strike was rated as bad, a step where the foot makes a small angle with the ground at heel contact was rated as marginal, and a clearly defined heel strike was rated as good. These class labels were applied based on the physiotherapist's experience with the target population. A GUI was developed to allow the physiotherapist to easily navigate through the video frame by frame, rate each step, and export the results. Figure 4–4 shows a screen shot of this GUI.

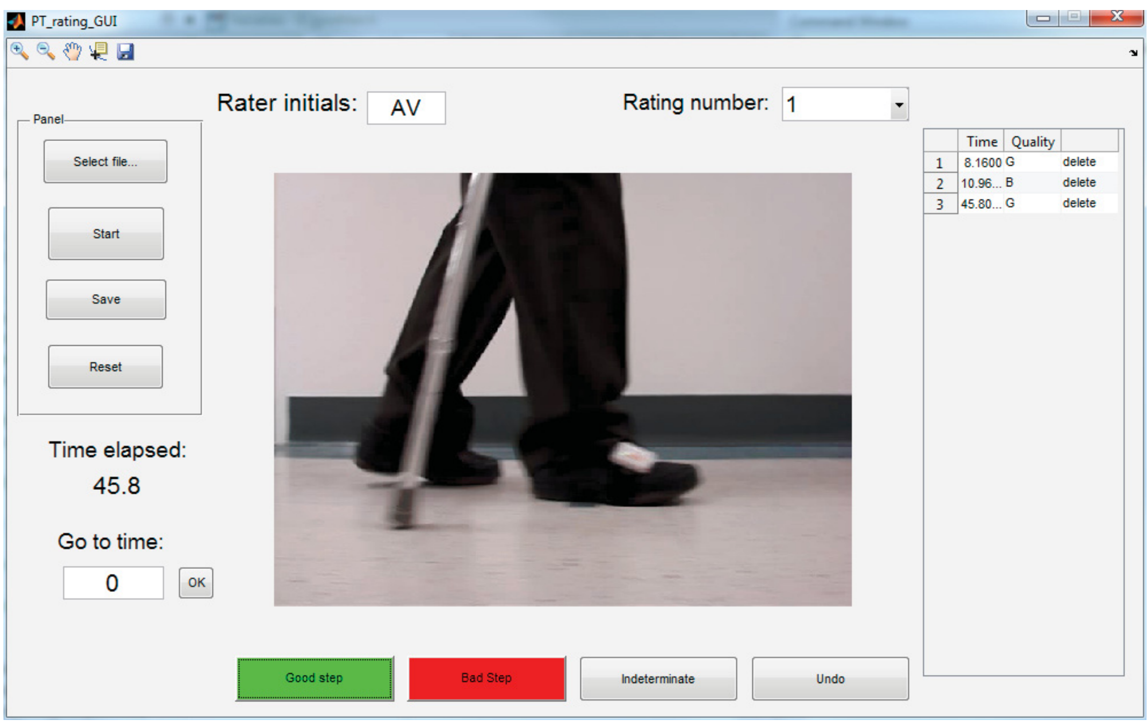

Figure 4–4: GUI for physiotherapist to rate each step from recorded video

4.3 Device orientation calibration

The nominal placement of the sensor on the foot, shown in Figure 3–2, is with the z-axis aligned with the ankle's medial-lateral axis. Since the algorithm measures the value of zGyro at initial contact, the performance of the algorithm will depend on accurate placement of the sensor module. Misalignment will result in partial projection of the ankle rotation onto the x and y axes, so that rotation about the medial-lateral axis will be split over x, y, and z axes instead of just the z axis. Hence the magnitude of the zGyro will be reduced, and classification accuracy will be affected.

A calibration routine was developed to correct for this. Since the sensor unit houses an accelerometer, the vertical direction can be detected on the basis of the acceleration due to gravity. The principle of operation for this calibration routine is to detect 'up' relative to the sensor coordinate system, and then rotate the measured vectors to align 'up' with the x axis. The calibration routine is outlined below:

1. Calculate mean x,y and z acceleration (a_x, a_y, a_z) for one second of data while subject is stationary with foot flat on the ground
2. Calculate axis and angle of rotation (given $\vec{g} = [0, 0, -1]$ and $\vec{g\prime} = \frac{[a_x, a_y, a_z]}{\sqrt{a_x^2 + a_y^2 + a_z^2}}$:

$$\vec{u} = \vec{g} \times \vec{g\prime} \tag{4.2}$$

$$\theta = \cos^{-1}(\vec{g} \cdot \vec{g\prime}) \tag{4.3}$$

3. Generate the rotation matrix:

$$R = \cos\theta \boldsymbol{I} + \sin\theta[\boldsymbol{u}]_{\times} + (1 - \cos\theta)\boldsymbol{u} \otimes \boldsymbol{u} \tag{4.4}$$

4. Rotate raw data:

$$\vec{v'} = R\vec{v}$$

A simulation study validating this procedure is given in Section 4.4.3.

4.4 Preliminary results

This section gives the main findings from the preliminary experiments, i.e. the classification performance of the algorithm described in Section 4.1.

4.4.1 P1

The aim of this set of experiments was to test the feature extraction hypothesis sufficiently enough to justify conducting a more thorough validation experiment. Figure 4–5 shows the distribution of ω_{IC} for all trials in the P1 experiments. There was a clear separation between good trials (as rated by physiotherapists, shown in green), marginal trials (yellow), and bad trials (red). Trials with better quality steps had a more negative ω_{IC} than marginal or bad trials. This supported our initial hypothesis (Section 4.1.1) and suggested that it was worthwhile to perform additional experiments to validate the algorithm.

4.4.2 P2

The feature extraction algorithm was applied to the raw data to extract ω_{IC}. To determine the success of the algorithm at correctly detecting steps, the video was analysed, and the time of each heel strike was manually recorded and compared with that determined by the algorithm. Of the 394 steps recorded, the algorithm was successful in detecting 393 steps, or 99.8%. In addition, no false steps were detected.

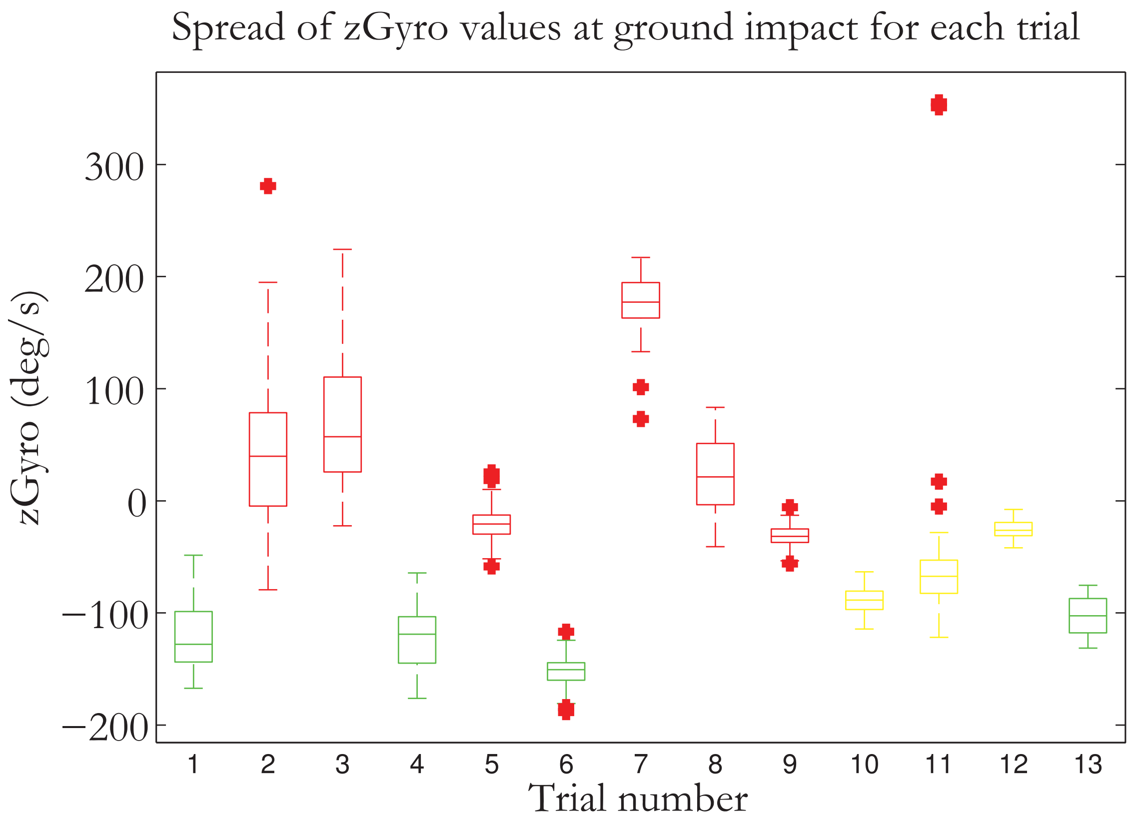

Figure 4–5: Distribution of ω_{IC} for each trial in P1 experiments. Trials are coloured by average step quality, green for good, red for bad, and yellow for marginal. The boxes give the highest and lowest values not considered to be outliers, the upper and lower quartiles, and the median. Outliers are shown as separate points.

Figure 4–6 shows the distribution of ω_{IC}. There is a good separation of the two classes, with some steps occurring on the wrong side of the decision boundary, determined using SVM for classification, as described in Section 3.3. Figure 4–7 shows the probability density function for each class of step, demonstrating a significant degree of overlap. Figure 4–8 shows the tradeoff between sensitivity and specificity defined by the ROC curve for classification. Table 4–1 gives other performance metrics.

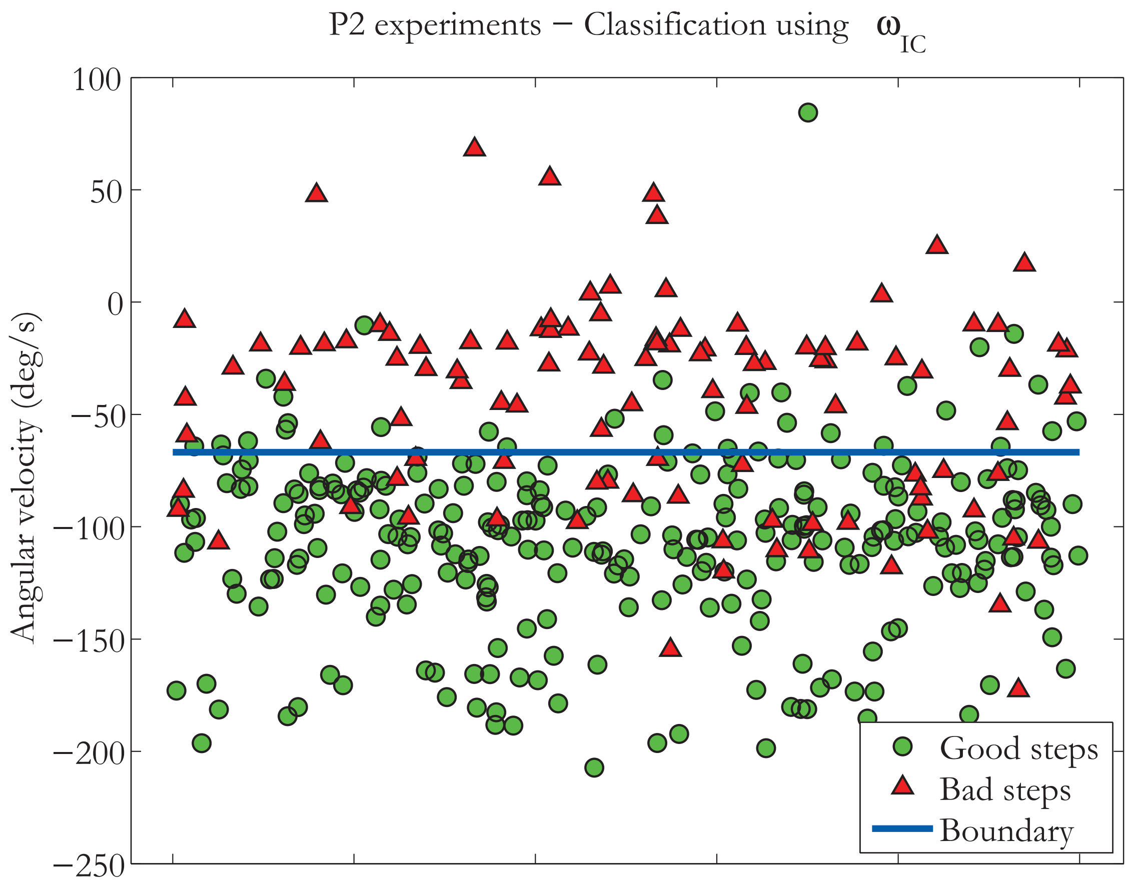

Figure 4–6: Scatter plot of individual steps in P2 experiments, coloured by gold standard rating. The classification boundary for optimal accuracy is also shown.

Table 4–1: Performance metrics for classification of P2 experiment set

Sensitivity at 75% specificity	0.764
Area under the ROC curve	0.847
Maximum accuracy	81.2%

4.4.3 Orientation calibration results

A simulation study exploring the effects of sensor orientation on classification performance was conducted. Three dimensional experimental data from the P2 experiment set was rotated by α degrees about the x axis, and then by β degrees about

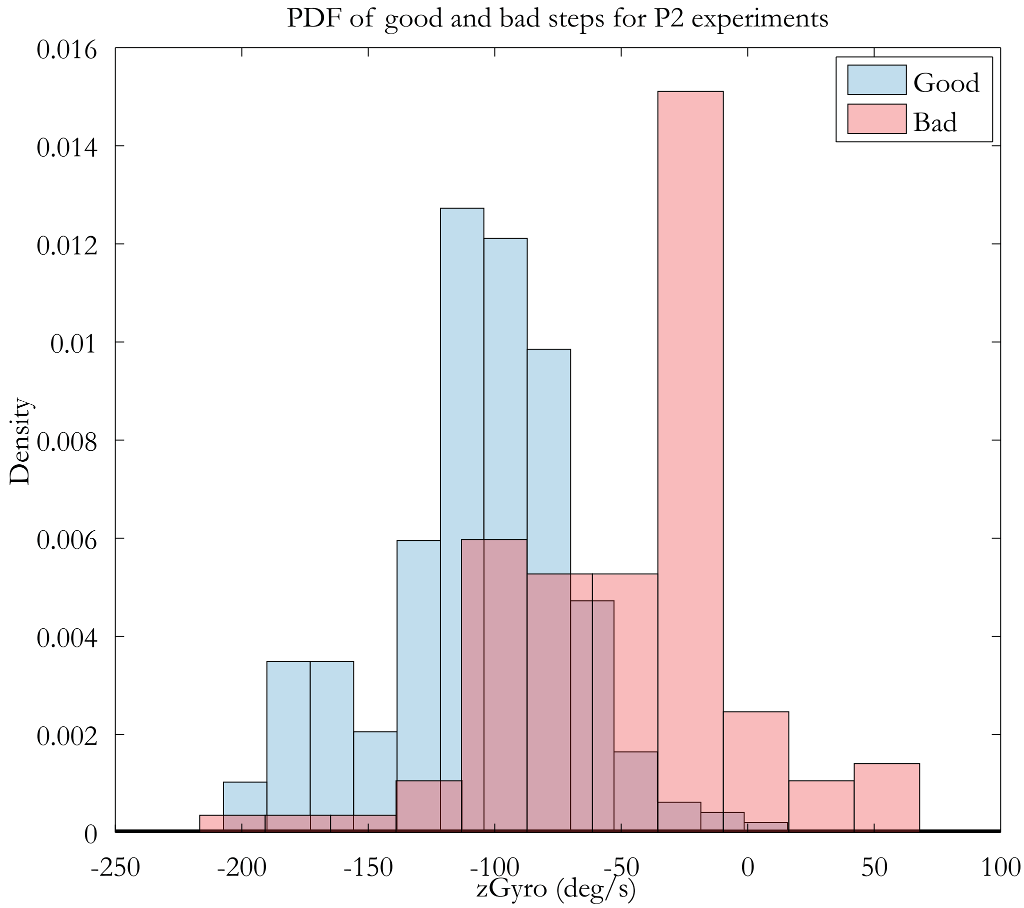

Figure 4–7: Probability density function for the P2 experiments. Each class of step is shown separately.

the y axis to produce data from a simulated misaligned sensor. The transformation was performed using intrinsic rotations, as per the equations below:

$$\vec{v}' = R\vec{v} \tag{4.5}$$

$$R = X(\alpha)Y(\beta)Z(\gamma) \tag{4.6}$$

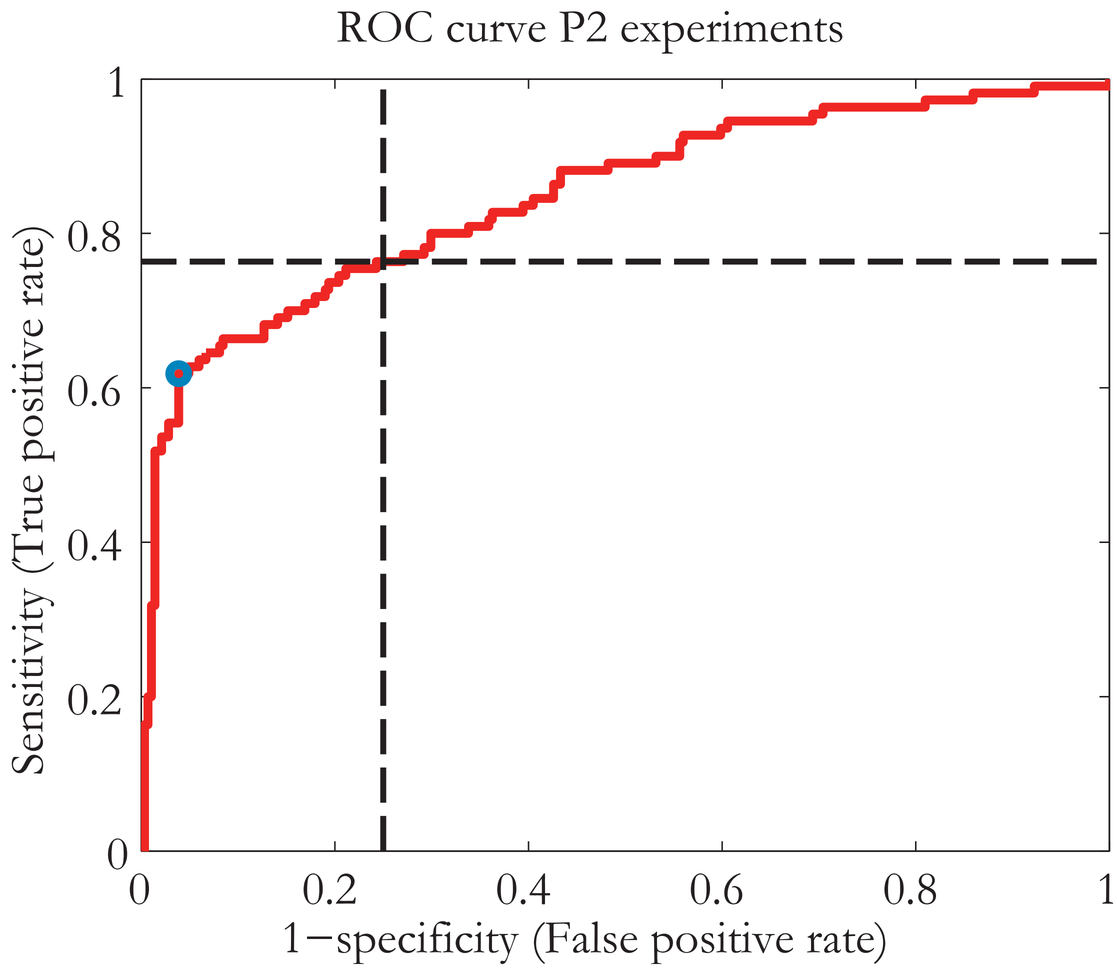

Figure 4–8: ROC curve for classification of P2 experiment data. The accuracy optimal operating point is highlighted in blue, and the cross-hairs show the sensitivity at 75% specificity.

$$R = \begin{bmatrix} c(\beta)c(\gamma) & -c(\beta)s(\gamma) & s(\beta) \\ c(\alpha)s(\gamma) + c(\gamma)s(\alpha)s(\beta) & c(\alpha)c(\gamma) - s(\alpha)s(\beta)s(\gamma) & -c(\beta)s(\alpha) \\ s(\alpha)s(\gamma) - c(\alpha)c(\gamma)s(\beta) & c(\gamma)s(\alpha) + c(\alpha)s(\beta)s(\gamma) & c(\alpha)c(\beta) \end{bmatrix} \tag{4.7}$$

Here v is the raw data, $v\prime$ is the rotated data, and R is the rotation matrix. The rotation matrix is defined using Tait-Bryan angles, also known as yaw, pitch and roll. The rotations α, β, and γ are performed sequentially about the x axis, the new y axis ($y\prime$) and finally the new z axis $z\prime\prime$ [63]. c and s in Equation 4.7 are shorthand for cos and sin.

Figure 4–9 shows some sample data, and the effect of a misaligned sensor on the signal. Each signal is distorted slightly with the rotations. The zGyro signal is of particular importance, and the rotated signal has high frequency spikes at initial contact, projected from the xGyro and yGyro signals. This would reduce the accuracy of the algorithm.

Figure 4–10 shows the results of this simulation study. Note that misalignment about the z axis γ was not considered, since the zGyro channel is not affected by a static rotation about the z axis. Following rotation of the vectors, features were extracted as usual (see Section 4.1). From the accuracy optimal decision boundary, an accuracy was determined, this is shown in Figure 4–10 as a function of α and β (note that cross validation was not performed for this study). Here we can see that the accuracy of the classification method drops off sharply as the misalignment angle increases. This is an issue since footwear geometry is by no means uniform from patient to patient, and positioning the device precisely parallel to the sagittal plane is unreasonable in some cases. Hence a correction for this phenomenon was sought.

A calibration procedure to correct for this effect was developed, as described in Section 4.3. Figure 4–11 shows the accuracy results after calibration. The same simulated data and procedure used in generating Figure 4–10 was used, except prior

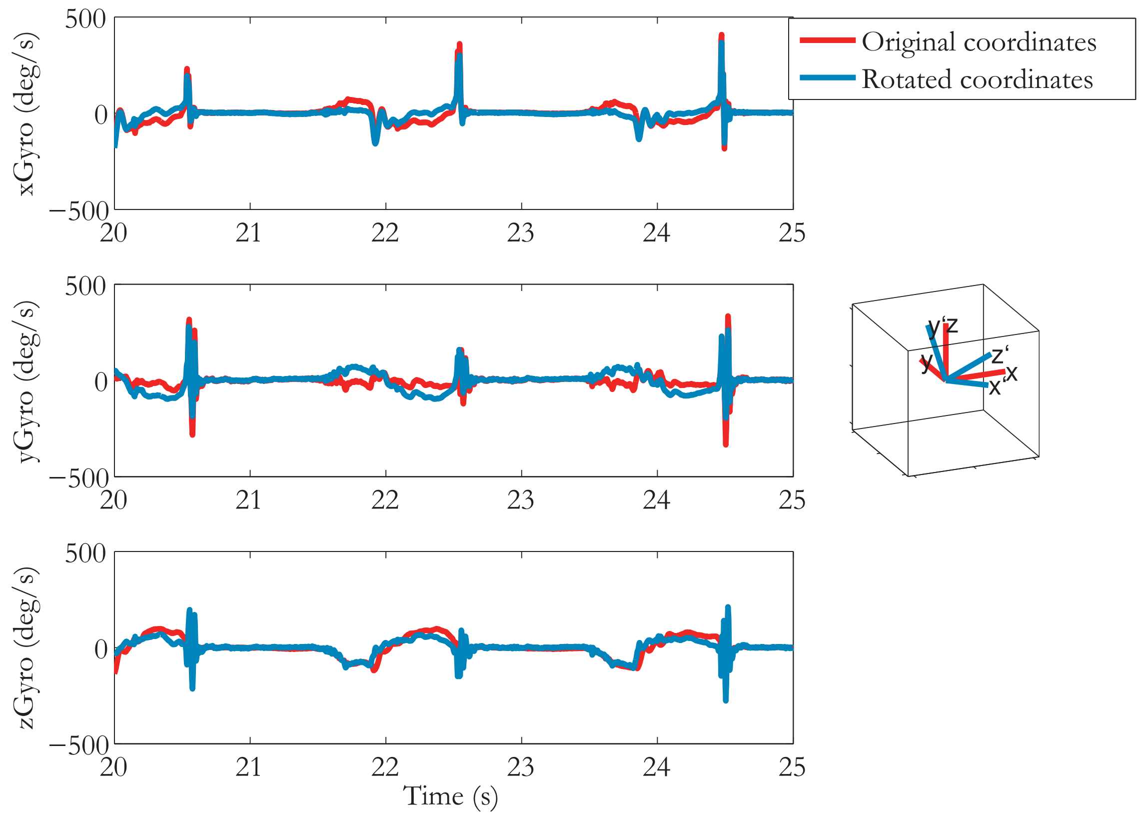

Figure 4–9: Sample data, raw (shown in red) and rotated (shown in blue) by 45 degrees about the x axis and 30 degrees about the y axis. Both coordinate systems are shown in the right hand panel.

to classification the data was calibrated using the mean acceleration signal to realign the x axis with gravity.

Figure 4–11 shows that the accuracy still reduces as a function of α. This is to be expected since rotation about the x axis does not change the gravity vector direction, so calibration with respect to gravity cannot fix misalignments about the x axis. While not ideal, this is not much of an issue since these misalignments are unlikely due to the geometry of the foot. Placement of the device anywhere from the

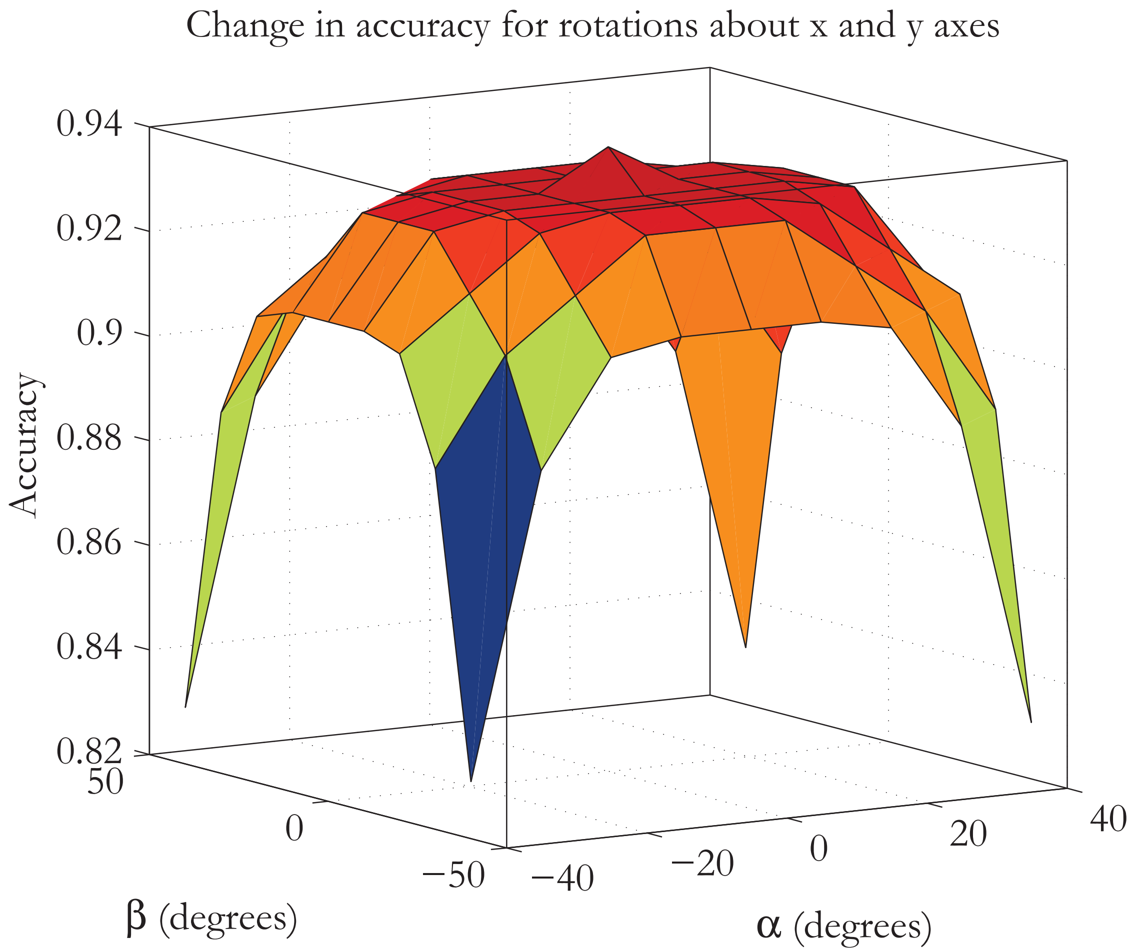

Figure 4–10: Simulation showing classification accuracy as a function of device misalignment about the x (α) and y (β) axes

ankle to the side of the toes should yield an accurate alignment about the x axis. This is not the case for alignment of the y axis, however, as a slight misalignment vertically can result in a large angular deviation of the y axis. Therefore, calibration to correct for this misalignment is crucial. As we can see in Figure 4–11, calibration results in an *accuracy that is independent of* β. This is an important result, showing that calibration is indeed effective at cancelling misalignment error for the y axis.

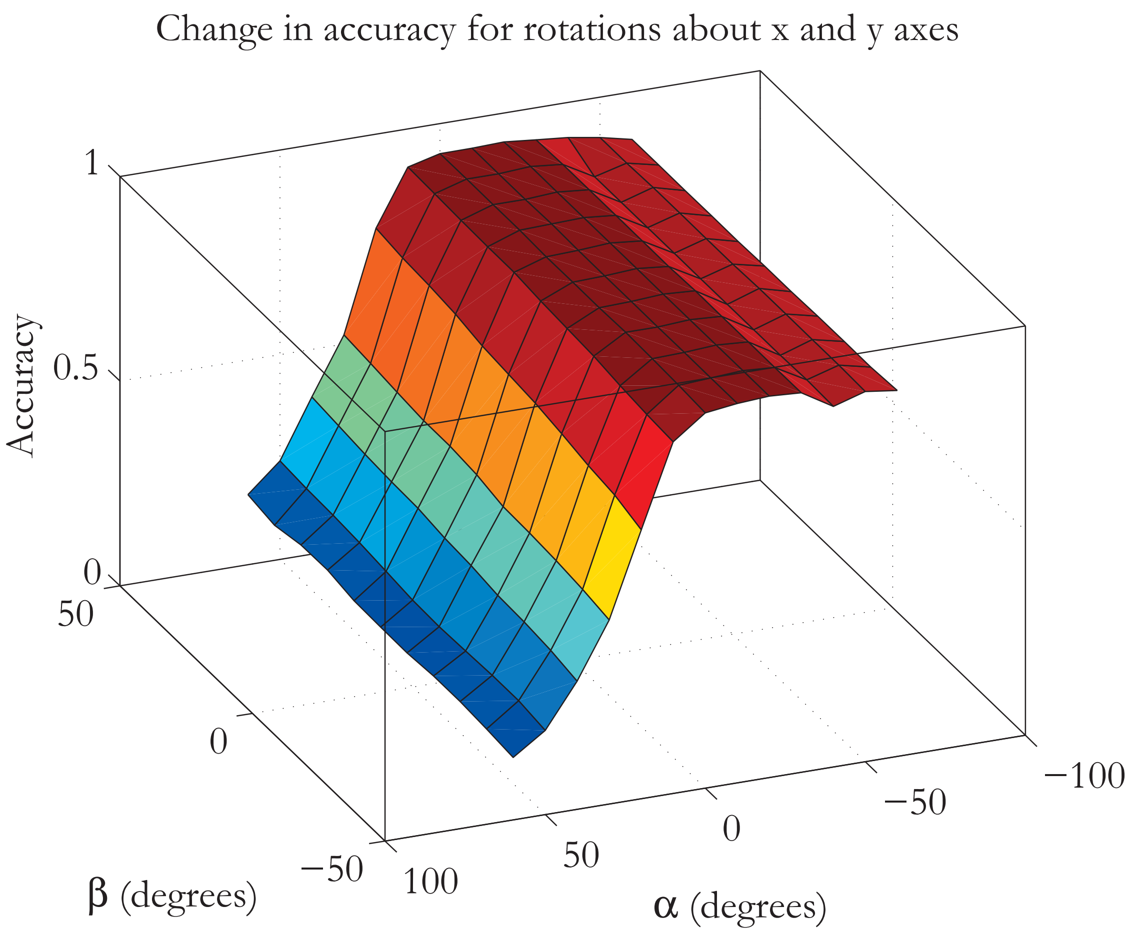

Figure 4–11: Simulation showing classification accuracy as a function of device misalignment about the x (α) and y (β) axes, after calibrating for orientation.

4.4.4 P3

Of the 483 steps recorded, 482 were detected by the algorithm, for a success rate of 99.8%. Figure 4–12 shows the distribution of initial contact temporal error for every step, determined by comparing the time of initial contact by manual analysis of the video to that determined by the algorithm. The error was measured in terms of the number of video frames, leading to the discrete nature of the temporal error. The majority of steps had errors of less than 3 frames (frame rate of 30fps). Errors larger than 3 frames were generally due to the algorithm detecting a minimum not coinciding with initial contact, and these steps often produced incorrect classifications.

Figure 4–13 shows the distribution of ω_{IC}, colour coded by step quality. It is apparent that the split of good/bad steps is less defined than the results from the P2 experiments, with significant overlap between the two classes. Figure 4–14 shows the probability density function for this set of steps, which highlights a considerable overlap between the good/bad classes. This is reflected in the ROC curve (Figure 4–15) which has a smaller area under the curve, suggesting poorer classification. Performance metrics are given in Table 4–2.

Table 4–2: Performance metrics for classification of P3 experiment set

Sensitivity at 75% specificity	0.815
Area under the ROC curve	0.856
Maximum accuracy	61.5%

The calibration procedure was applied for the P3 data set, however it was possible to perform the reverse transformation to determine the effect calibration has on the classification performance. Table 4–3 shows the maximum accuracy and ROC curve area for the uncalibrated and calibrated data. Both metrics are slightly better with calibration, confirming that it was worthwhile persisting with this method in the clinical study.

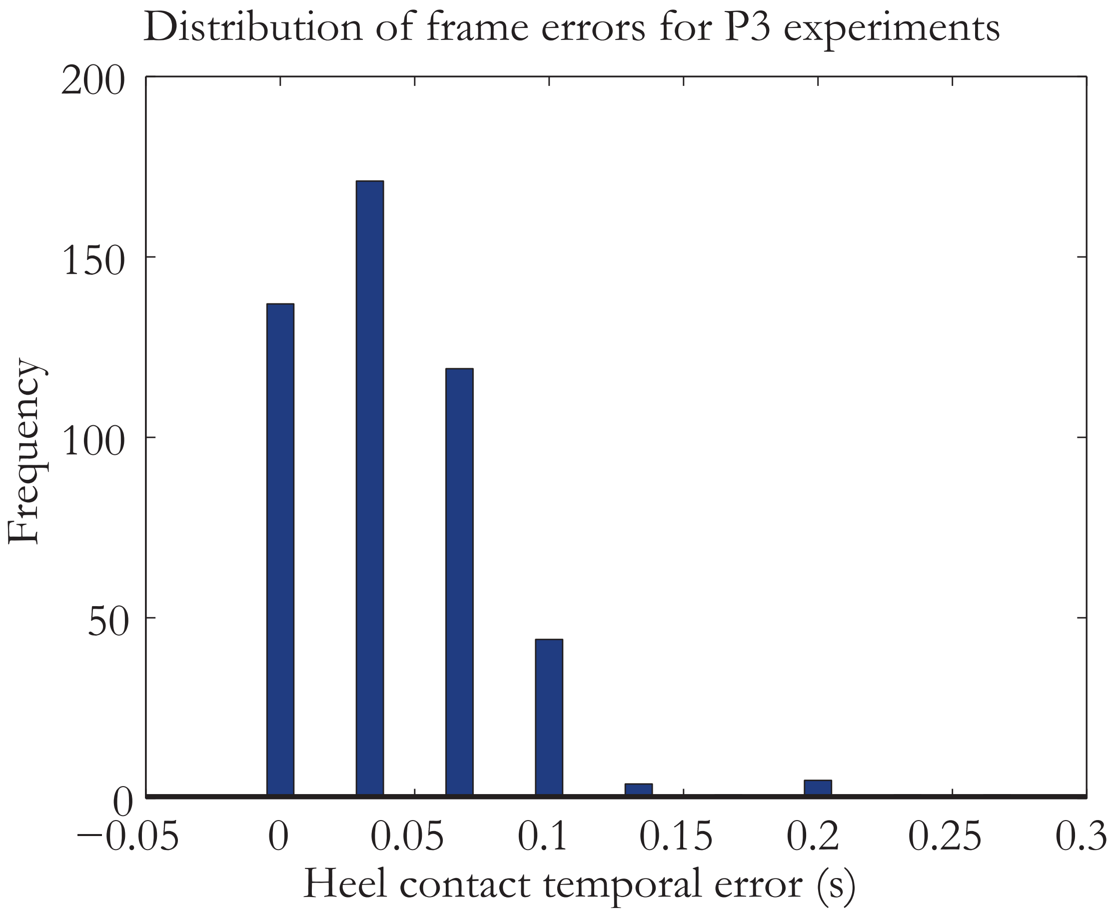

Figure 4–12: Temporal error of heel contact detection when compared to manual video analysis. Note that error is calculated in terms of number of video frames, hence the discrete nature of the results.

Table 4–3: P3 data set - comparison between calibrated and uncalibrated results

	Uncalibrated	Calibrated
Accuracy	60.7%	61.5%
ROC curve area	0.846	0.856

4.5 Discussion

Results from the preliminary experiments show that the angular velocity of the foot at initial contact is effective at distinguishing between good and bad steps. Both

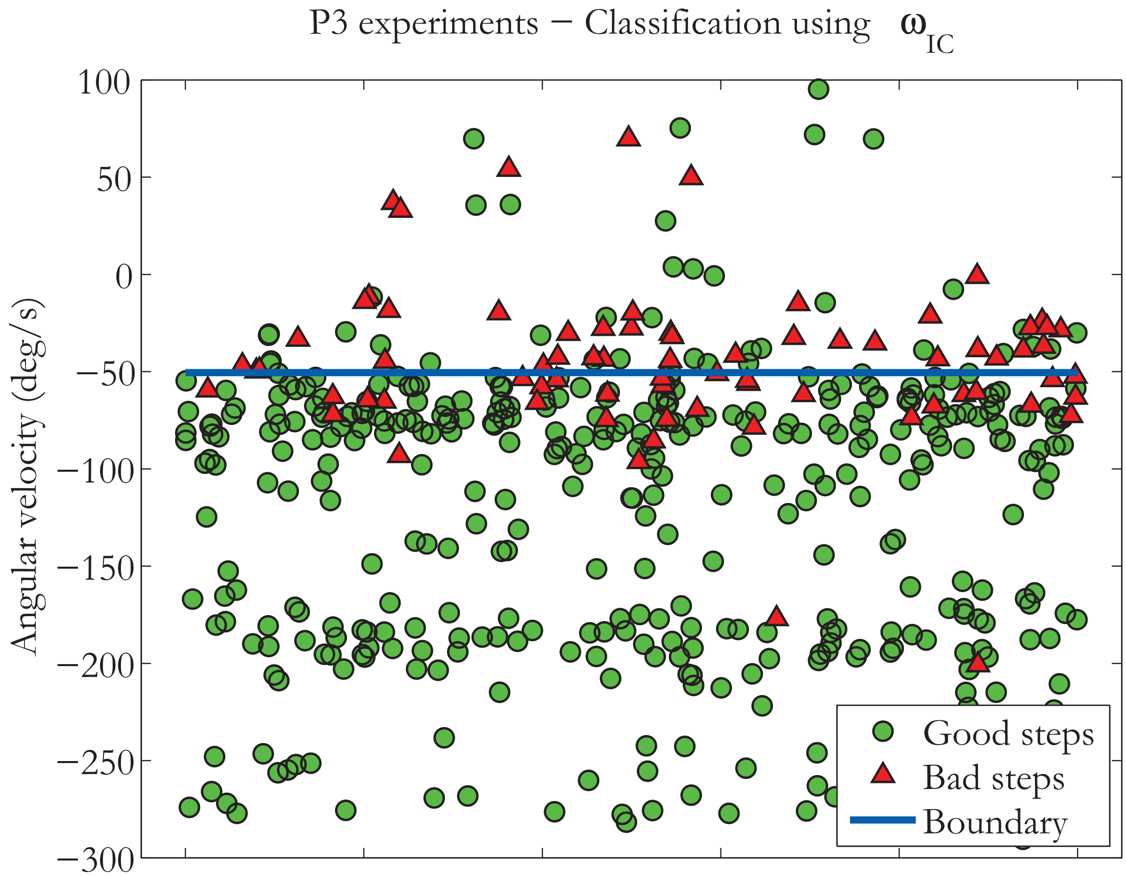

Figure 4–13: Scatter plot of individual steps in P3 experiments, coloured by gold standard rating. The classification boundary for optimal accuracy is also shown.

the P2 and P3 experiments showed reasonable classification results (P2 ROC curve area of 0.847, P3 ROC curve area of 0.856). In addition, the algorithm developed was successful at extracting this feature, with only 2 missed steps over the 877 total steps. However, there was a significant proportion of misclassified steps, with a P2 misclassification rate of 18.8% and a P3 misclassification rate of 38.4%. Some of these misclassified steps were due to algorithm error (note the temporal error distribution in Figure 4–12), however, this only accounts for 11% of steps.

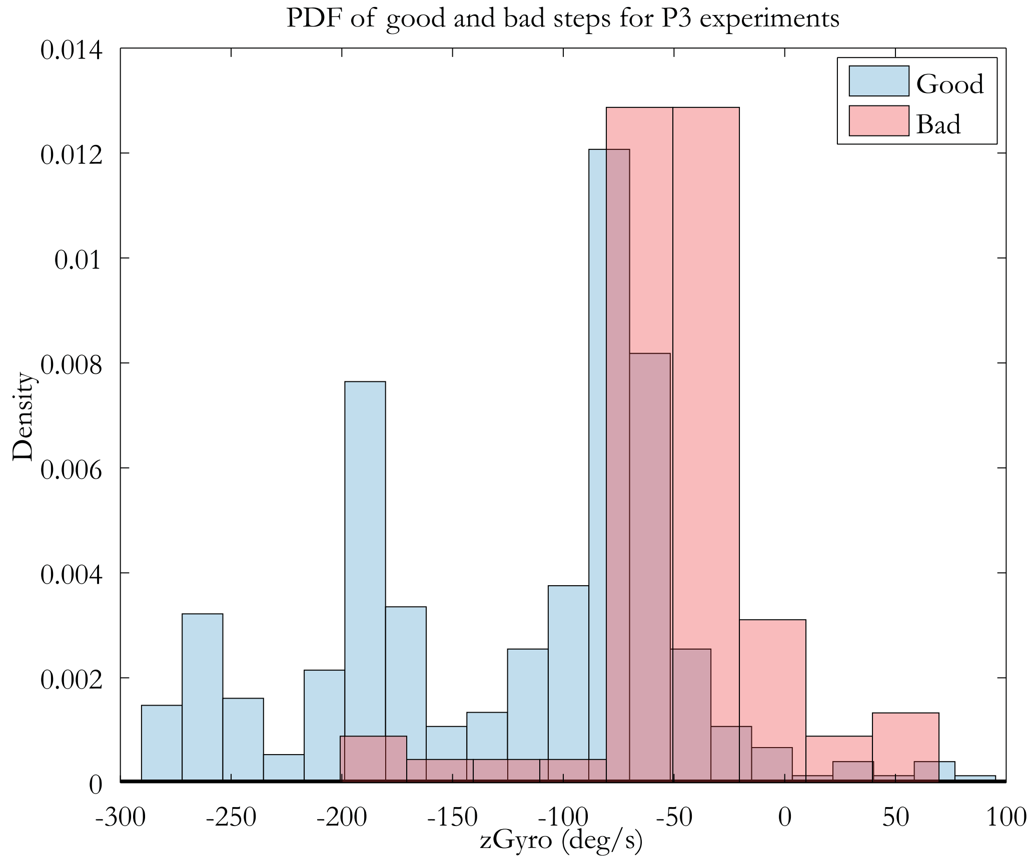

Figure 4–14: Probability density function for the P3 experiments. Each class of step is shown separately

One reason for the incorrect classifications could be physiotherapist error. Although the physiotherapist rating has been used as the gold standard, the rating can vary from rater to rater, and also from rating to rating for an individual rater. To investigate the intra-rater repeatability, the rating process for the P2 set was repeated, yielding a κ value of 70% using Cohen's κ as a statistical gauge [64]. This represents only moderate repeatability, which may explain the misclassified steps.

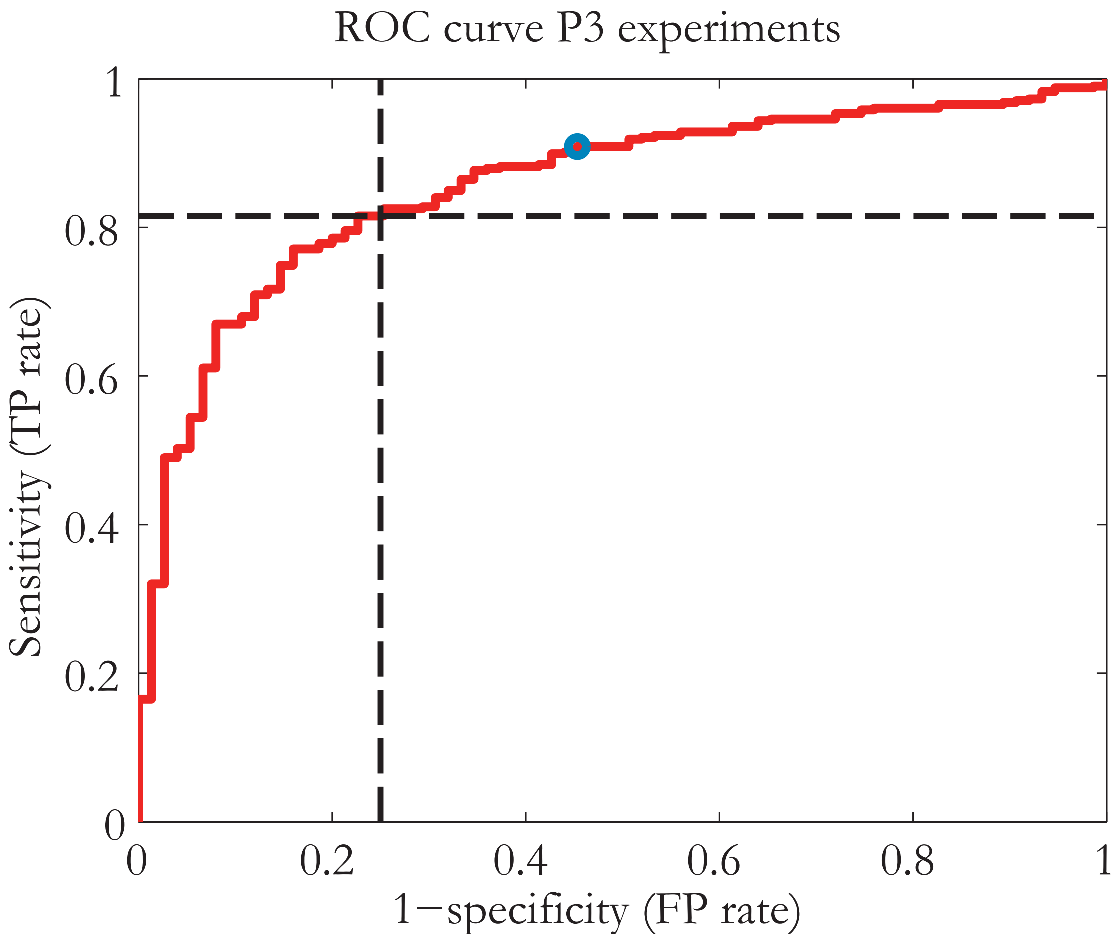

Figure 4–15: ROC curve for classification of P3 experiment data. The accuracy optimal operating point is highlighted in blue, and the cross-hairs show the sensitivity at 75% specificity.

The high proportion of misclassified steps, particularly for the P3 set, is concerning. It must be noted that the data gathered in these preliminary experiments was highly unbalanced towards good steps, with only 16% of the steps in P2 and P3 rated bad. In addition, the majority of the bad steps were gathered from mimicked abnormal gait. While this was useful for developing the algorithm, in order to validate the algorithm, true abnormal gait data gathered from the target population is necessary. This was the purpose of the clinical experiments, and the details are presented in the following chapter.

Regarding the calibration procedure, the simulation study presented in this chapter showed that calibrating the device for orientation does not affect the accuracy at the ideal placement (z axis parallel to the medial-lateral axis), while accuracy is significantly improved by calibration for a misaligned sensor. In particular, calibration makes the accuracy independent of rotation about the y axis, which is the most likely misalignment in practical use. Although for this project, care was taken to place the sensor accurately for all experiments, reducing the advantage of calibration, it was still worth implementing this feature for the clinical experiments since there is no cost in accuracy.

CHAPTER 5
Clinical pilot study

Chapter 4 described the development of the algorithm using data from preliminary experiments. Results demonstrated good classification performance, however, the experiments included only mimicked abnormal gait. This chapter presents a clinical study to evaluate the algorithm; experiments were conducted with the target population, i.e. elderly patients being treated for gait abnormalities by a physiotherapist. The main purpose of the clinical experiments was to assess the accuracy of the algorithm at classifying individual steps as good or bad.

5.1 Methods

The same Shimmer device used in the preliminary experiments was used as the sensing module for the clinical study. The PC which was used to record data in the preliminary experiments was replaced with a biofeedback module that implemented the algorithm in real time. The biofeedback module is described below:

5.1.1 Biofeedback module

An LG Nexus 4 smartphone was used for the biofeedback module. Its key specifications:

- Android 4.4.4 operating system (KitKat)
- 2GB RAM
- Qualcomm Snapdragon SF Pro 1.5GHz CPU
- True HD IPS Plus capacitive touchscreen

- 4.7” display
- Inbuilt loudspeaker
- Bluetooth capabilities
- Size 133.9 x 68.7 x 9.1 mm

An Android application was developed to communicate with the sensing module. The development process was simplified with the use of the Android driver supplied by the Shimmer manufacturer.

5.1.2 Android GUI

Raw 6 axis data was streamed to the smartphone via bluetooth, and the algorithm was applied to the raw data using the smartphone's CPU. An Android application was developed using the Java programming language with the Android Software Development Kit. Figure 5–1 shows a screenshot of the Android application.

The application performs the following functions:

1. Real-time classification of steps as good or bad - data is streamed point by point at 51.2 samples/second. A filtered version of the zGyro channel is plotted, and step markers are plotted in colour.
2. Biofeedback to user on step quality - the application optionally produces feedback in the form of an audible tone for every good step (positive biofeedback). Positive biofeedback was chosen as opposed to feedback for every bad step upon consultation with the supervising physiotherapist; in her experience positive biofeedback was more effective.
3. Calibration for sensor orientation (see Section 4.3 for details).

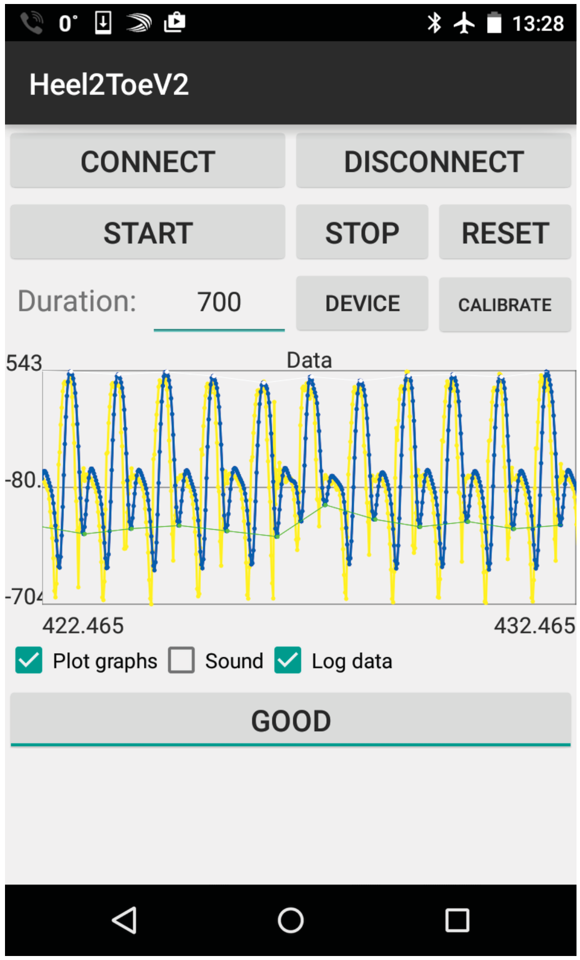

Figure 5–1: Screenshot of Android application for data recording and feedback

4. Data logging - Raw data was logged to a `.dat` file (a generic text-format file with tab separated values), and information about each step (time of step, quality of step, angular velocity of step at initial contact) was logged to a separate `.dat` file, alongside the calibration parameters.

Graph plotting, biofeedback, and data logging are all optional and can be toggled on or off.

5.1.3 Clinical experiments

The target population for this study comprised elderly persons attending a rehabilitation program or being seen by a physiotherapist. Subjects were identified as gait impaired by the treating physiotherapist at the Royal Victoria Hospital site of the McGill University Health Centre through the Geriatric Day Hospital. The inclusion requirements were:

- 65 years or older
- Medically stable
- Able to walk at least 15 metres with or without walking aids

Ethics approval was obtained from the Research Ethics Board. All subjects gave voluntary informed consent.

Data was gathered by attaching the Shimmer device to the subject's foot, performing the orientation calibration described in Section 4.3, and instructing the subject to walk up and down a 15 metre corridor. Raw data was logged to the smartphone, as well as information on step quality, completed by our algorithm in real time. Video of the trial was recorded by a camera mounted on a trolley that was manually pushed to determine physiotherapist ratings.

Table 5–1 shows the age, sex, and relevant medical conditions (where available) for each of the 12 subjects. Most subjects completed 6 trials, with each trial comprising a 'there and back' along the corridor. Some subjects failed to complete all six due to fatigue. Feedback from the smartphone (in the form of an audible tone for each good step) was turned on for the second half of the trials for each subject, however subjects were not informed about the significance of the feedback.

Table 5–1: Patient list and info for clinical experiments

Patient ID	Sex	Age	Medical conditions
01	F	74	
02	M	79	
03	F	66	
04	M	89	Macular degeneration Parietal lobe stroke
05	M	86	Parkinson's disease Mild dementia Macular degeneration
06	M	70	
07	M	74	Guillain-Barré syndrome Parkinson's disease
08	F	67	
09	M	91	Spinal stenosis Hip fracture
10	F	87	
11	M	87	Transient ischemic attack
12	M	89	

5.1.4 Analysis

Step labels were assigned by the physiotherapists using video analysis as described in Section 4.2. The classification results presented in the next section were determined using support vector machines as described in Section 3.3. All accuracy values reported were determined using 2-fold cross validation, as described in Section 3.3.5.

Several gait variables for each subject were recorded. These were:

- Number of steps (as determined by the algorithm)
- Good step percentage (as determined by the physiotherapist)
- Good step percentage (as determined by the algorithm)

- Average cadence
- Standard deviation of cadence
- Average ω_{IC}
- Standard deviation of ω_{IC}

In order to determine the average and standard deviation of cadence, periods where the subject was not walking had to be removed. This was done by selecting 'gait blocks' from the kinematic data. A gait block was defined as a contiguous segment of data with less than 3 seconds between adjacent detected initial contacts. Figure 5–2 shows a sample kinematic data record in which two gait blocks were determined. The time between consecutive initial contacts (in seconds) was used to calculate cadence for each step:

$$Cadence = \frac{120}{\Delta T}$$

The number of steps was equal to the number of detected initial contacts within the gait blocks. Values of ω_{IC} values were determined using the algorithm from Section 4.1.

5.2 Results

The results in this chapter are divided into three sections. The first section gives the physiotherapist ratings of all the steps for each subject. The second section concerns the accuracy of the algorithm for detecting initial contact. The last section presents the classification performance of the ω_{IC}.

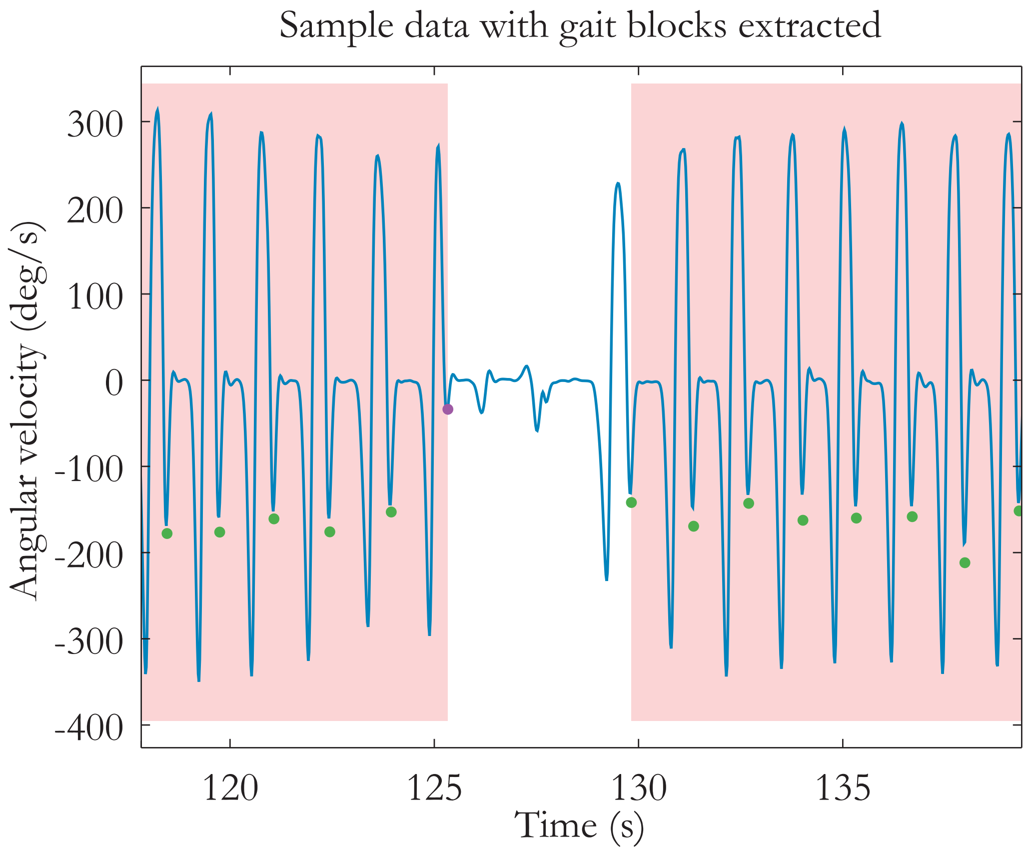

Figure 5–2: Sample zGyro data showing detected gait blocks in pink. Detected steps (initial contact) are shown as points, green steps are good, purple are bad.

5.2.1 Physiotherapist ratings

Table 5–2 shows the good/bad step split, as rated by the physiotherapist. The data set was quite well balanced, with only 58.5% of steps rated good.

Table 5–2: Breakdown of step quality over all trials, as categorized by the physiotherapist

Good steps	1271
Bad steps	903
Total	**2174**

Table 5–3 shows the physiotherapist ratings for each subject.

Table 5–3: Step quality as rated by physiotherapist for individual subjects

Subject ID	No. of good steps	No. of bad steps	Total steps	Good step percentage
01	191	30	221	86.4%
02	171	7	178	96.1%
03	168	24	192	87.5%
04	173	81	254	68.1%
05	147	26	173	85.0%
06	104	2	106	98.1%
07	97	23	120	80.8%
08	2	152	154	1.3%
09	1	160	161	0.6%
10	126	156	282	44.7%
11	0	242	242	0.0%
12	91	0	91	100.0%

5.2.2 Initial contact detection

The physiotherapist rated found only two steps that were not detected by the algorithm, yielding an algorithm detection success rate of 99.9%.

Although the algorithm detected almost every step, there were some cases where it failed to detect the minimum associated with initial contact accurately. Rather, it detected a minimum occurring during swing, as illustrated in Figure 5–3. This minimum was usually caused by the foot scraping the ground during mid-swing, or occasionally when the swinging leg clipped the stance leg. The result of this false minimum was always a value of ω_{IC} much higher than normal (in the 0 deg/s to 100 deg/s range), which caused the algorithm to classify the step as bad regardless of the quality of heel strike. These steps are visible as outliers in Figure 5–4. There were only 22 steps (1.0% of all steps in the clinical experiments) that missed the true minimum. This error in detecting heel contact appears to be a problem, but labelling these steps as bad is not a major issue since clipping the ground or the opposite leg during the swing phase is a negative feature, and attaching a 'bad' label to such steps is entirely appropriate. Additionally, only three of these steps were misclassified, indicating that the physiotherapist agrees with the 'bad' label for the majority of these missed initial contact steps.

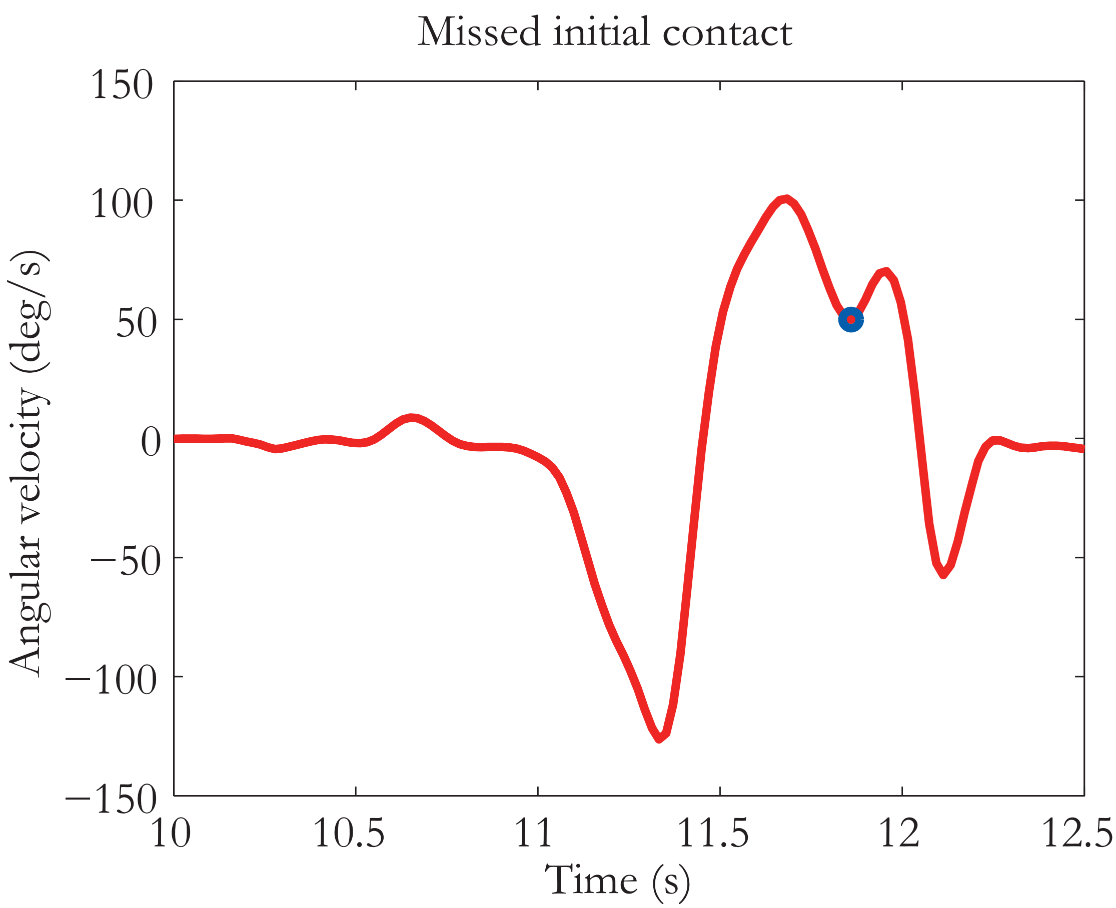

Figure 5–3: Example of missed initial contact. The blue point shows the false local minimum detected by the algorithm, which occurs during peak swing as opposed to initial contact.

In Chapter 4, an analysis of the temporal error between the detected initial contact and the actual initial contact from the preliminary data set was presented. Due to the much larger sample size of the clinical experiments, this analysis is not feasible for this data set.

5.2.3 Classification results

Figure 5–4 shows the distribution of ω_{IC} for each step, coloured by step quality for all the clinical data. The accuracy optimal boundary is shown in blue, determined through classification, as described in Section 3.3. The accuracy with this boundary was 85.5%. Figure 5–5 shows the associated probability density function, showing the overlap between the two classes. Figure 5–6 shows the corresponding ROC curve in Panel A. Panel B shows the boundary curve, which gives the threshold value for any given sensitivity/specificity operating point. This ROC curve demonstrates very good classifier performance, with an area under the curve of 0.956. The maximum accuracy is 85.5%, and the sensitivity at 75% specificity is 96.2%, both very good results.

Table 5–4: Performance metrics for classification of clinical experiment set

Sensitivity at 75% specificity	0.962
Area under the ROC curve	0.956
Maximum accuracy	85.8%

We examined the misclassified steps to understand the reasons for misclassification. Figures 5–7 and 5–8 show scatter plots of good and bad steps for each subject individually; Figures 5–9 and 5–10 show the associated frequency distribution functions. These figures demonstrate that the two classes are well separated for most

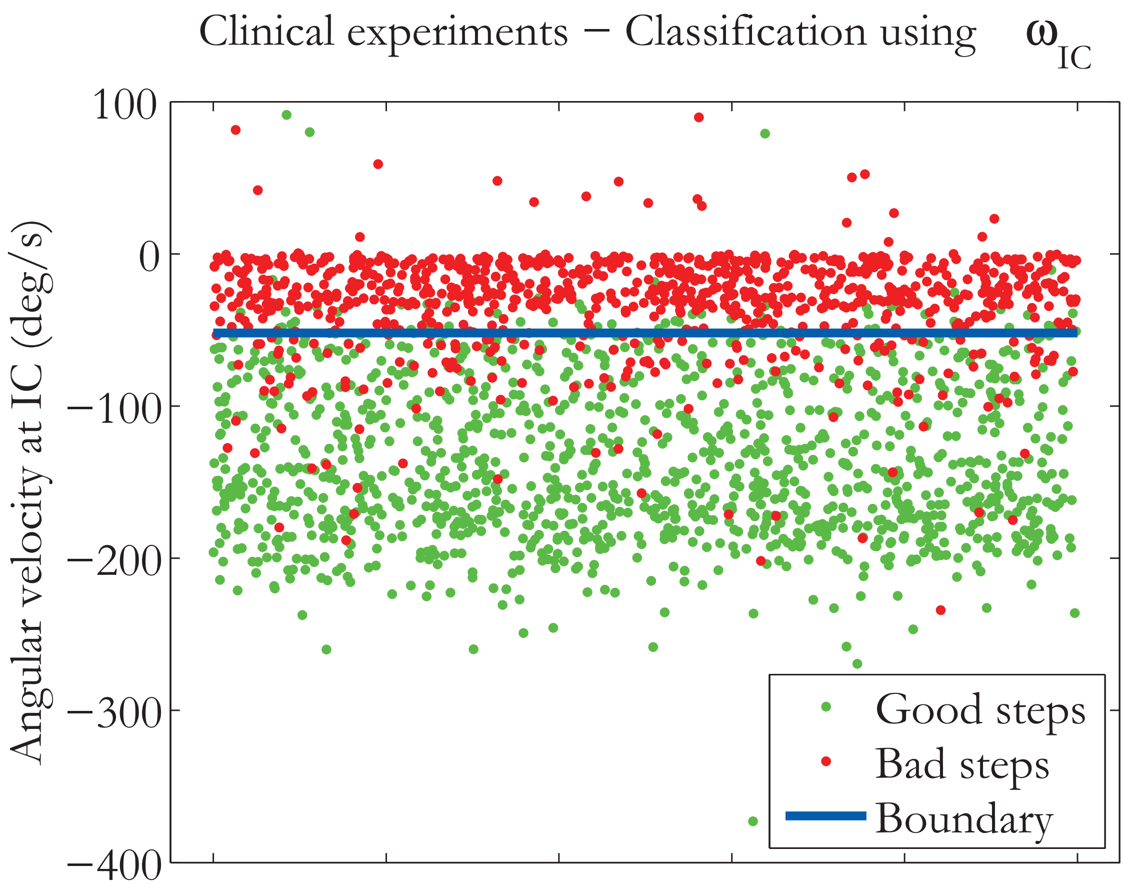

Figure 5–4: Scatter plot of individual steps for clinical experiments, coloured by physiotherapist rating. The classification boundary for optimal accuracy is also shown.

subjects. However, two subjects, 04 and 10, have distributions that overlap significantly, demonstrating that ω_{IC} is a poor feature for classifying steps for these two subjects.

One potential method to improve classification performance would be to use a unique boundary for each patient. To investigate this idea of a patient-specific boundary, steps were classified separately for each subject, with a distinct boundary determined for each, and a corresponding accuracy. The results of this analysis are

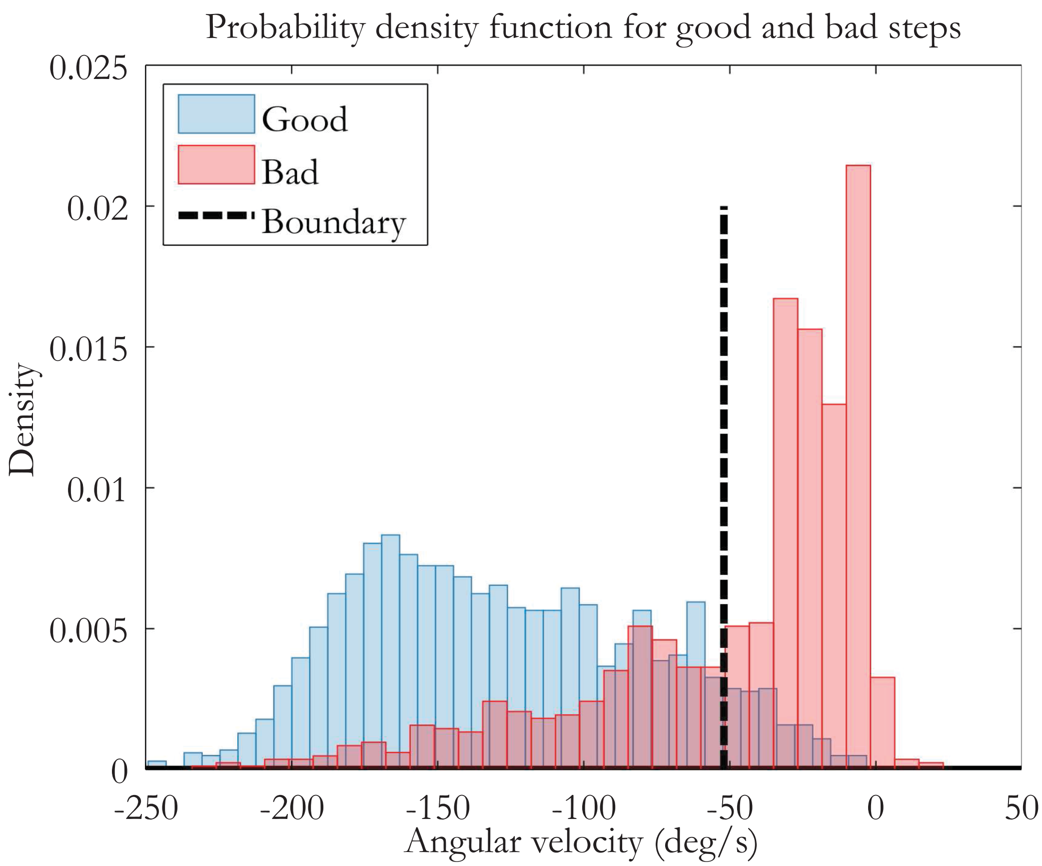

Figure 5–5: Probability density function for clinical experiments. Each class of step is shown separately, along with the boundary for optimal accuracy.

shown in Table 5–5. Note that in calculating the accuracy for these cases, cross validation could not be performed due to the small sample size and unbalanced class sets. Hence, for comparison purposes, the corresponding accuracy for the ensemble case is reported without cross validation as well.

Table 5–5 shows that the classification accuracy is slightly better overall when using a patient specific boundary, with a total accuracy of 91.6% compared with 89.8%. However, the fact that the accuracy is only marginally better suggests that

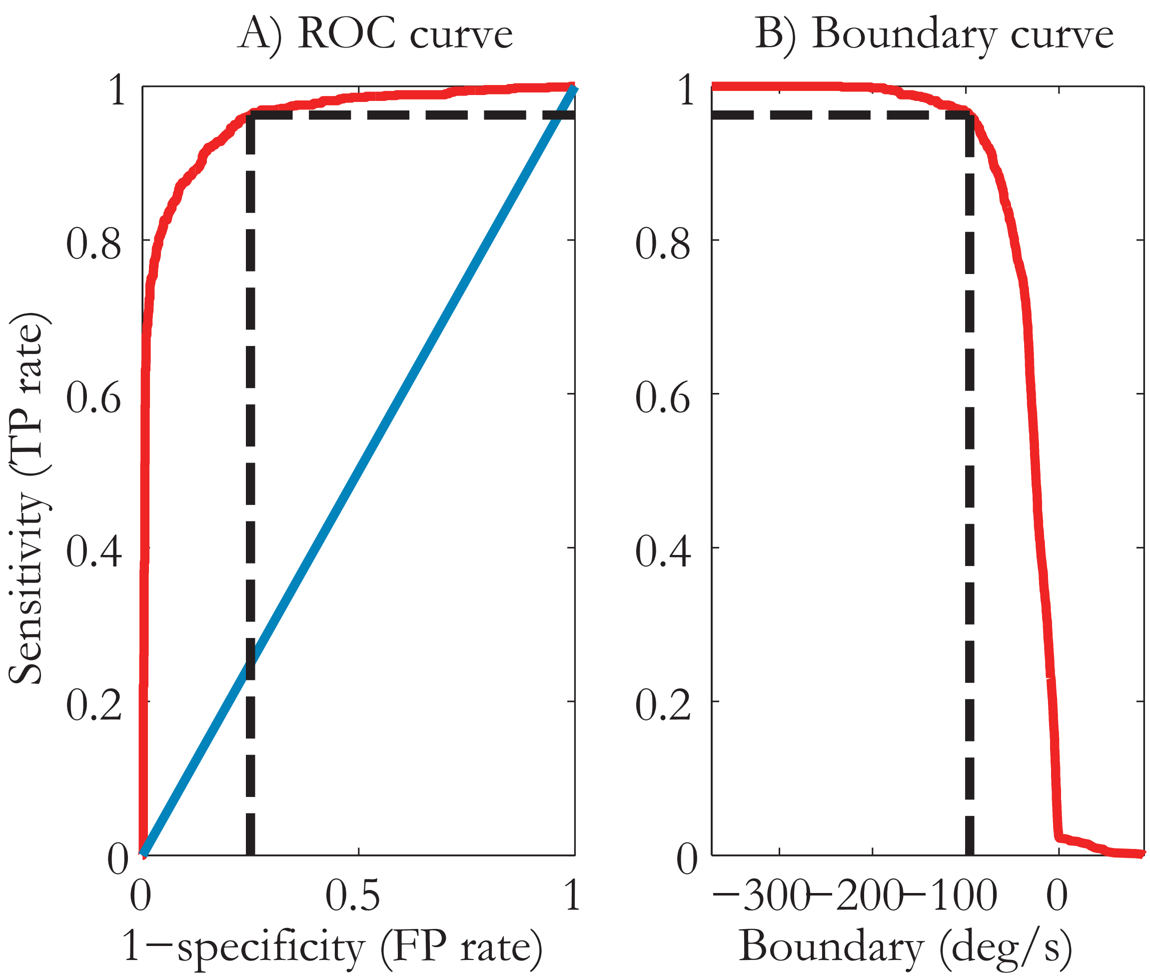

Figure 5–6: ROC curve for clinical experiments (red curve in panel A). The blue line shows the classification performance due to chance. The dashed black line shows the operating point for 75% specificity, with the associated boundary value on panel B.

using a patient specific boundary will not correct the majority of the misclassified steps. The subjects that benefited the most from a patient specific boundary were Subjects 01 and 07, with a 4.5% and 5% improvement respectively, which is relatively minor. The high overlap in the frequency distribution for Subjects 04 and 10 is evident here as they have the lowest accuracies (68.1% and 77.0% pooled accuracy

Table 5–5: Boundary and accuracy results for patient specific boundaries. The accuracy from the ensemble boundary (Pooled acc.), the accuracy from a patient specific boundary (Separate acc.), and the improvement that results from using a patient specific boundary (Acc. difference) are shown for each subject.

Patient ID	Boundary (deg/s)	Pooled acc.	Separate acc.	Acc. difference
01	-90.8	89.1%	93.7%	4.5%
02	-109.8	97.2%	98.9%	1.7%
03	-85.3	91.2%	92.7%	1.5%
04	-30.15	68.1%	71.7%	3.5%
05	-53.9	96.0%	96.5%	0.6%
06	-22.7	98.1%	99.1%	0.9%
07	-63.8	90.0%	95.0%	5.0%
08	-68.9	96.1%	98.7%	2.6%
09	-46.5	99.4%	99.4%	0.0%
10	-52.0	77.0%	77.0%	0.0%
11	-57.8	99.6%	100%	0.4%
12	-52.9	100.0%	100%	0.0%
Total	-52.0	89.8%	91.6%	1.8%

respectively). These accuracies were not improved much with a patient specific boundary, indicating that it is not the location of the boundary that is the issue for these subjects.

To investigate why subjects 04 and 10 differ from the rest of the subjects, a number of variables from each subject were examined (as described in Section 5.1.4). These are shown in Table 5–6.

None of the variables for Patients 04 and 10 recorded in Table 5–6 stand out as outliers. Figure 5–11 shows the relationship between cadence and angular velocity for each patient, together with their scaled standard deviations. This figure shows that the recorded measures for Patients 04 and 10 are in the same vicinity as the rest of the patients.

Table 5–6: Measured variables for all patients, averaged over entire experiment. Number of steps, good step percentage (algorithm), average and standard deviation of cadence, average and standard deviation of ω_{IC} were all determined using kinematic data. Good step percentage (PT) was the percentage of good steps as determined by the physiotherapist.

Patient ID	Number of steps	Good step % (PT)	Good step % (algorithm)	Avg. Cadence (steps/min)	Std. Cadence (steps/min)	Avg. ω_{IC} (deg/s)	Std. ω_{IC} (deg/s)
01	229	86.4	79.9	78.0	6.8	-131.5	45.8
02	198	96.1	85.4	89.2	12.0	-151.3	54.0
03	214	87.5	86.0	92.0	14.3	-163.8	61.9
04	296	68.1	3.38	93.3	9.0	-60.1	30.3
05	192	85.0	62.0	97.1	9.2	-116.4	55.5
06	140	98.1	88.6	85.6	11.9	-162.6	47.9
07	161	80.8	42.9	118.5	13.6	-98.6	35.8
08	264	1.3	0.0	86.9	8.5	-28.8	8.9
09	221	0.62	0.0	73.0	9.0	-20.0	11.7
10	420	44.7	0.2	85.2	7.3	-52.9	25.0
11	324	0.0	0.0	76.7	9.1	-4.0	13.5
12	158	100.0	82.9	91.1	14.8	-153.2	53.3

Figure 5–12 shows the relationship between the percentage of good steps and average ω_{IC}. Good step percentage follows a sigmoidal curve, as expected. Patients 04 and 10 lie in the middle of the sigmoid, at the steepest section. At this location, step quality is highly sensitive to changes in angular velocity. Small errors in angular velocity detection and physiotherapist rating errors will have an amplified effect. It is possible that performance is poor for these two subjects because the subjects' average ω_{IC} is very close to the boundary between good and bad steps.

5.3 Summary

Results from experiments conducted on the target population show that the algorithm for feature extraction can accurately detect initial contact in real time, and the angular velocity at initial contact is capable of distinguishing between heel first and flat foot contact with high classification accuracy. Classification performance is very strong for 10 out of 12 subjects. Some accuracy is lost for patients 04 and 10, who have an average angular velocity close to the boundary between the two classes. This is to be expected because small errors in calculating angular velocity in this region can easily result in an incorrect classification. Using a patient specific boundary does not improve classification accuracy by much (1.8% improvement overall). While it is possible that the inclusion of additional features may improve classification performance, the current performance is more than satisfactory. And ROC curve area of 0.956 overall suggests that this algorithm is suitable for clinical use.

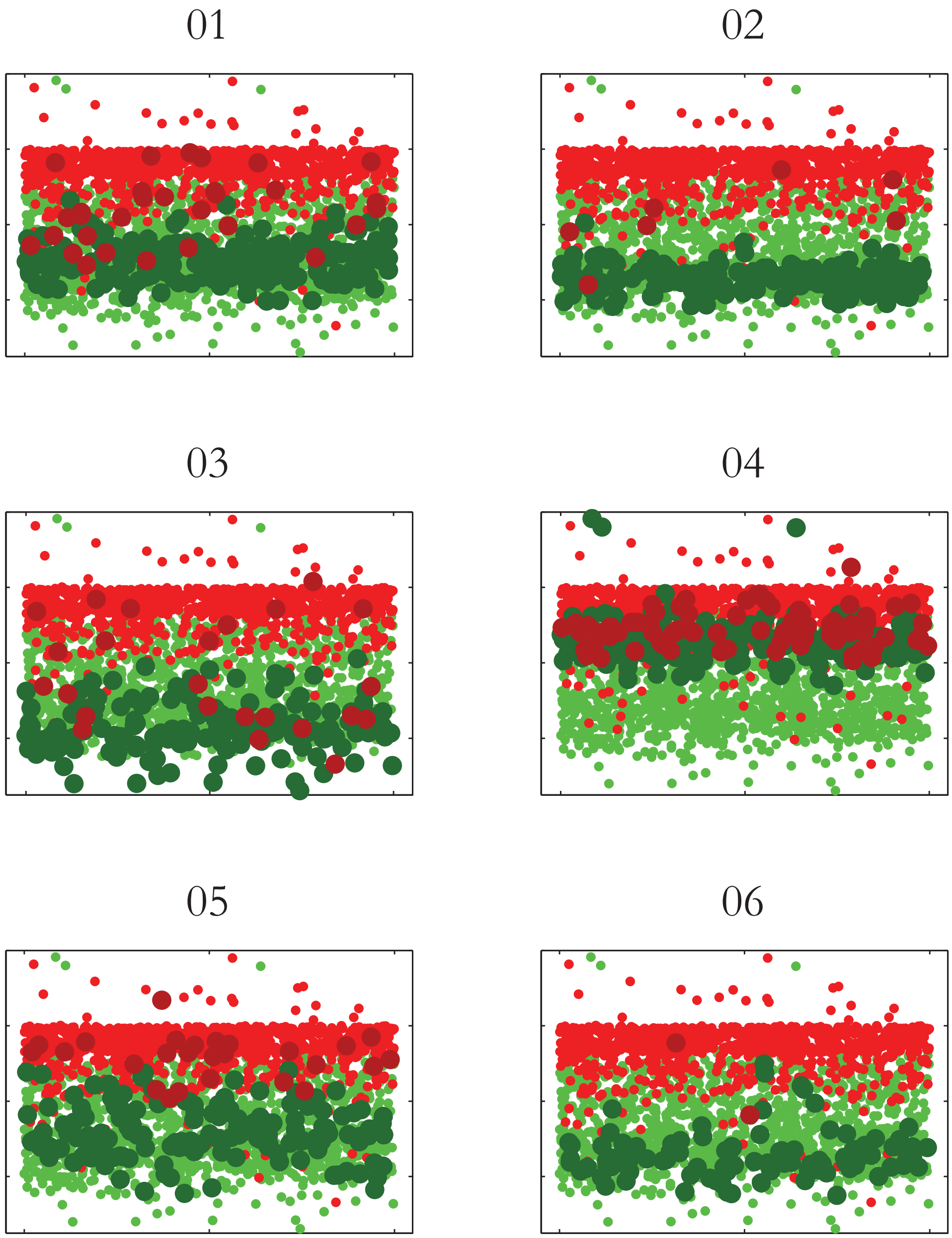

Figure 5–7: Distribution of steps, coloured by step quality (good steps are green, bad steps are red). The y axis shows ω_{IC}. The smaller, lighter points show the ensemble, while the darker points are steps from individual subjects. Each panel shows a different subject (subject ID given in title, subjects 01 to 06).

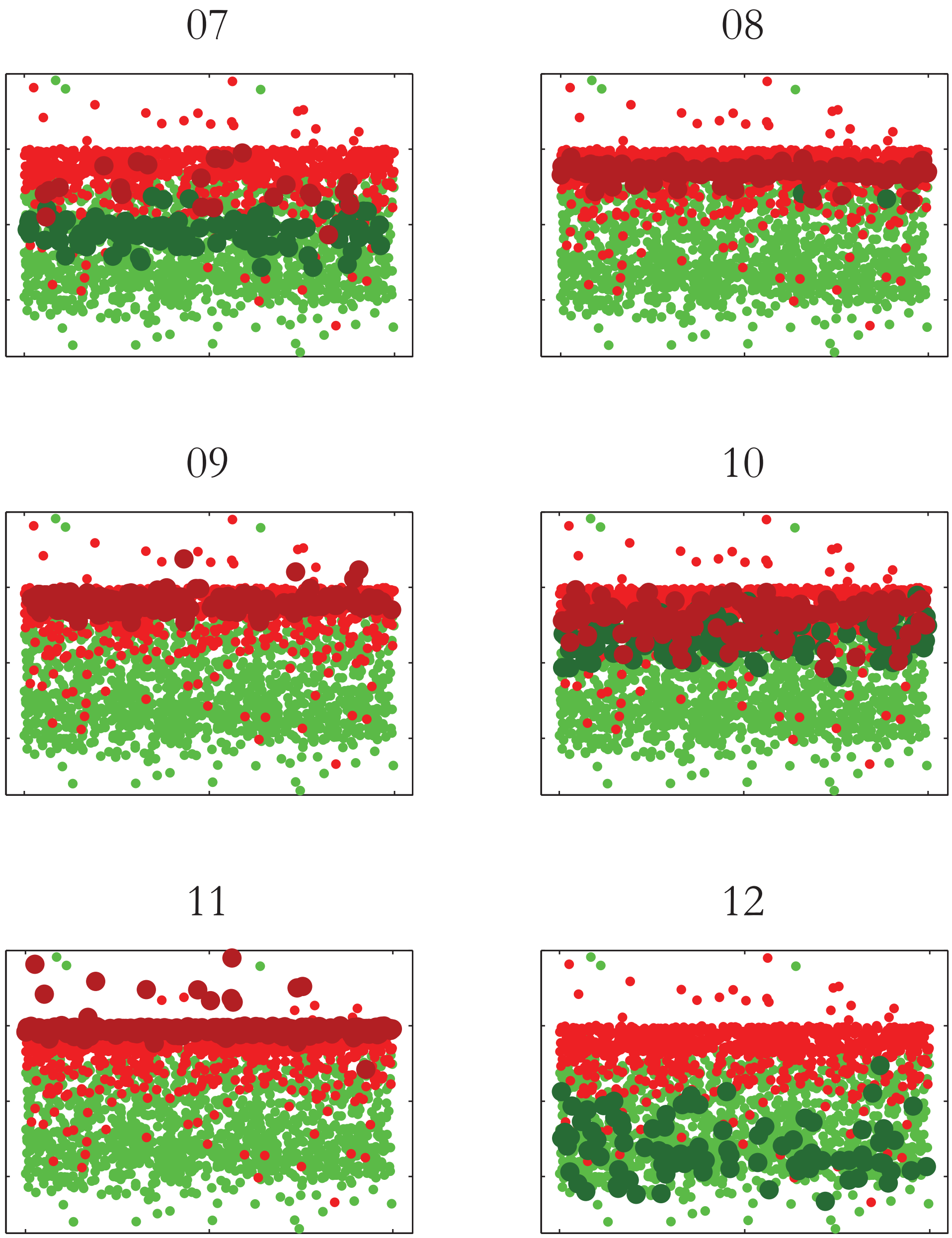

Figure 5–8: Distribution of steps, coloured by step quality (good steps are green, bad steps are red). The y axis shows ω_{IC}. The smaller, lighter points show the ensemble, while the darker points are steps from individual subjects. Each panel shows a different subject (subject ID given in title, subjects 07 to 12).

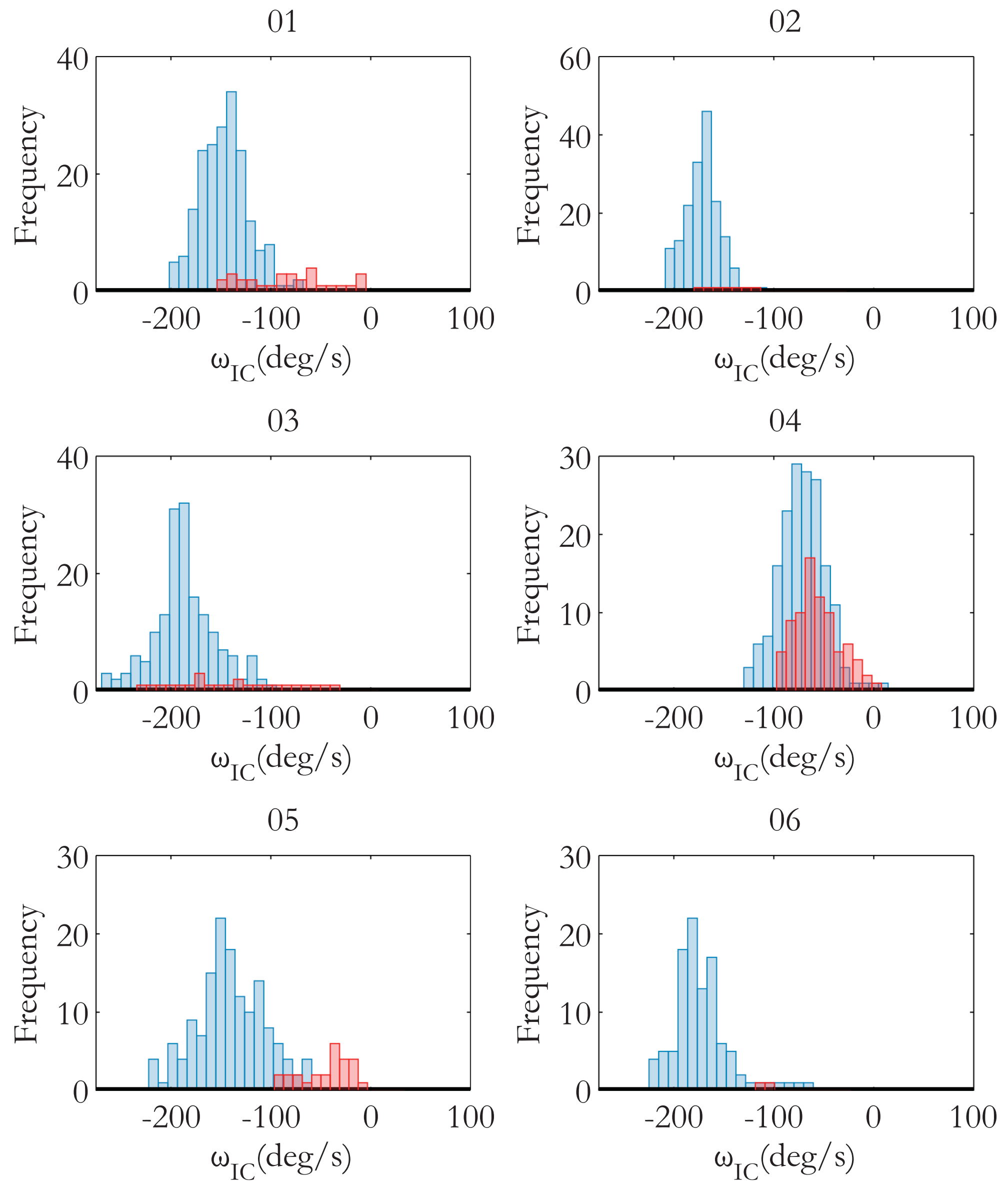

Figure 5–9: Probability frequency distribution of ω_{IC} for individual subjects 01 to 06. Each class of step is shown separately, good steps in blue and bad steps in red.

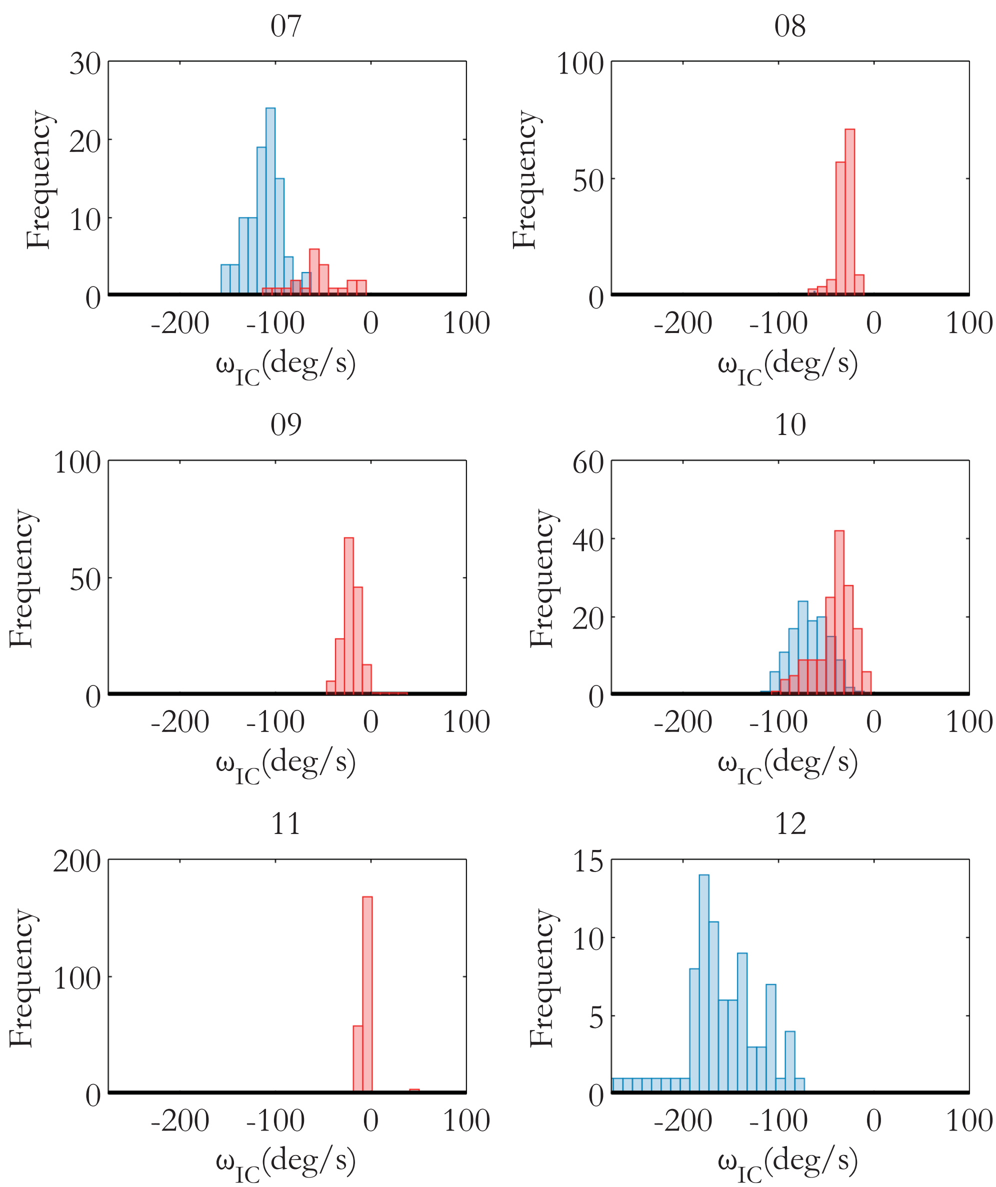

Figure 5–10: Probability frequency distribution of ω_{IC} for individual subjects 07 to 12. Each class of step is shown separately, good steps in blue and bad steps in red.

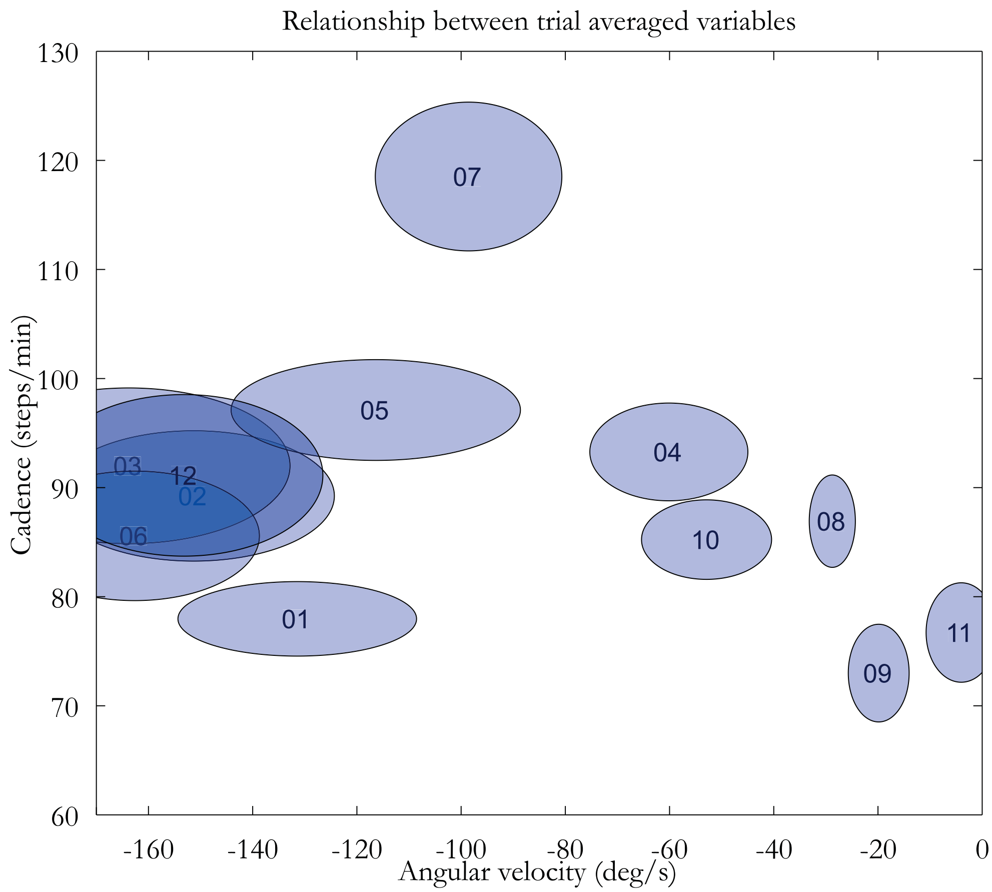

Figure 5–11: Scatter plot of ω_{IC} against cadence, showing the average for each subject. Shaded ellipses show half a standard deviation in each direction.

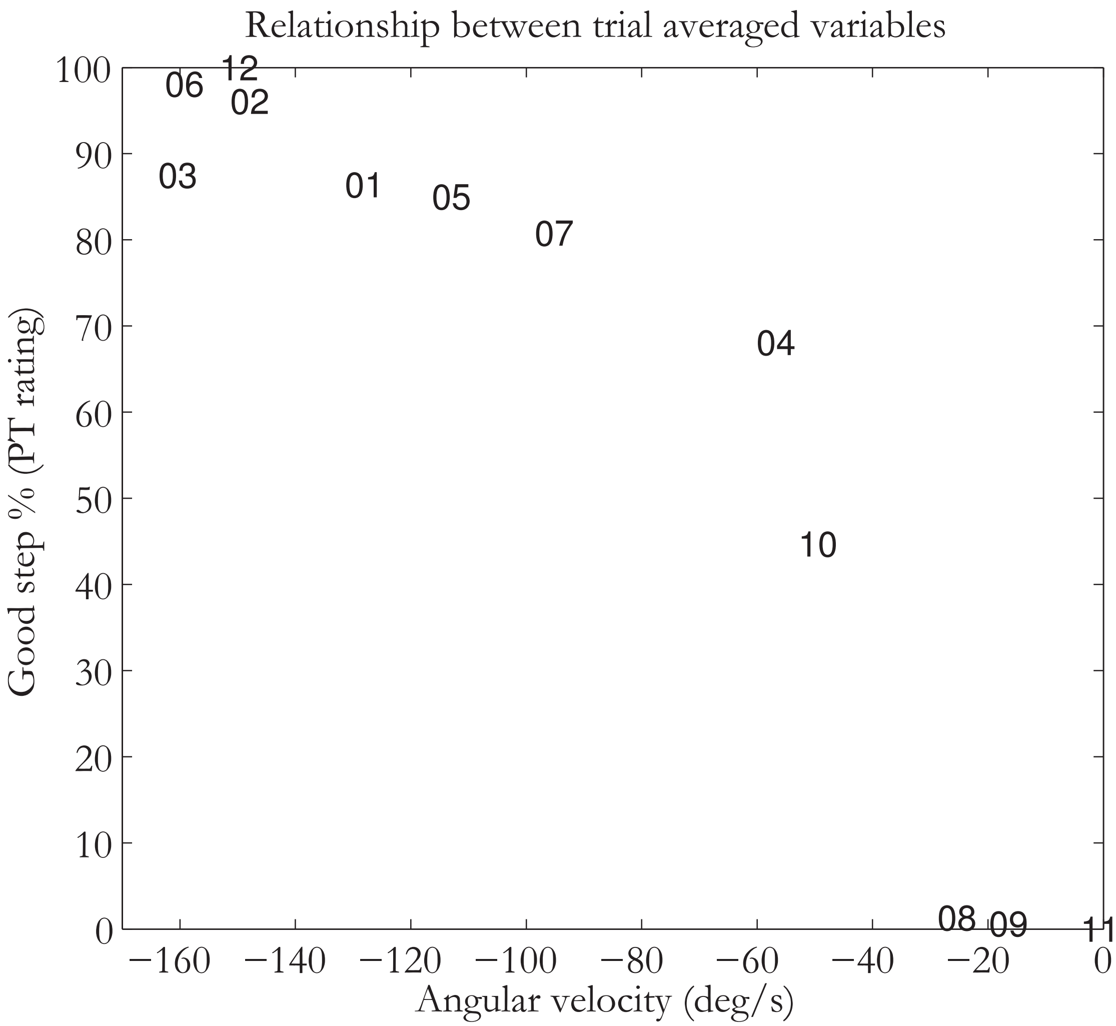

Figure 5–12: Scatter plot showing percentage of good steps (as rated by the physiotherapist) as a function of averaged angular velocity. Each point is a separate patient.

CHAPTER 6
Discussion

The results in Chapters 4 and 5 show that the major outcome of this project is a novel, real-time algorithm that accurately distinguishes between heel-to-toe and shuffling gaits on a step-by-step basis using angular velocity of the foot. Many previous studies have used wearable sensors for gait monitoring, however, most have relied on time-frequency analyses to detect individual steps, rendering their algorithms unsuitable for real-time use. Previous studies which have presented methods to detect individual steps have only been tested on healthy gait, while studies dealing with classification of abnormal gait have classified subjects as a whole, not on an individual step by step basis. Only a few studies have incorporated real time feedback, and these have been focused on minimizing trunk sway or limiting knee motion. Feedback for heel-to-toe versus shuffling gait has not been studied. Overall, the combination of the following features makes the algorithm presented in this thesis novel:

- Detection of individual steps
- Classification of individual steps as good or bad
- Feedback to user in real time
- Strong classification results for a range of abnormal gaits
- Cheap and portable hardware
- Incorporation of the algorithm into a smartphone application

The angular velocity of the foot at initial contact distinguishes between good and bad steps with high classification accuracy for the clinical dataset, with a maximum accuracy of 85.5%, and a ROC curve area of 0.956. The classification results demonstrate very strong classification performance. This is despite subjective physiotherapist labels with moderate repeatability - as stated in Chapter 4, the P2 physiotherapist ratings were repeated to establish intra-rater repeatability, and a κ value of 70% was found. This indicates that the ratings were not consistent, making classification more difficult. The fact that classification results were very good for 10 out of 12 subjects despite the low repeatability of the gold standard is very promising.

In addition, the two classes have an artificial class boundary, in the sense that there are no distinct clusters among the steps and each step falls along a continuum of step quality. This also makes classification more difficult, since there is no space between the two classes. The overall performance suggests that the algorithm and classification feature is suitable for rehabilitation purposes.

To implement this algorithm for clinical use, a suitable boundary must be chosen. In Chapter 5, the main performance metric reported was maximum accuracy. This is the accuracy at the accuracy optimal boundary, and is useful for measuring classifier performance, along with ROC curve area. However, a different boundary may be more suitable for rehabilitation purposes. Step quality is actually a continuous variable and not a discrete good/bad rating. The angle of the foot at initial contact is used by physiotherapists as a measure of step quality. The larger the angle the foot makes with the ground, the better the quality of step. Therefore, marginally good steps, i.e. steps which are close to the boundary, still have room for improvement.

Thus, to improve gait in the long term it may be more useful to rate these marginally good steps as bad. With positive feedback implemented, i.e. a beep for every good step, a beep would reinforce a walking pattern, while the absence of a beep will effect a change towards a more defined heel strike. The aim of the device is not to simply increase the percentage of good steps, but to improve heel-to-toe gait overall, i.e. increase the angle at heel strike. This is best achieved by enforcing a stricter good/bad boundary to trigger improvement for a greater proportion of steps. Choosing an operating point weighted towards sensitivity ensures that most bad steps are classified correctly. The trade-off is that some good steps will also be rated as bad, but since these are likely to be marginal steps anyway, this is preferable. The supervising physiotherapist for this project suggested that a specificity of 75% would be acceptable to maximise sensitivity. The sensitivity at this specificity, as reported in Chapter 5, was 96.2%. The boundary associated with this 75% specificity, 96.2% sensitivity was -95.9 deg/s, and would be appropriate for clinical use.

The clinical results were significantly better than the preliminary experiments, which had accuracy ratings of 81.2% and 61.5% for P2 and P3 respectively. The low accuracy with the P3 data was a particular concern; indeed examining the P3 dataset alone would lead to the conclusion that the chosen feature was poor. However, the fact that the results with the clinical data were much better alleviates those concerns. There were two main differences between the preliminary experiments and the clinical experiments: 1) subjects walked on a treadmill in the preliminary experiments and overground in the clinical experiments, and 2) bad steps were mimicked in the preliminary experiments. The kinematic differences between overground and treadmill walking are minor [65]. Hence, the most likely cause of the discrepancy is the mimicked abnormal gait in the preliminary experiments. While the subjects who walked with a mimicked shuffle were trained physiotherapists familiar with abnormal gait, it is possible that their gait did not accurately represent the target population.

Figure 6–1 shows the average angular velocity of the foot for mimicked abnormal gait (P3) and true abnormal gait (clinical). The mean and standard deviations were calculated only from steps that were rated bad by the physiotherapist. This figure shows some key differences:

- Mimicked abnormal gait had a more negative angular velocity at initial contact (0% gait cycle)
- Angular velocity of the foot at toe off was more negative for mimicked abnormal gait ($\approx$ 60% gait cycle)
- Peak swing angular velocity was higher for mimicked gait (80% gait cycle)

The first point in particular had an effect on the algorithm. This could explain the poor accuracy for the P3 experiments, since the bad steps had a more negative ω_{IC} than what should be expected from the target population, blurring the boundary between good and bad steps.

6.1 Classification improvements

The majority of the misclassified steps belonged to Subjects 04 and 10, with accuracies of 68.1% and 77.0% respectively. It is clear that ω_{IC} was not a suitable

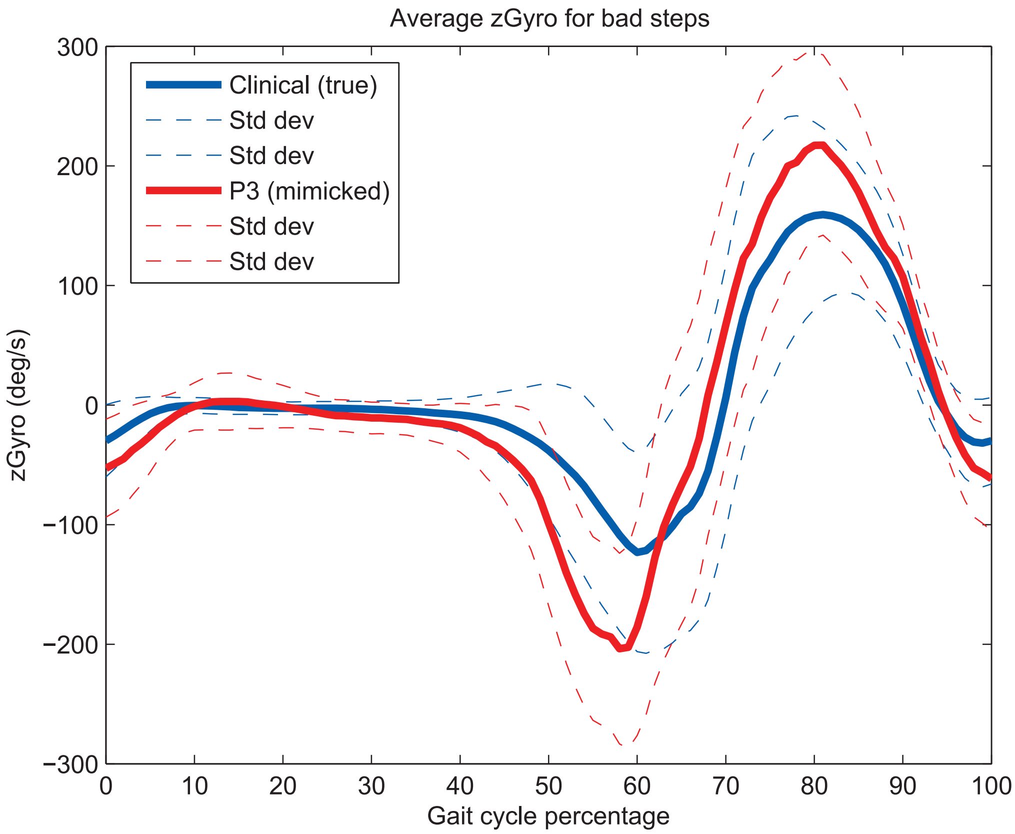

Figure 6–1: Average foot trajectory for bad steps - comparing true bad steps to mimicked bad steps. Blue curve shows the mean angular velocity of the foot for clinical subjects, red curve shows the mean angular velocity of the foot for mimicked bad steps in the P3 data set. Dashed lines show one standard deviation from the mean.

feature for these subjects, as there was significant overlap in the frequency distributions of good and bad steps (Figures 5–9 and 5–10). Using a patient specific boundary does not help for these subjects, since the main issue is the overlap between the two classes, not the location of the boundary. The accuracy improvement using a patient specific boundary was only 1.8% overall.

Table 5–6 was compiled in an attempt to find additional distinguishing features. The values produced for Subjects 04 and 10 did not stand out for all these parameters, suggesting that these parameters will not help as distinguishing features. Figure 6–2 shows the average angular velocity over all steps for each subject, normalised from 0% to 100% of the gait cycle. The angular velocity over the entire gait cycle for each subject was compared to see if there were any characteristics of the gait cycle that may have made classification difficult. However, the curves for Subjects 04 and 10 were similar to the rest of the subjects.

It was noted in Section 5.2.2 that for some steps the algorithm did not detect initial contact correctly. Of the 22 steps which had an erroneous initial contact detection, only 3 of these belonged to Subject 04, and Subject 10 had zero. Therefore we can conclude that detection of initial contact was not the issue for these subjects.

A reason for the failure of ω_{IC} for classifying Subjects 04 and 10 has not been found yet. The average ω_{IC} for these subjects was very close to the boundary, so it is possible that their steps were marginal and difficult for the physiotherapist to rate.

The classifier used in this work is support vector machines, as described in Section 3.3. Generally, different classifiers produce different results depending on the dimension of the feature set and number of classes. Support vector machines was initially chosen due to its strong performance in binary classification, and ability for expansion to a higher dimension feature set. If additional features were incorporated, it would have been pertinent to evaluate the performance of other classifiers to see if the performance could be improved. However, when only one feature is used, all

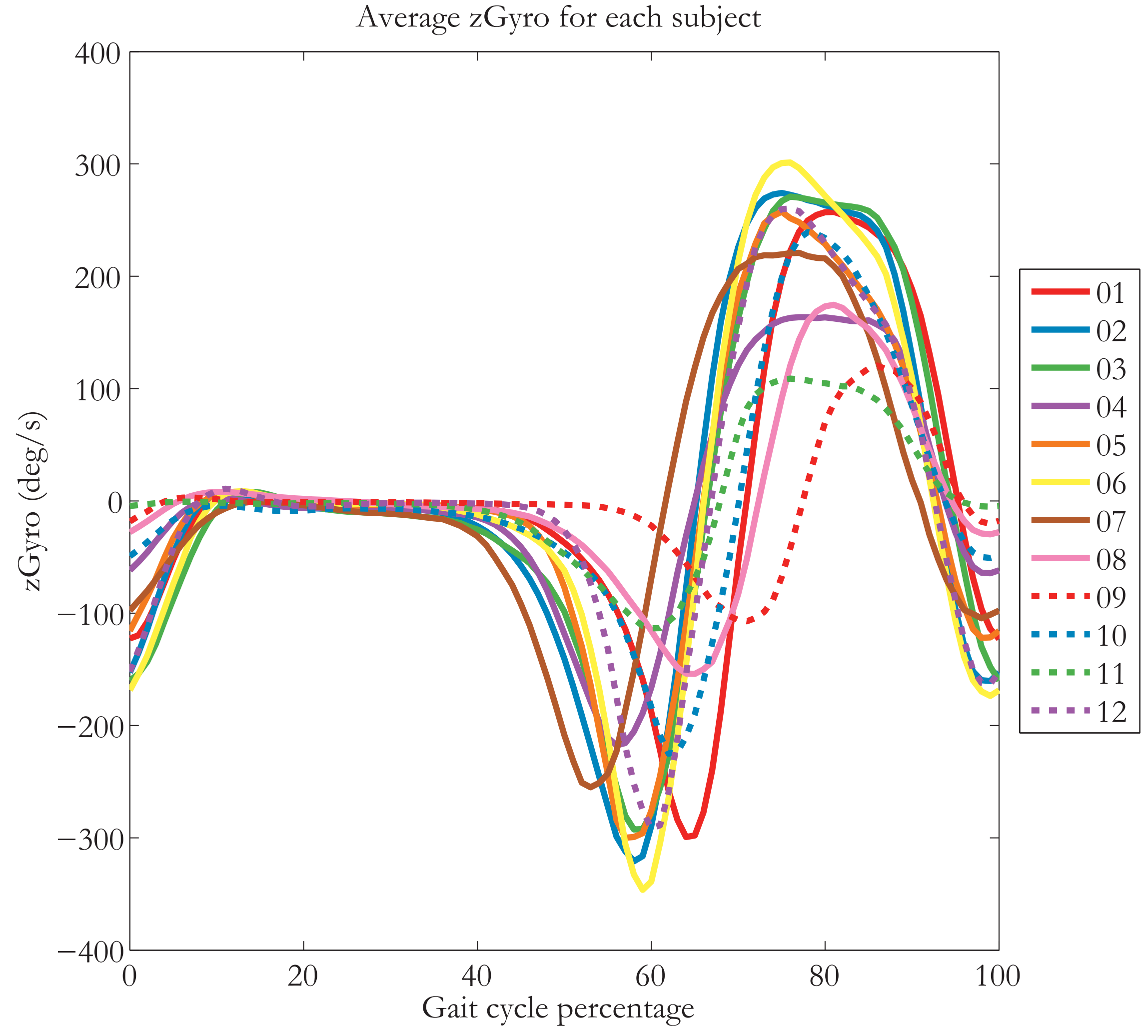

Figure 6–2: Average angular velocity gait profile for each subject for bad steps. Angular velocity for each individual step is first normalised in the time domain, and then averaged.

classifiers are identical, so in this case no improvement can be gained by investigating other classifiers.

6.2 Future work

While improvements to the classification performance may be possible with changes to the feature extraction algorithm or the addition of different features for classification, we have shown that the current algorithm performs very well at distinguishing between the two classes of steps. This is the 'reliability' phase of the device development, and shows the device is capable of evaluating step quality.

The next phase of device development is the 'validity' phase, which relates to whether the device, with biofeedback activated, can improve gait in the target population, both in the short term and long term. Short term validity relates to gait improvement over the course of three therapy sessions, while long term validity relates to permanent gait improvement. Analysis of the short term effects will be determined by assessing how gait parameters change with feedback on versus feedback off, with the patient instructed about the significance of feedback (i.e. instructed to maximise the number of good steps by listening for the feedback tone). The gait parameters will be both qualitative and quantitative, including, but not limited to, overall step quality (device rating), gait speed, and cadence. The distribution of ω_{IC} will also be compared. Gait improvements will be analysed further over consecutive visits to the clinic. Experiments on the short term validity phase are currently under way. The hypothesis for the validity phase is that incorporating device feedback together with information about what constitutes a good step will increase the proportion of good steps compared with no feedback.

Regarding the long term validity, subjects should be instructed to operate the device regularly in a home environment. Gait parameters can be logged over these sessions, and gait quality evaluated by a physiotherapist before this intervention and afterwards. As well as recording gait improvement, this experiment would be an ideal platform for a usability study, where patients and clinicians can be surveyed about the sensor ergonomics and interface intuitiveness. This experiment has not yet been designed.

As stated in Section 5.1.2, positive biofeedback has been used to relay the step quality to the patient, i.e. an audible tone for every good step, but no feedback for bad steps. This was chosen under the theory that patients will respond better to encouragement than reproach. However, there is no evidence to support this claim in this case, so it may be worth investigating which approach is more effective (positive biofeedback versus negative biofeedback).

Currently, the smartphone application was designed for operation by a trained clinician, and as such, the user interface is more technical than necessary with superfluous settings and graphing features. Since the ultimate aim is for home use by the patients directly, an interface redesign is necessary to simplify operation. In addition, extra features may be added upon consultation with clinicians, such as an adjustable difficulty scale to customise the classification on a patient by patient case.

6.3 Conclusion

The research presented in this thesis shows that using a single inertial measurement unit to distinguish between heel-to-toe gait and shuffling gait is feasible, with very good classification results (maximum accuracy of 85.8%, 96.2% sensitivity at nominal 75% specificity, and area under the ROC curve of 0.956). The feature extraction and classification algorithm can run in real time on a smart phone and provide feedback on individual step quality to the user. A user interface in the form of an Android application was developed to provide feedback (in the form of an audible tone) and log data for clinicians. The ultimate goal of this project is to develop this device for gait rehabilitation,, and with the accuracy of the device at detecting step quality now established, the next major step will be to determine the effectiveness of the device as a tool for rehabilitation.

References

[1] D. A. Winter, *Biomechanics and motor control of human gait: normal, elderly and pathological*, 1991.

[2] V. T. Inman, H. J. Ralston, and F. Todd, *Human walking*. Williams & Wilkins, 1981.

[3] Association, Rancho Los Amigos Medical Center. Professional Staff and Service, Rancho Los Amigos Medical Center. Pathokinesiology and Dept, Rancho Los Amigos Medical Center. Physical Therapy, *Observational Gait Analysis Handbook*. Professional Staff Association, Rancho Los Amigos Medical Center, 1989.

[4] D. A. Winter, *Biomechanics and motor control of human movement*. John Wiley & Sons, 2009.

[5] C. Kirtley. (2014) Clinical gait analysis. [Online]. Available: http://www.clinicalgaitanalysis.com

[6] M. W. Whittle, "Clinical gait analysis: A review," *Human Movement Science*, vol. 15, no. 3, pp. 369–387, 1996.

[7] E.-J. Marey, *Le mouvement*. Masson, 1894.

[8] R. G. Aptekar, F. Ford, and E. E. Bleck, "Light patterns as a means of assessing and recording gait. i: Methods and results in normal children," *Developmental Medicine & Child Neurology*, vol. 18, no. 1, pp. 31–36, 1976.

[9] D. Grieve, "The assessment of gait." *Physiotherapy*, vol. 55, no. 11, p. 452, 1969.

[10] D. A. Winter, A. O. Quanbury, D. A. Hobson, H. G. Sidwall, G. Reimer, B. G. Trenholm, T. Steinke, and H. Shlosser, "Kinematics of normal locomotion—a statistical study based on tv data," *Journal of biomechanics*, vol. 7, no. 6, pp. 479–486, 1974.

[11] A. Schmid, P. W. Duncan, S. Studenski, S. M. Lai, L. Richards, S. Perera, and S. S. Wu, "Improvements in speed-based gait classifications are meaningful," *Stroke*, vol. 38, no. 7, pp. 2096–2100, 2007.

[12] D. A. Winter, A. E. Patla, J. S. Frank, and S. E. Walt, "Biomechanical walking pattern changes in the fit and healthy elderly," *Physical therapy*, vol. 70, no. 6, pp. 340–347, 1990.

[13] M. P. Murray, R. C. Kory, and B. H. Clarkson, "Walking patterns in healthy old men," *Journal of Gerontology*, vol. 24, no. 2, pp. 169–178, 1969.

[14] B. Gillis, K. Gilroy, H. Lawley, L. Mott, and J. Wall, "Slow walking speeds in healthy young and elderly females," *Physiotherapy Canada*, vol. 38, no. 6, pp. 350–352, 1986.

[15] L. W. Lamoreux, "Kinematic measurements in the study of human walking." *Bulletin of prosthetics research*, vol. 10, no. 15, pp. 3–84, 1970.

[16] W. Weber and E. F. Weber, *Mechanik der menschlichen Gehwerkzeuge: eine anatomisch-physiologische Untersuchung*. Dietrich, 1836, vol. 1.

[17] J. Verghese, A. LeValley, C. B. Hall, M. J. Katz, A. F. Ambrose, and R. B. Lipton, "Epidemiology of gait disorders in community-residing older adults," *J Am Geriatr Soc*, vol. 54, no. 2, pp. 255–261, 2006.

[18] B. Salzman, "Gait and balance disorders in older adults," *Am Fam Physician*, vol. 82, no. 1, pp. 61–68, 2010.

[19] J. Hough, M. McHenry, and L. Kammer, "Gait disorders in the elderly." *American family physician*, vol. 35, no. 6, pp. 191–196, 1987.

[20] B. R. Bloem, J. Haan, A. M. Lagaay, W. van Beek, A. R. Wintzen, and R. A. Roos, "Investigation of gait in elderly subjects over 88 years of age," *Journal of geriatric psychiatry and neurology*, vol. 5, no. 2, pp. 78–84, 1992.

[21] N. B. Alexander and A. Goldberg, "Gait disorders: search for multiple causes." *Cleveland Clinic journal of medicine*, vol. 72, no. 7, pp. 586–586, 2005.

[22] M. E. Morris, F. Huxham, J. McGinley, K. Dodd, and R. Iansek, "The biomechanics and motor control of gait in parkinson disease," *Clinical Biomechanics*, vol. 16, no. 6, pp. 459–470, 2001.

[23] S. Kimmeskamp and E. M. Hennig, "Heel to toe motion characteristics in parkinson patients during free walking," *Clinical biomechanics*, vol. 16, no. 9, pp. 806–812, 2001.

[24] Canadian Society of Exercise Physiologists, "Canadian physical activity guidelines for older adults - 65 years and older," www.csep.ca/guidelines, 2014.

[25] S. Figueiredo, "Nordic walking: A new training for frail elderly," Master's thesis, McGill University, August 2009.

[26] M. E. Morris, C. L. Martin, and M. L. Schenkman, "Striding out with parkinson disease: evidence-based physical therapy for gait disorders," *Physical Therapy*, vol. 90, no. 2, pp. 280–288, 2010.

[27] T. C. Rubinstein, N. Giladi, and J. M. Hausdorff, "The power of cueing to circumvent dopamine deficits: a review of physical therapy treatment of gait disturbances in parkinson's disease," *Movement Disorders*, vol. 17, no. 6, pp. 1148–1160, 2002.

[28] S. S. Travis, H. H. Duncan, and W. J. McAuley, "Mall walking. an effective mental health intervention for older adults." *Journal of psychosocial nursing and mental health services*, vol. 34, no. 8, pp. 36–38, 1996.

[29] N. Salbach, N. Mayo, S. Wood-Dauphinee, J. Hanley, C. Richards, and R. Cote, "A task-orientated intervention enhances walking distance and speed in the first year post stroke: a randomized controlled trial," *Clinical rehabilitation*, vol. 18, no. 5, pp. 509–519, 2004.

[30] H. Moffet, J.-P. Collet, S. H. Shapiro, G. Paradis, F. Marquis, and L. Roy, "Effectiveness of intensive rehabilitation on functional ability and quality of life after first total knee arthroplasty: a single-blind randomized controlled trial," *Archives of physical medicine and rehabilitation*, vol. 85, no. 4, pp. 546–556, 2004.

[31] K. K. Mangione, R. L. Craik, S. S. Tomlinson, and K. M. Palombaro, "Can elderly patients who have had a hip fracture perform moderate-to high-intensity exercise at home?" *Physical Therapy*, vol. 85, no. 8, pp. 727–739, 2005.

[32] H. Park, K. J. Kim, T. Komatsu, S. K. Park, and Y. Mutoh, "Effect of combined exercise training on bone, body balance, and gait ability: a randomized

controlled study in community-dwelling elderly women," *Journal of bone and mineral metabolism*, vol. 26, no. 3, pp. 254–259, 2008.

[33] T. B. Moeslund, A. Hilton, and V. Krüger, "A survey of advances in vision-based human motion capture and analysis," *Computer vision and image understanding*, vol. 104, no. 2, pp. 90–126, 2006.

[34] R. G. Cutlip, C. Mancinelli, F. Huber, and J. DiPasquale, "Evaluation of an instrumented walkway for measurement of the kinematic parameters of gait," *Gait & posture*, vol. 12, no. 2, pp. 134–138, 2000.

[35] K. Aminian, B. Najafi, C. Büla, P.-F. Leyvraz, and P. Robert, "Spatio-temporal parameters of gait measured by an ambulatory system using miniature gyroscopes," *Journal of biomechanics*, vol. 35, no. 5, pp. 689–699, 2002.

[36] S. R. Hundza, W. R. Hook, C. R. Harris, S. V. Mahajan, P. A. Leslie, C. A. Spani, L. G. Spalteholz, B. J. Birch, D. T. Commandeur, and N. J. Livingston, "Accurate and reliable gait cycle detection in parkinson's disease," 2013.

[37] W. Zijlstra and A. L. Hof, "Assessment of spatio-temporal gait parameters from trunk accelerations during human walking," *Gait & posture*, vol. 18, no. 2, pp. 1–10, 2003.

[38] M. Aung, S. Thies, L. Kenney, D. Howard, R. Selles, A. Findlow, and J. Goulermas, "Automated detection of instantaneous gait events using time frequency analysis and manifold embedding." *IEEE trans. on neural sys. and rehab. eng.*, 2013.

[39] C. M. O'Connor, S. K. Thorpe, M. J. O'Malley, and C. L. Vaughan, "Automatic detection of gait events using kinematic data," *Gait & posture*, vol. 25, no. 3, pp. 469–474, 2007.

[40] E. Desailly, Y. Daniel, P. Sardain, and P. Lacouture, "Foot contact event detection using kinematic data in cerebral palsy children and normal adults gait," *Gait & posture*, vol. 29, no. 1, pp. 76–80, 2009.

[41] A. Sharenkov, A. N. Agres, J. F. Funk, G. N. Duda, and H. Boeth, "Automatic initial contact detection during overground walking for clinical use," *Gait & posture*, vol. 40, no. 4, pp. 730–734, 2014.

[42] A. Findlow, J. Goulermas, C. Nester, D. Howard, and L. Kenney, "Predicting lower limb joint kinematics using wearable motion sensors," *Gait & posture*, vol. 28, no. 1, pp. 120–126, 2008.

[43] H. Rouhani, J. Favre, X. Crevoisier, and K. Aminian, "Measurement of multi-segment foot joint angles during gait using a wearable system," *Journal of biomechanical engineering*, vol. 134, no. 6, p. 061006, 2012.

[44] M. Yoneyama, Y. Kurihara, K. Watanabe, and H. Mitoma, "Accelerometry-based gait analysis and its application to parkinson's disease assessment. part 1: Detection of stride eventaccelerometry-based gait analysis and its application to parkinson's disease assessment." 2013.

[45] A. Dalton, H. Khalil, M. Busse, A. Rosser, R. van Deursen, and G. ÓLaighin, "Analysis of gait and balance through a single triaxial accelerometer in presymptomatic and symptomatic huntington's disease," *Gait & posture*, vol. 37, no. 1, pp. 49–54, 2013.

[46] Y. Guo, D. Wu, G. Liu, G. Zhao, B. Huang, and L. Wang, "A low-cost body inertial-sensing network for practical gait discrimination of hemiplegia patients," *TELEMEDICINE and e-HEALTH*, vol. 18, no. 10, pp. 748–754, 2012.

[47] A. Adkin, B. Bloem, and J. Allum, "Trunk sway measurements during stance and gait tasks in parkinson's disease," *Gait & posture*, vol. 22, no. 3, pp. 240–249, 2005.

[48] P. B. Shull, W. Jirattigalachote, M. A. Hunt, M. R. Cutkosky, and S. L. Delp, "Quantified self and human movement: A review on the clinical impact of wearable sensing and feedback for gait analysis and intervention," *Gait & posture*, vol. 40, no. 1, pp. 11–19, 2014.

[49] L. L. Verhoeff, C. G. Horlings, L. J. Janssen, S. A. Bridenbaugh, and J. H. Allum, "Effects of biofeedback on trunk sway during dual tasking in the healthy young and elderly," *Gait & posture*, vol. 30, no. 1, pp. 76–81, 2009.

[50] L. J. Janssen, L. L. Verhoeff, C. G. Horlings, and J. H. Allum, "Directional effects of biofeedback on trunk sway during gait tasks in healthy young subjects," *Gait & posture*, vol. 29, no. 4, pp. 575–581, 2009.

[51] J. R. Davis, M. G. Carpenter, R. Tschanz, S. Meyes, D. Debrunner, J. Burger, and J. H. Allum, "Trunk sway reductions in young and older adults using multimodal biofeedback," *Gait & posture*, vol. 31, no. 4, pp. 465–472, 2010.

[52] F. B. Horak, M. Dozza, R. Peterka, L. Chiari, and C. Wall III, "Vibrotactile biofeedback improves tandem gait in patients with unilateral vestibular loss," *Annals of the New York Academy of Sciences*, vol. 1164, no. 1, pp. 279–281, 2009.

[53] P. B. Shull, K. L. Lurie, M. R. Cutkosky, and T. F. Besier, "Training multiparameter gaits to reduce the knee adduction moment with data-driven models and haptic feedback," *Journal of biomechanics*, vol. 44, no. 8, pp. 1605–1609, 2011.

[54] N. Basaglia, N. Mazzini, P. Boldrini, P. Bacciglieri, E. Contenti, and G. Ferraresi, "Biofeedback treatment of genu-recurvatum using an electrogoniometric device with an acoustic signal. one-year follow-up." *Scandinavian journal of rehabilitation medicine*, vol. 21, no. 3, pp. 125–130, 1988.

[55] R. Koheil and A. R. Mandel, "Joint position biofeedback facilitation of physical therapy in gait training," *American Journal of Physical Medicine & Rehabilitation*, vol. 59, no. 6, pp. 288–297, 1980.

[56] A. Faivre, M. Dahan, B. Parratte, and G. Monnier, "Instrumented shoes for pathological gait assessment," *Mechanics research communications*, vol. 31, no. 5, pp. 627–632, 2004.

[57] Shimmer Research, "Shimmer 2r motion development kit," www.shimmersensing.com, 2010.

[58] *9DOF Calibration Application User Manual*, Rev 1.2b ed., Shimmer Research, 2013.

[59] G. Pendharkar, G. R. Naik, and H. T. Nguyen, "Using blind source separation on accelerometry data to analyze and distinguish the toe walking gait from normal gait in itw children," *Biomedical Signal Processing and Control*, vol. 13, pp. 41–49, 2014.

[60] S. Rogers and M. Girolami, *A first course in machine learning.* CRC Press, 2011.

[61] User:ZackWeinberg, “Svm separating hyperplanes (svg),” http://commons.wikimedia.org/wiki/File:Svm_separating_hyperplanes_(SVG).svg#mediaviewer/File:Svm_separating_hyperplanes_(SVG).svg, 2014, based on PNG version by User:Cyc - This file was derived from:Svm_separating_hyperplanes.png. Licensed under Creative Commons Attribution-Share Alike 3.0 via Wikimedia Commons.

[62] K. H. Zou, A. J. O’Malley, and L. Mauri, “Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models,” *Circulation*, vol. 115, no. 5, pp. 654–657, 2007.

[63] R. E. Roberson and R. Schwertassek, *Dynamics of multibody systems*. Springer-Verlag Berlin, 1988, vol. 18.

[64] J. M. Wood, “Understanding and computing cohen’s kappa: A tutorial.” *WebPsychEmpiricist. Web Journal at http://wpe. info/*, 2007.

[65] S. J. Lee and J. Hidler, “Biomechanics of overground vs. treadmill walking in healthy individuals,” *Journal of Applied Physiology*, vol. 104, no. 3, pp. 747–755, 2008.