Design and Implementation of an Artificial Intelligence-Driven Gait Phase Recognition System for Orthotic Knee Control

John Farah

Thesis submitted to the Faculty of Engineering
in partial fulfillment of the requirements for the degree of

MASTER OF APPLIED SCIENCE
in
Biomedical Engineering

Ottawa-Carleton Institute for Biomedical Engineering

University of Ottawa

Ottawa, Ontario

May 2018

Abstract

Microprocessor-controlled stance-control knee-ankle-foot orthoses (M-SCKAFO) can have multiple sensors at all lower-limb segments. This causes M-SCKAFO to be bulky and expensive, with complex control systems. Stance-control systems with sensors local to the knee-joint component would provide a modular orthosis component for easier orthotist-customization and personalization for users with knee-extensor weakness. A gait phase recognition model (GPR) is essential for a fast, accurate, and generalizable real-time orthosis-control.

This thesis designed, developed, and evaluated a machine learning-based GPR model for intelligent M-SCKAFO control. The model used gait signals that mimicked thigh inertial sensor and knee angle. Machine learning was implemented to identify gait phases across multiple surface conditions and walking speeds. Thigh-segment angular velocity, thigh-segment acceleration, and knee angle were calculated from 30 able-bodied participants for level and up, down, right-cross, and left-cross slopes at 0.8, 0.6, 0.4 m/s, and self-paced speeds (1.33 m/s, SD = 0.04 m/s). A logistic model tree (LMT) was built with a set of 20 signal features extracted from 0.1s sliding windows. The GPR model determined the walking state and was fed through a "transition sequence verification and correction" (TSVC) algorithm to deal with continuous states. The GPR model was evaluated on a different data set from 12 able-bodied individuals that completed the same walking protocol (validation set).

Gait phases were classified successfully regardless of surface-level, walking speed, and individual walking variability. The LMT had a tree size of 1643 nodes with 822 leaf nodes. The GPR model produced overall classification accuracy of 98.4% and increased to 98.7% when TSVC was applied. Results also demonstrated evidence of strong model-generalizability with GPR accuracy of 90.6% and increased to 98.6% when TSVC was applied, on the validation set.

This research demonstrated that local sensor signals from thigh and knee, integrated with machine intelligence algorithms, provided viable GPR suitable for real-time orthosis-control. The logistic decision tree model and feature selection approach were computationally efficient for real-time GPR and gave reliable, robust, and generalizable results across multiple surfaces, walking speeds, and individual walking variability. GPR also benefitted from transition sequence verification and correction algorithms, providing enhanced gait phase classification performance.

Table of Contents

List of Figures

List of Tables

Abbreviations and Definitions

Acc	Thigh-segment acceleration
AI	Artificial intelligence
AngVel	Thigh-segment angular Velocity
ANN	Artificial neural network
CAGH	Correlation along gravity and heading directions
CAREN	Computer assisted rehabilitation environment
CART	Classification and regression tree
CFS	Correlation-based feature selection
DT	Decision tree
EMG	Electromyography
FCBF	Fast correlation-based filter
FO	Foot-off
FN	False negative
FP	False positive
HAR	Human activity recognition
HMM	Hidden Markov model
GPR	Gait phase recognition
GRF	Ground reaction force
IC	Initial contact
IEEE	Institute of Electrical and Electronics Engineers
IMU	Inertial measurement unit
JNER	Journal of NeuroEngineering and Rehabilitation
KA	Knee angle
KAFO	Knee-ankle-foot orthosis
LMT	Logistic model tree
LR	Loading response
M-SCKAFO	Microprocessor-controlled stance-control knee-ankle-foot orthosis
Matlab	High-level technical computing language and interactive environment
MCC	Matthew's correlation coefficient
MeMeA	International Symposium on Medical Measurements and Application
MKF	Maximum knee flexion
MLP	Multi-layer perceptron
MS	Mid-stance
NB	Naïve bayes
OWVS	Ottawalk-variable speed
PO	Push-off
Prec	Precision
RF	Random forest
ROC	Receiver operator characteristic

ROM	Range of motion
SCKAFO	Stance-control knee-ankle-foot orthosis
Sens	Sensitivity
Spec	Specificity
SD	Standard deviation
SVM	Support vector machine
SP	Self-paced
TN	True negative
TOHRC	The Ottawa Hospital Rehabilitation Centre
TP	True positive
TS	Terminal stance
TSw	Terminal swing
WEKA	Waikato environment for knowledge and analysis
WMMS	Wearable mobility monitoring system
5-FCV	Five-fold cross validation
10-FCV	Ten-fold cross validation
6MWT	Six-minute walk test

Acknowledgments

I would like to thank my supervisors, Dr. Edward D. Lemaire and Dr. Natalie Baddour for their highly valued guidance, direction, and expertise throughout the completion of this thesis. I greatly appreciate everything you both have done for me. You have helped me grow as a student, professional, and as an individual.

I would also like to acknowledge Jawaad Bhatti, Blatchford & Sons, for his technical expertise in robotics engineering, and the Natural Sciences and Engineering Research Council of Canada Discovery Grant for funding the research of this thesis.

Thank you to Brandon Fournier, Andrew Herbert-Copley, Andrew Smith, and Hossein Ghozildeh.

My greatest appreciation stems from my wonderful family. Thank you to my hard-working father David Farah, loving mother Julianne Farah, wonderful sister Diane Farah, and amazing brother Sami Farah, for their endless guidance, support, and love throughout my life. Thank you to my lovely girlfriend Eliana Hobeiche for her continuous support, strength, encouragement, patience, and love.

Thanks to all my friends for their continuous encouragement and companionship.

I would also like to acknowledge the following individuals for making my experience at The Ottawa Hospital Rehabilitation Centre an absolute blast.

The Ottawa Hospital Rehabilitation Centre Students/Staff

- Amaia Hernandez
- Emily Sinitski

Finally, I would like to thank the volunteers who donated their time to participate in this research.

1. Introduction

1.1 Introduction

Walking is a complex and dynamic motor task that is important for many daily living activities. Walking analysis can provide crucial information regarding a person's health status and functional capacity. Quantifying walking parameters by gait phase is necessary for deeper understanding of human locomotion and assistive device control [1], [2]. During locomotion, the knee joint and knee extensor and flexor muscles act as a shock absorber, provide and maintain stability during stance, and help with limb progression [3], [4]. The knee extensors (quadriceps muscles) absorb forces during body-weight acceptance, resist knee flexion during stance, and facilitate limb progression for taking the next step. If the knee extensor muscles are not sufficiently strong for these tasks, inefficient walking can occur with abnormal gait patterns [4], [5], energy consumption can increase [3], [6], [7], and knee collapse can occur due to the inability to counteract external knee flexion moments, increasing fall risk.

Muscle weakness can arise from muscular disease, peripheral or central neurological diseases, spinal cord injury, or trauma [4], [7], [8]. Individuals with knee extensor muscle weakness are prescribed a Knee-Ankle-Foot Orthosis (KAFO), which is a full leg brace that maintains the knee in full extension. Inhibiting knee flexion eliminates knee collapse but causes users to develop additional irregular gait patterns as compensatory mechanisms; such as, hip-hike, lateral sway, and vaulting, which can be chronically detrimental to the human body [4], [9].

A Stance-Control KAFO (SCKAFO) prohibits knee flexion and provides support during stance while permitting free knee motion during swing, thereby allowing more natural swing limb progression. SCKAFO offer improved mobility over conventional fixed-knee KAFO [9]–[12]. The SCKAFO control system allows knee joints to switch between stance and swing modes during locomotion.

Many SCKAFO designs are mechanically controlled, with some designs requiring the knee to reach full extension before locking the mechanism (i.e., knee flexion resistance cannot engage with a partially flexed knee). These SCKAFO present functional limitations for knee support reliability and difficulty with non-level surfaces in daily-living environments.

The most advanced SCKAFO are controlled using microprocessors. Microprocessor-controlled SCKAFO (M-SCKAFO) combine sensors data such as inertial measurement units (IMU) and pressure sensors to achieve gait phase recognition and stance-control. Rule-based and intelligent algorithms use multiple sensors to determine when the knee requires stability (stance) or free motion (swing) and are often more reliable for walking on different terrains. Some M-SCKAFO control systems can switch between multiple settings (i.e., full lock, partial resistance, free motion) to more easily negotiate between slower walking speeds, uneven ground conditions, stair descent/ascent, sit-to-stand, and stand-to-sit modes. Variable-resistance during knee-flexion also allows for stumble recovery and fall prevention by enabling a user to catch themselves during a fall. To provide sufficient stance-control functionality, knowing where the leg is during a stride is central. Therefore, real-time Gait Phase Recognition (GPR) becomes a crucial task for lower-limb orthosis-control.

In this thesis, an intelligent system that uses a hydraulic valve for variable flexion-resistance for orthosis-control was designed and implemented for the Ottawalk-Variable Speed (OWVS) knee joint. The OWVS was developed at the Ottawa Hospital Rehabilitation Centre (TOHRC) in conjunction with The Blatchford Group.

GPR using machine-learning classification with a rule-based algorithm, in conjunction with a local sensor system at the thigh and knee, was used to determine when to switch between different knee flexion resistance settings during the gait cycle. Thigh and knee kinematic data were measured with Vicon 3D motion analysis to mimic signals obtained from an IMU placed on an instrumented orthosis. All data were linearly interpolated to 200Hz to mimic an IMU's internal sampling rate for M-SCKAFO control systems. The novel intelligent stance-control system was sufficiently generalizable for different walking speeds and ground surface conditions experienced in real-world walking scenarios.

1.2 Rationale

A conventional mechanically controlled SCKAFO cannot distinguish between different terrains and sloped surfaces and may require external knee flexion/extension moments to enable a locked or open setting. People that require an orthosis for knee extensor weakness may not have sufficient quadriceps strength to fully extend their leg, or to do so reliably over a full day of walking during regular daily activities. For example, a person with knee extensor weakness may have difficulty achieving full leg extension to lock their orthosis when walking uphill. Many of these devices are also not able to lock and provide stability for an already flexed leg, making falls inevitable after a stumble.

In contrast to M-SCKAFO, mechanically controlled SCKAFO can be aesthetically pleasing, do not require external power, and are smaller (i.e., fit under trousers). However, conventional SCKAFO lack the reliability and overall functionality offered by the onboard computing and system design of the M-SCKAFO. Microprocessor-controlled orthoses could sense different terrains and, depending on the system, toggle between gait modes to provide variable resistances and in-stance dampening during the gait cycle. M-SCKAFO could provide better comfort and more natural gait, enhancing a user's experience and increasing their quality of life throughout daily living activities. While current commercially available M-SCKAFO have stance-control functionality suitable for daily activities, they are driven by systems that require sensors on all orthotic-segments (i.e., thigh, knee, shank, footplates), need external power, and require central fabrication. This results in very expensive devices that are heavy, bulky, and cannot fit under clothing, thus decreasing their aesthetic appeal.

The performance ceiling for high-end orthoses depends on their control systems. An orthosis that offers versatility and performance to the end-user is considered the better product. The best M-SCKAFO on the market, the Otto Bock C-Brace, can negotiate between level ground walking, sloped surfaces, stairs, and sit/stand motions. The C-Brace uses multiple sensors on the orthosis, requires central fabrication, and is purchased as a whole. The C-Brace's lateral component, containing the knee joint, microprocessor, and battery, makes the device too bulky to fit under trousers. An ideal stance-control system would use localize sensors to the thigh and knee to produce a system with the same if not better functionality, allowing for different foot and ankle options to be used and local fabrication by the orthotist.

By taking advantage of machine learning, intelligence algorithms integrated with inertial sensors, the system could be robust enough to eliminate the need for manual switching between level and uneven ground. Research is needed to determine whether a suitable local sensor control system can be achieved using machine learning techniques. A modular M-SCKAFO unit would provide the freedom to customize thigh, shank, and foot segments, and possibly reduce the weight and overall system cost by simpler manufacturability.

1.3 Objectives and Hypotheses

The objectives and hypotheses for this thesis are:

1. Evaluate a GPR model with various machine learning classifiers using only knee angle and thigh-segment kinematics as input.

   Hypothesis:

   a. Knee angle and thigh-segment kinematics with machine learning algorithms will identify gait phases with overall classification performances >98%.
   b. Knee angle features do not contribute to level walking GPR; classification performance without knee angle features will not change.

2. Select features independently of a classifier to identify feature subsets that enhance GPR performance. Determine the best GPR model across different surfaces and walking speeds.

   Hypothesis:

   a. Feature selection will reduce redundancy, data dimensionality, and provide a feature subset with the most relevant features for real-time GPR.
   b. Training models with different surfaces will have different feature sets, following feature selection.

c. Models trained on all surfaces and walking speeds will be generalizable and achieve the best classification performance compared to training models without all surfaces and walking speeds.

3. Evaluate a combination of logistic regression decision tree and rule-based transition sequence algorithm to improve classification performance for real-time orthosis-control.

   Hypothesis:

   a. Integrating signal features from the thigh and knee with a logistic model decision tree will identify gait phases across surfaces and walking speeds, and be generalizable to new data.

   b. Implementing a “transition sequence verification and correction” algorithm will enhance classification results.

1.4 Thesis Contributions

This thesis produced several contributions in the gait phase identification, machine intelligence, and orthosis-control fields. Machine learning was used to develop a novel high-performance gait phase identification model that uses localized sensor signals at the thigh and knee. Useful signal features, independent of the classifier, were extracted from small data windows for real-time application and were shown to be appropriate for machine learning implementation. This study also provided evidence that gait recognition systems and smart orthoses would benefit from training on multiple walking conditions and speeds across real-world walking scenarios, rather than just level ground and experimenter-selected walking speeds. In addition, implementing a gait phase transition sequence verification and correction algorithm improved model performance and was generalizable to new data. By using local sensors at the thigh and knee, sensor system complexity is reduced and enables modular unit production for microprocessor-controlled stance-control knee-ankle-foot orthoses.

The local sensor machine learning-based GPR model enables a modular M-SCKAFO unit that does not require specific components distal to the knee unit. This model exceeded the target performance measures for multiple walking speeds and surface conditions that would be experienced through daily walking activities. The local sensor stance-control system facilitates orthosis-customizability and manufacturing, making the product increasingly cost-effective for end-users whilst maintaining functionality and additional benefits such as aesthetic appeal.

1.5 Thesis outline

This thesis manuscript is divided into six chapters. Chapter 2 is a literature review that describes knee joint function, normal gait, pathological gait, and knee extensor weakness, machine learning classifiers, gait phase recognition models, and orthosis solutions (i.e., KAFO, SCKAFO, M-SCKAFO).

Chapter 3 contains a full manuscript presented at the IEEE International Symposium on Medical Measurements and Applications (MeMeA) in Rochester, Minnesota, USA (May 2017) and addresses objective 1. This chapter investigates viable machine learning classifiers to identify a classifier to be used for the final model research.

Chapter 4 contains a journal manuscript submitted to IEEE Transactions on Neural Systems and Rehabilitation Engineering, and addresses objective 2. This chapter describes signal features and training model that are best for gait phase identification for thigh-segment and knee angle kinematics for multiple walking speeds and surfaces.

Chapter 5 contains a journal manuscript that will be submitted to the Journal of Neuro-Engineering and Rehabilitation and addresses objective 3. This chapter describes the optimized GPR model integrated with rule-based algorithms for M-SCKAFO implementation and evaluation.

Chapter 6 presents a thesis summary, overview of the thesis contributions and suggestions for future work.

2. Literature Review

This chapter introduces normal gait, a brief overview of pathological gait, the concept of gait phase recognition (GPR), and reviews ubiquitous lower-limb orthosis solutions for knee extensor weakness. The motivation, equipment, and methods used to control stance-control knee-ankle-foot orthoses are explored, as well as the challenges and need for additional work in this unique area. Reviews of GPR studies that use machine learning, and orthosis control systems that use mechanical-control and microprocessor-control for stance-control orthoses are presented.

2.1 Normal Gait

Gait patterns can be characterized by a sequence of events associated with translational motion along the planes of progression; including, forces, velocities, power, and energy changes during walking while adapting to different environments and terrains.

Human gait occurs in a periodic and repetitive cycle involving limb and torso movements [4]. A normal gait cycle begins at IC, often referred to as foot strike, and continues until the same foot makes contact with ground again directly after swing. The gait cycle, or stride, is divided into two main phases: stance and swing (Figure 2.1). A stride starts with one foot-ground contact and ends with subsequent ground contact of the same foot [3], [4], [13]. Stance phase is when the foot remains in contact with the ground, accounting for approximately 60% of a full stride. Stance sub-phases are initial contact (IC), loading response (LR), mid-stance (MS), and terminal stance (TS) [14]. Swing phase is the period after stance phase when the foot leaves the ground and swings through to the following ground contact. Swing sub-phases are pre-swing, initial swing, mid-swing, and terminal swing.

Stance can be divided into three main tasks: weight acceptance, single limb support, and swing limb advancement [3], [4], [13]. Weight acceptance consists of IC and LR, includes shock absorption, achieves limb stability to prevent collapse, and supports contralateral limb progression. To prevent forward falling and limb collapse due to forward impetus, knee and ankle muscles activate to maintain stability (co-activation of antagonist muscle groups). Following weight-acceptance, single limb support involves MS and TS and provides body balance while the contralateral limb is in swing. Swing limb advancement lifts the limb into the air and propels it forward into swing. Knee extensor and flexor activation are essential during stance [3], [4], [13]. Without sufficient strength, a person would develop irregular gait patterns, experience knee-collapse, and fall.

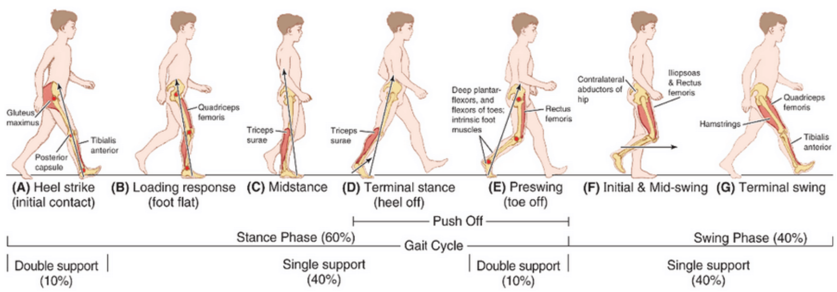

Figure 2.1: Gait cycle phases, sub-phases, and events (adopted from [14]).

2.2 Knee Joint Function

The knee joint is the largest synovial joint, connecting the thigh bone (femur) to the shin bone (tibia) [15]. Figure 2.2 illustrates the knee's primary muscle groups, the flexors (hamstrings) and extensors (quadriceps muscles). The flexors and extensors work together and against one another to stabilize the knee during loading response and for swing limb progression. The knee muscles play a crucial role in human locomotion and, thus, muscle weakness can cause detrimental effects to the human body [3], [4], [9], [13].

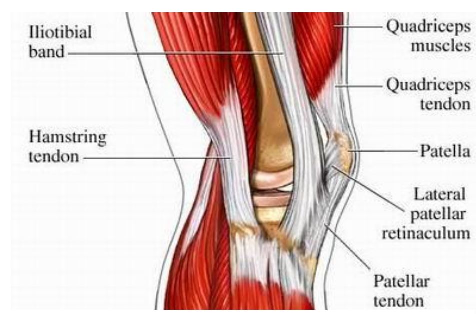

Figure 2.2: Human knee joint

The knee has a large range of motion (ROM), primarily in the sagittal plane, referring to flexion and extension. Knee extensors and flexors are active for stance stability, foot clearance, and swing limb advancement [3]. These muscles control the flexion and extension rate, induced by the Ground Reaction Force (GRF) in stance and leg inertia in swing [7] (Figure 2.3). For instance, when the GRF vector acts anterior to the knee during initial contact, knee flexors must engage to counteract the external moment and stabilize the knee. Similarly, when a GRF vector acts posterior to the knee during loading response, terminal stance, and pre-swing, knee extensors must engage to maintain stability and prepare for swing limb progression (Figure 2.4).

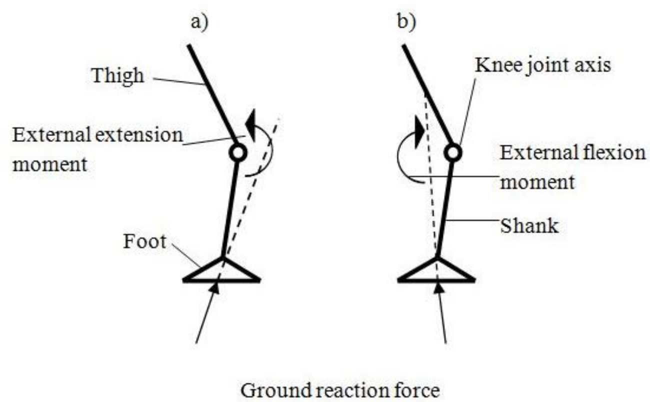

Figure 2.3: a) External extension moment from ground reaction force vector anterior to the knee. b) External flexion moment from ground reaction force vector posterior to the knee (adopted from [4]).

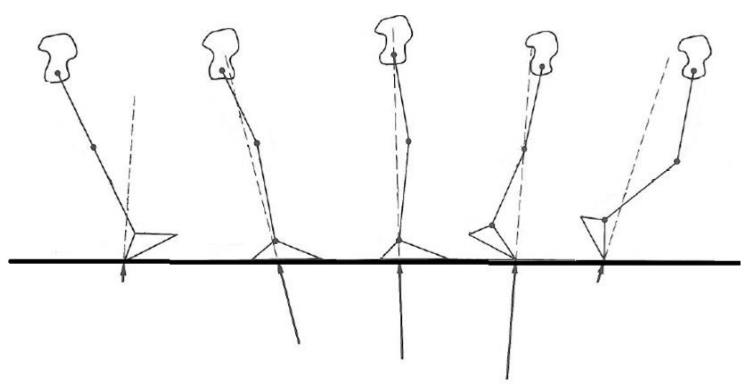

Figure 2.4: Ground reaction force vectors at initial contact, loading response, mid-stance, terminal stance, and pre-swing (adopted from [4]).

2.2.1 **Level Ground Walking**

The knee flexes twice during a gait cycle, during loading response and mid-stance for shock absorption and during early swing for foot clearance and swing limb progression [3], [4], [13], [16]. Normal knee motion has a range of 0°-60° [3] (Figure 2.5). At foot strike, the knee is flexed at about 5° and rapidly continues to flex to ~20° for limb and body stabilization during loading response. The knee begins to flex before swing and reaches about ~60° for toe and foot clearance during swing limb advancement [3], [4], [13].

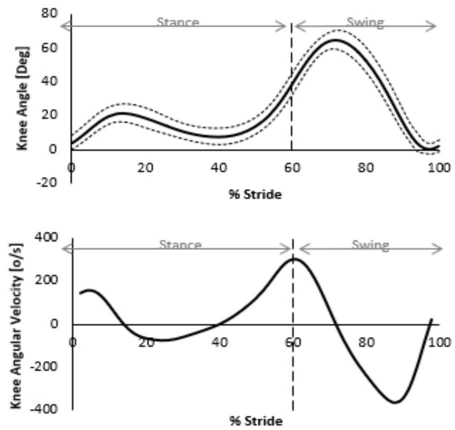

Figure 2.5: Knee angle (°) (top) and angular velocity (°/s) (bottom) as a percent of gait cycle [14].

Knee flexion angular velocity is low during mid-stance [3] but increases to a maximum of ~350°/s at foot off. The knee then extends, reaching a peak >350°/s at terminal swing. Table 2-1 summarizes the role of extensors and flexors during a normal gait cycle.

Table 2-1: Summary of normal knee function during a stride.

Phase	**Task**	**Events**	**Cycle (%)**	**Description**	**Functional Interpretation**
Stance	Weight Acceptance	Initial contact	0-2%	Fully extended (0°)	Stabilization and weight acceptance
		Loading response	2-12%	Flexion to ~20° (~300°/s)	Shock absorption and load maintenance
	Single Limb Support	Mid-stance	12-31%	Gradual extension (~150°/s)	Stabilized weight bearing and limb progression
		Terminal stance	31-50%	Active maximum extension (5°)	Maximize step length and stability, onset of swing
	Swing Limb Advancement	Pre-swing	50-62%	Passive flexion (~350°/s)	Preparation of swing
Swing		Initial swing	62-75%	Active maximum flexion (60°)	Toe and foot clearance for limb progression
		Mid-swing	75-87%	Passive extension (~350°/s)	Knee begins to extend to advance limb and prepare
		Terminal swing	87-100%	Fully extended (0°)	Knee extensors eccentrically contract for limb deceleration to prepare for foot strike

2.2.2 **Pathological Gait and Knee Extensor Weakness**

Muscular pathologies such as poliomyelitis, Guillain-Barre syndrome, muscular dystrophy, and muscular atrophy [3] may lead to knee extensor weakness. The ability to extend and flex the knee is crucial for weight acceptance and support during stance [3], [4], [13], [17]. Knee extensors primarily control the rate of knee flexion, stabilizing the knee during LR by counteracting external flexion moments induced by GRF [4], [17]. Without the knee extensors, the limb cannot support body weight without irregular gait patterns (i.e., knee hyperextension, hip-hiking, circumduction, vaulting [10]). Co-activation of antagonist muscles (flexors, extensors) is required to stabilize and support an individual's weight during stance phase.

For stair climbing, someone with severe knee extensor weakness cannot perform step over step locomotion. Compensatory gait mechanisms/modalities for stair climbing involve a fully extended affected knee whereas the contralateral sound limb flexes and extends for step ascension.

2.3 Knee-Ankle-Foot Orthoses

Knee-Ankle-Foot Orthoses (KAFO) are lower extremity assistive walking devices that allow individuals with knee extensor weakness to support themselves during stance [1], [9], [10], [12], [18]–[21]. Over time, people using conventional fixed-knee KAFO begin to adopt unnatural gait patterns to compensate for the restrictive knee range of motion. Compensatory walking strategies that allow the braced limb to clear the foot include hip-hike (raising the affected limb), circumduction (swinging the leg lateral around the body), lateral trunk sway, and vaulting (contralateral ankle plantar flexion) [4], [5], [9], [12]. Immobilized knee-walking reduces walking efficiency by 24% [22], leading to higher energy expenditure, and can be chronically detrimental to the body [6], [23]; including, joint dysfunction and soft-tissue damage at the hip and lower back. In addition, people walk slower [24] and may experience fatigue, mobility loss, and pain. Since foot clearance during swing is harder to achieve, walking with fixed-knee KAFO becomes increasingly difficult for different ground levels such as sloped surfaces (i.e., incline, decline, cross-slopes) and stairs. Free knee motion during the gait is essential for natural walking across daily living activities. Due to these functional limitations and issues, KAFO rejection can reach 60% [6], [19]. Prominent issues for accepting an orthosis are weight, size, noise, and cost [9].

2.4 Stance-Control Knee-Ankle-Foot Orthoses

Stance-Control KAFO (SCKAFO) are full leg braces that support the leg during weight bearing, but differ from fixed-knee KAFO by permitting free-knee motion during swing. According to Yakimovich et al. [10], their SCKAFO provided a 488% increase in knee ROM in contrast to walking with a fixed-knee KAFO. Although irregular gait patterns are still present due to fully extended and locked knee during stance, SCKAFO offer many advantages over fixed-knee KAFO for people with knee-extensor weakness [9], [10], [12], [18], [19]. By simply allowing free knee rotation during swing, SCKAFO improve mobility, provide more natural gait, lessen compensatory mechanisms and associated joint loads that increase energy consumption, increase foot clearance, increase walking speed, and have greater user satisfaction.

The most important and challenging task for SCKAFO designs is determining when to engage the locking mechanism to provide support and unlock to permit free knee motion. Mechanically-controlled SCKAFO have advantages and disadvantages [9], [19]. Weighted/spring-loaded pawls (e.g. Otto Bock Free Walk, Becker UTX [9]) or belt clamping design [9], [10], [12] use leg position to lock the knee at initial contact and disengage at foot-off. Most mechanically-controlled SCKAFO require full knee extension to engage knee-lock [9], [19], [25]; therefore, individuals who cannot extend their leg at each step have unreliable stance-control. Mechanical SCKAFO that cannot engage at any angle would not resist knee flexion during a stumble and, with the absence of knee flexion resistance, would result in a fall.

The Horton Stance-Control-Orthosis [9], [19], [26] and the Ottawalk Belt-Clamping joint experimental design [10], [12] use GRF to engage the locking mechanism at any knee angle. These techniques use foot latches/switches such as pushrods and pressure sensors located at the orthotic-foot segment to engage the locking mechanism when the foot contacts the ground. Persons with knee-extensor weakness that have sufficient hip-flexion control can also achieve stance-control through other means, such as angular velocity based approaches [7], [18], [25], [27]. The angular velocity based approach is based on the premise that knee-flexion angular velocity is greater during a stumble than during walking [18], [25]. The stumble recovery system engages when a pre-set knee-flexion angular velocity threshold is surpassed. Lemaire et al. [18], [25] demonstrated that it was possible to remove the need for mechanical and electronic components located at the SCKAFO ankle and foot segments while providing knee-flexion resistance at any knee angle. The design used rotary hydraulics to resist knee flexion whenever a preset angular velocity threshold was surpassed.

The main advantages of a mechanically controlled SCKAFO include free knee motion during swing, use of relatively simple control systems that do not require external power, a low profile, low weight, and the ability to fit under trousers. Unfortunately, mechanically controlled SCKAFO can have inconsistent stance/swing phase recognition leading to unreliable locking and less functional versatility across different walking conditions. Mechanically controlled SCKAFO may function poorly between different gait modes, terrains, and daily living environments (i.e., stairs, ramps, uneven ground), as well as a range of walking speeds [1], [9], [18], [19], [24], [25].

2.4.1 **Microprocessor-Controlled Stance-Control Knee-Ankle-Foot Orthoses**

According to the SCKAFO design review by Tian et al. [19], KAFO fall under three categories: passive KAFO (fixed-knee), conventional SCKAFO (mechanically-controlled), and dynamic KAFO. Dynamic KAFO can use pneumatics actuators, linear springs, and hydraulics to control swing and stance phases [19], and can also employ microprocessors to dictate when to engage/disengage the locking mechanism. Emerging microprocessor-controlled SCKAFO (M-SCKAFO) use various electronic sensors to guide and stabilize the knee during gait [1], [6], [7], [9], [19]–[21], [27]–[35] and can provide enhanced function, versatility, and rehabilitation (i.e., partial flexion resistances) across daily walking activities.

Integrating pressure sensors, strain gauges, joint angle sensors, and inertial measurement units (IMU, accelerometer and gyroscope) with an onboard processor provides all the benefits that come with SCKAFO in addition to increased reliability for engaging/disengaging between stance and swing phases. Electronic sensors coupled with computation algorithms and signal processing determines whether a user is in stance or swing. Various control algorithms determine limb orientation and/or position in the gait cycle and prompt the knee joint mechanism to switch to a locked, free, or partial-resistance setting. Therefore, to enable appropriate function, activity classification and control from wearable sensor data becomes essential.

M-SCKAFO also have the advantage of providing more gait modes (i.e., ramps, curbs, and stairs), providing stumble recovery through fall detection algorithms, and sensing and navigating different walking terrains. These orthoses rely on complex algorithms that combine the sensors that may have high computational costs, require an external power source and need to be charged regularly. Since multiple sensors may be located at different orthosis segments, devices can become bulky (i.e., cannot fit under pants), aesthetically unappealing, more expensive, and cannot be personalized.

2.4.1.1 Becker Orthopedic 9001 E-Knee

The Becker Orthopedic 9001 E-Knee's (Figure 2.6) [35] uses an electromagnetic system that consists of a magnetically activated one-way dog clutch, two ratchet plates, a magnetic coil, a tension spring, a foot switch with a pressure sensor, and a battery pack. The pressure sensor detects weight bearing to dictate swing and stance phases. The coil becomes energized and connects two circular ratchet plates together when stance is detected. The two ratchet plates have teeth that restrict angular motion to free knee extension [9], [19], [35].

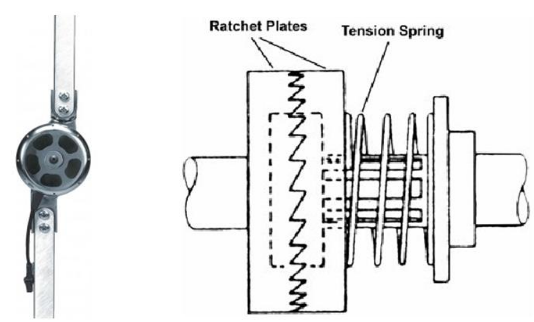

Figure 2.6: Becker Orthopedic 9001 E-Knee joint and electromagnetic system schematic.

Becker Orthopedic 9001 E-Knee advantages include reliable stance and swing knee-locking engagement/disengagement. In addition, the knee joint does not permit knee flexion when the two ratchet plates are connected, allowing for knee flexion resistance at any knee angle. However, ratchet devices have limited flexion resistance positions. Specifically, this knee joint consists of 60 ratchet teeth and therefore allows up to 6° between each knee flexion locking positions [9]. Moreover, when the two ratchet plates are engaged during weight bearing, they generate clicking sounds during knee rotation [9] and can result in orthosis abandonment for end-users.

2.4.1.2 Biologically-Based Actuation System KAFO Design

Cullel et al. [7] and Moreno et al. [27] proposed a KAFO with actuated knee and ankle joints. Their actuator system design attempted to mimic the leg muscles during gait, be modular and adaptable to different people with lower limb weakness. They approximated a linear relationship elastic performance between angular position and torque at the knee-load-displacement [7]. The actuator system was applied to a KAFO with a four-bar linkage knee joint with two compression springs that simulates cruciate ligament movement, guided by the knee's anatomical movement (Figure 2.7). An electromagnetic solenoid switches between modes as a function of the gait cycle. One compression spring has a higher spring constant than the other. The spring stores energy during weight-bearing and returns energy during terminal stance to help extend the affected limb. The second spring compresses during swing and helps extend the knee at terminal swing. Their rule-based condition system used a knee goniometer for angle, gyroscope for shank angular velocity, gyroscope for foot angular velocity, and an accelerometer on the foot segment.

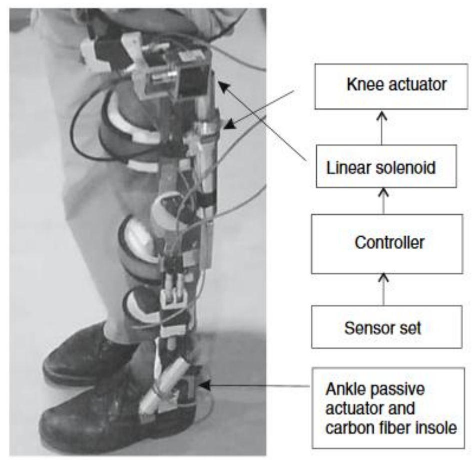

Figure 2.7: Biologically-based actuation KAFO system design by Cullel & Moreno et al.et al. [5], [28].

Advantages for this KAFO include versatility and customizability due to its biologically-based control system, reliable lock mechanism, switching functionality, knee stiffness adaptation with the automatic system, and partial knee-assistance during swing [7], [9], [19], [27]. Unfortunately, the KAFO was large, bulky and lacked aesthetic appeal.

2.4.1.3 Otto Bock Sensor Walk

Kaufman, et al. [6] and Irby, et al. [30] developed a dynamic SCKAFO that used a wrap-spring clutch that was commercialized as the Otto Bock Sensor Walk (Figure 2.8). The knee locked and unlocked according to a solenoid that transmitted torque during knee-flexion and tightened around two clutch hubs to stop knee-flexion. The solenoid loosened during knee-extension, providing unhampered knee motion during swing. A microcontroller interpreted knee angle sensor data and foot plantar pressure through pressure sensors at the heel and forefoot [6], [9], [19], [30]. The sensors detected when the limb was in stance or swing and controlled the solenoid accordingly.

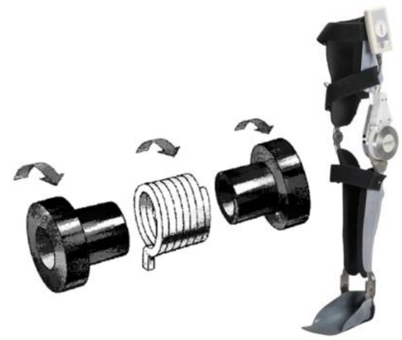

Figure 2.8: Otto Bock Sensor Walk SCKAFO with wrap-clutch mechanism.

This M-SCKAFO resisted knee-flexion at any point throughout the gait cycle due to weight-bearing detection, had a reliable knee locking mechanism control, and was very strong. This system can be used for patients weighing up to 136 kg [9] . The Sensor Walk's drawbacks include being heavy and bulky, therefore it is cosmetically unappealing. The device was also difficult to take on and off [9]. The Sensor Walk has been removed from the Otto Bock orthosis line of products and has been replaced by its successor, the C-Brace (discussed in Section 2.4.1.7).

2.4.1.4 Quasi-Passive Compliant SCKAFO

The quasi-passive compliant SCKAFO (Figure 2.9) is an experimental device [36], [37] that was inspired by moment-angle analysis of knee function behavior and approximated by a linear torsional spring during stance. A friction-based latching mechanism locked and unlocked the knee joint and a pulley wrapped around the knee joint to pull and convert linear stiffness to torsional stiffness for knee-control [19], [36], [37]. This orthosis design determined the gait phase using an instrumented shoe insole with force-sensitive resistors and integrated conductive polymers [37]. Foot sensor input was then put through a finite state machine algorithm, implemented by a microcontroller that separated the gait cycle into weight acceptance, terminal stance, and swing phases. Knee control switched to support setting during weight acceptance and disengaged during terminal stance. Throughout swing, the controller monitored knee angular velocity to identify flexion and extension. The spring disengaged support during swing and engaged during extension [19], [36], [37].

The device is quasi-passive since the control system uses a small actuator to lock and unlock the spring component but requires no other actuation. Advantages include reliable and compliant knee control, and little power and computational cost to control. Drawbacks of this device include a large knee joint component and a large control unit. The stance-control system was experimentally implemented on level ground and healthy able-bodied participants [37], suggesting that evaluation is needed on uneven ground and with people who have knee extensor weakness.

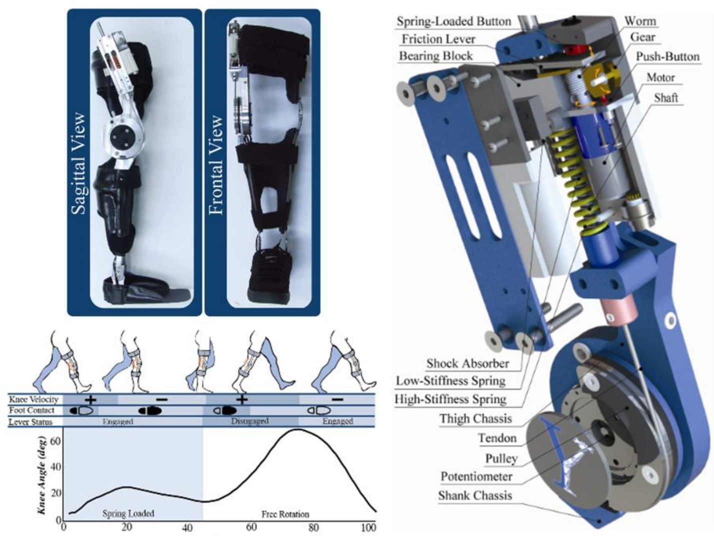

Figure 2.9: Quasi-Passive Compliant SCKAFO, control algorithm, and exploded view of the mechanical knee joint system [36], [37].

2.4.1.5 Otto Bock E-Mag Active

The Otto Bock E-Mag Active [33] (Figure 2.10) is electronically-controlled by an electromagnet. A friction wedge locks and unlocks the knee according to an accelerometer and gyroscope located within the M-SCKAFO thigh portion. Combined with a computational algorithm, the E-Mag Active engages to lock the knee during weight-bearing and releases during swing. E-Mag Active accommodates various terrains and ground elevations [32], [33].

Advantages of the Otto Bock E-Mag Active M-SCKAFO include its aesthetic appeal; stance-control functionality, even along different terrains; and orthotist-customizable to accommodate a person's gait patterns. An orthotist can adjust the device to the user if their gait improves or worsens over time. A disadvantage of this orthosis is that a person must have sufficient hip flexion strength to swing their limb forward and be able to fully extend their knee. Without the capacity for full knee extension, the locking mechanism will not engage at initial contact and disengage at terminal stance [32].

Figure 2.10: Otto Bock E-Mag Active SCKAFO (adopted from [33]).

2.4.1.6 Pneumatically-Powered KAFO

Sawicki & Ferris [21] developed a laboratory-based dynamic KAFO that was powered by artificial pneumatic muscles (Figure 2.11). Four pneumatic muscles moved the knee and two others moved the ankle [19], [21]. Electromyographic signals from the vastus lateralis and medial hamstrings controlled the extensors and flexors, respectively [19], [21].

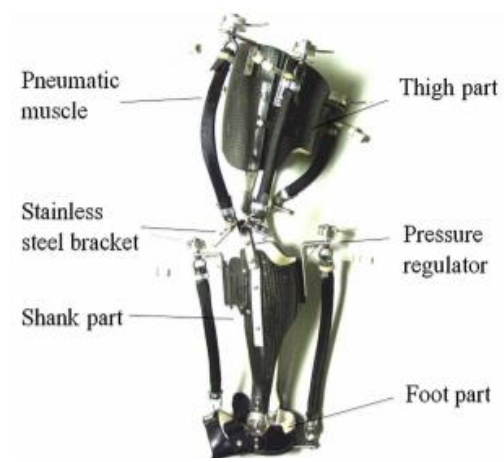

Figure 2.11: Pneumatically Powered KAFO [20].

The stance-control system had two modes programmed into the onboard microprocessor. The first mode co-activated artificial extensors and flexors. During weight-acceptance, the two flexors provide flexion moments and the two extensors provide extension moments to resist knee flexion and collapse. During mid-stance phase the artificial extensors initiated knee extension and the flexors began to deplete their air-reserves to gradually reduce the flexion moment about the knee [19], [21]. In the second mode, the system did not allow simultaneous flexor and extensor activation. During the weight-acceptance phase, the flexors provided support. During stance, the extensors activated to initiate extension. Once the extensors begin to activate, the flexors started depleting their air-reserves to allow for swing-limb progression during swing.

Pneumatically-powered KAFO used EMG signals to activate and deactivate the artificial muscles, which can be problematic for people with neurological conditions and therefore poor EMG signals. Sawicki & Ferris [21] thought that EMG was an effective way to scale orthosis-assistance because of its biologically-based nature, potentially leading to reduction of muscle recruitment compared to kinematics-based control systems, and enabling easier synergy of nervous system and orthosis control for novel motor tasks. Unfortunately, using surface EMGs as a medium for robotic control can also be inconsistent during walking due to motion artifacts. This design required an external air source for the artificial muscles and was too heavy to be portable. The device’s complex control system also required a large power source. In addition, pneumatic muscle components surrounding the orthosis made the device big, heavy, and bulky.

2.4.1.7 Otto Bock C-Brace

The Otto Bock C-Brace [38] (Figure 2.12) is currently the most versatile commercial M-SCKAFO. This orthosis has a hydraulic actuator that acts as a linear damper to guide knee control. Currently, C-Brace is the only M-SCKAFO on the market that provides knee flexion dampening (i.e., gradually allows the knee to flex during stance), by varying fluid flow rates. Two sets of knee sensors determine knee angle

Figure 2.12: The Otto Bock C-Brace [39].

and the dorsal shank spring's strain-gauge (ankle-moment sensor) determines weight-bearing [19], [20], [38]. During gait, limb position and gait phase information are sent to the microprocessor to control the knee by changing valve settings to optimize flexion or extension resistance.

The greatest advantage of the Otto Bock C-Brace is its ability to provide appropriate knee flexion and extension resistances throughout the gait cycle. This allows the user to walk with more natural gait compared to conventional lock/unlock orthoses and navigate different terrains (i.e., stairs ascent/descent) and ground conditions (i.e., ramps) [19], [38], [39]. C-Brace can detect different walking terrains and adjust valve settings accordingly in real-time, enabling seamless and natural gait. Moreover, since knee-flexion resistance can be applied at any angle, the device also has a stumble control when sensors detect a moment of instability [39]. Drawbacks for the C-Brace include its size since the lateral unit containing the knee joint and microprocessor are too large to fit under clothes. In addition, the device must be central fabricated, removing the potential for orthotist-customizability and user-personalization.

2.4.1.8 M-SCKAFO Review Summary

Table 2-2 summarizes M-SCKAFO systems, sensors, advantages, and disadvantages.

Table 2-2: A summary of microprocessor-controlled SCKAFO

Device/Model	**System**	**Sensors**	**Sensor Placement**	**Advantages**	**Disadvantages**
Becker Orthopedic 9001 E-Knee	Pressure sensor detects weight bearing. Electromagnetic coil energizes to connect two ratchet plates. Ratchet plate teeth restrict movement to free knee extension	1	Foot	Reliable locking and unlocking throughout stance and swing	Loud, finite knee flexion lock settings (6° between each)
Cullel/Moreno Actuator Design	Two compression spring actuators absorb and release energy, providing dynamic knee-extension resistance	4	Foot, shank, knee	Reliable locking and unlocking. Adaptable to individuals. Partial knee-extension assistance during stance and swing	Large, bulky, cosmetically unappealing

Otto Bock Sensor Walk	Wrap-spring clutch tightens around two hubs during knee flexion, loosens during knee extension. Knee angle and pressure sensors	3	Knee, foot	Reliable locking and unlocking. Locks at any knee angle. Users up to 136kg	Bulky, heavy, cosmetically unappealing, and difficult to don and doff
Quasi-Passive Compliant SCKAFO	Pulley pulls during stance to torsional stiffness that engages a friction-based latching mechanism. Foot sensors inside sole	4	Foot	Reliable and compliant knee control. Uses little actuation, power, computation.	Large, bulky, only tested for level ground
Otto Bock E-Mag Active	Electromagnet friction wedge locks and unlocks the knee. Inertial sensors on the thigh	2	Thigh	Aesthetically appealing, reliable lock mechanism, orthotist-customizable	Capacity for full knee extension, sufficient hip flexion strength, cannot lock at any knee angle
Pneumatically-Powered KAFO	Four pneumatic muscles for knee and two for ankle. Muscle control according to 6 EMG and limb position	6	Thigh, shank (one for each artificial muscle)	Variable damping, user-customized stance-control, mimics leg muscles	Big, heavy, bulky, requires external air supply too heavy to be portable. Surface EMG can be inconsistent
Otto Bock C-Brace	Hydraulic linear damper (flexion and extension). Sensors at knee and ankle	2	Knee, dorsal shank	Variable knee resistive moments, detects multiple terrains, stumble control	Large, lateral unit makes the device bulky, cannot wear under pants, must be central fabricated, expensive

2.5 Gait Phase Recognition

2.5.1 Conventional Gait Analysis and Measurement Systems

Gait phases can be identified using typical laboratory motion analysis systems.

- Only marker-based
- Marker-based + force plate
- Instrumented walkway (treadmills + force plate)

2.5.2 Inertial Sensor Gait Phase Recognition Systems

Wearable technology is quickly emerging in the biomedical field for daily activity self-monitoring and can be used for gait phase identification. Inertial sensors provide a portable, lightweight, low cost, and robust gait monitoring system for a wide range of gait parameter estimation applications [1], [40]–[61]. For example, [42] and [43] developed Smartphone applications that provided gait parameters such as cadence, step timing, trunk movements, and gait symmetry from inertial sensor data during a six-minute walk test.

Liu, et al. [44] detected initial contact, loading response, mid-stance, terminal stance, pre-swing, initial swing, mid-swing and terminal swing using a two-axis accelerometer and three gyroscopes on the foot, shank and thigh. While this study was successful in providing a wearable sensor system for detecting gait phases, multiple sensors were required on various lower extremity limbs. Williamson, et al. [62] calculated knee angle in real time by combining accelerometer and gyroscope signals and then determining the difference of tilt between the thigh and shank.

Senden, et al. [41] evaluated acceleration-based gait using accelerometers and presented an acceleration-based gait database for healthy participants including cadence, speed, asymmetry and irregularity [41]. Limb orientation and joint angles can be measured with two pairs of sensors. Additional studies proposed quaternion data from inertial sensors mounted to specific body segments to assess limb 3D orientation during gait [40], [45], [63]. Joint angles can be derived from body segment orientations and accurate gait phase estimation can be achieved through angular velocity integration [44]. However, integrating angular velocity data gave rise to integration drift and required periodic recalibration. Willemsen, et al. [64] determined lower extremity segment angles without integration, mitigating integration drift by assuming rigid-body dynamics, using only a pair of one dimensional accelerometers but required low-pass filtering for more precise measurements, thus introducing signal delay. Noise and signal delay are not ideal for real-time control systems like orthoses.

Pappas, et al. [56] report a reliable gait phase identification system that uses a rule-based algorithm. Gait phases were identified by prior characterization of inertial measurement signals. Their system depended on foot angular velocity, and three force sensitive resistors to determine when weight-bearing occurred. To increase gait phase detection reliability and accuracy, they developed probable gait phase transitions (Figure 2.13) to determine the possible states (gait phase) of change during the gait cycle. This GPR system worked in real-time on a microprocessor, which is important for orthosis-control applications.

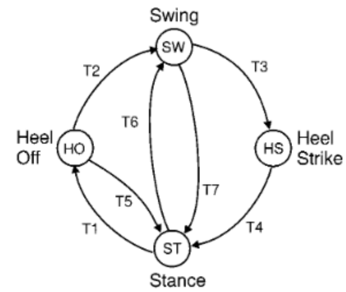

Figure 2.13: Gait phase detection transition states (adopted from [47]).

Tong and Grant [65] attached uniaxial gyroscopes to the thighs and shanks and through computational algorithms calculated knee inclinations from both signals, and using high-pass filters removed drift and re-calibrated the angles for each stride. Signals were highly correlated with motion capture output, suggesting that gyroscopes were a practical approach for walking analysis.

Abhayasinghe and Murray [53] developed a gait phase recognition system using a single IMU placed in a trouser pocket. They used accelerometer and gyroscope sensor fusion to determine thigh angles during gait and then identified IC, LR, MS, TS, pre-swing, and swing. This study was important for lower limb orthosis-control since many SCKAFO and other robotic prostheses can benefit from real-time gait phase recognition using modular electronic components located only at the thigh [1].

Gorsic, et al. [57] developed a gait phase identification system for a lower-limb robotic prosthesis. IMU attached to body segments and shoe insole sensors determined left-stance, left-right double stance, right-stance, and right-left double stance with 90% accuracy.

2.5.3 **Machine Learning Classifiers**

Machine learning is a branch of computer science, data analysis, and statistics that is crucial for pattern recognition, classification and regression. Machine learning can be considered as a computational learning scheme that enables computer algorithms to make decisions and inferences about data to accomplish specific tasks. Depending on the task, machine learning can be applied to computer prediction tasks, decision-making, computer vision and a wide range of applications in data mining and research, including gait analysis.

Examples of machine learning classifiers include, but are not limited to, support vector machines (SVM), *k*-nearest neighbour (*k*NN), hidden Markov models (HMM), naive bayes, multi-layer perceptrons (MLP) neural networks, decision trees (DT) and random forests (RF). Classifiers work as complex mathematical models that have the potential to produce reliable and effective classification. Classifiers work and 'learn' in different ways. Supervised learning is when machine learning algorithms learn from pre-existing examples. Inputs are labelled with correct outputs, and classifiers train from those examples by inferred functions that appropriately transcribe new inputs to desired results. Unsupervised learning is when the classifier makes its own inferences, finds hidden patterns, and groupings with unlabelled data. An example of unsupervised learning is cluster analysis.

Taking advantage of today's computational resources, sophisticated machine learning algorithms can provide quick and precise classification for different activities. In this section, examples of machine learning classifiers are described.

2.5.3.1 Support Vector Machines

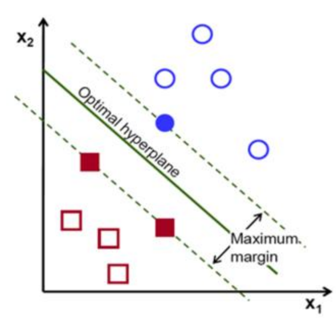

Figure 2.14: Optimized linear hyperplane separating two class distributions [75].

Support vector machines (SVM) are supervised learning models associated with classification and regression [66], using a set of different learning algorithms to determine an optimal hyperplane. The hyperplane output categorizes training data sets [67]. SVM are inherently binary classifiers, meaning that they can only distinguish between two specific categories or class memberships. However, algorithm manipulation can be done for multi-class SVMs, thereby distinguishing between more than two classes. Support vectors are training data points primarily used to determine the separation between categories. The optimal hyperplane is a probabilistic linear classifier that is computed to find the largest minimum distance between support vectors [67]. Figure 2.14 illustrates a visual representation of a support vector machine hyperplane separation.

In some cases, a more complex hyperplane is needed to distinguish between data sets, such as a curve. A curved hyperplane is computed using a Kernel functions, by mapping training data into higher dimensional feature spaces, replacing dot products with kernel functions in the mathematical process [67].

2.5.3.2 Neural Networks

In machine learning, artificial neural networks (ANN) are computational algorithms inspired by biological neurons. ANN classifiers use supervised learning and are comprised of nodes known as neurons, processing elements and layers that are connected and make up the neural network [68]. Each constituent of a neural network has its own role in the protocol for information processing used for data classification. Depending on the level of complexity, a neural network has highly interconnected elements and layers that process data in parallel to classify data and solve problems [69]. Neural networks can be structured differently; for example, feed-forward propagation, back-propagation, and perceptrons [68]–[70]. Feed-forward propagation allows input signals to travel one way, from input to output, whereas feedback neural networks allow signals to travel backwards to determine a state of equilibrium [69]. Perceptrons are neural networks that have weighted inputs and are chosen by gradient-based optimization algorithms [70]. Figure 2.15 shows an example of a neural network system with multiple inputs, three hidden layers, and four outputs for four possible category classifications.

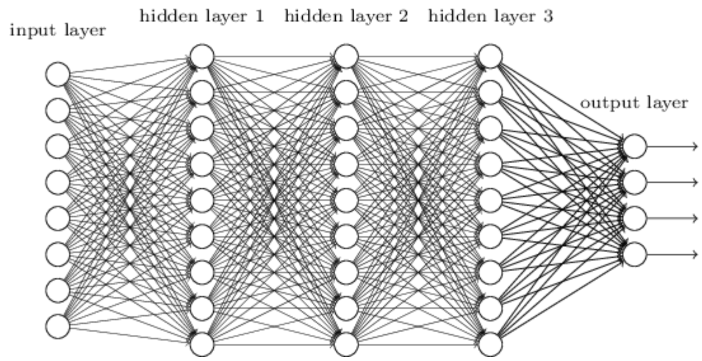

Figure 2.15: Neural network diagram showing input (features), hidden, and output layers.

2.5.3.1 Decision Trees, Random Forests, and CARTs

Decision trees (DT) are supervised machine learning models that resemble tree structures [71]. They consist of nodes that split off into additional nodes, forming branches through successive and directional links, and continue to branch out until a terminal (leaf) node (Figure 2.16). DT classification begins at the root node, located at the top of the tree, and consists of parent nodes (with child nodes beneath them) that contain a particular rule or condition. The condition ends when the branch reaches a leaf node with no further branches, and provides a classification output. Rules are typically in the form of "yes" or "no", "true" or "false" that define thresholds and condition statements specific to the classification task, and dictate the tree's pathway.

DT have many benefits over other machine learning classifiers such as neural networks. Interpretability has two manifestations. First, we can interpret the decision for any particular test pattern as the condition connection for any branch leading to a particular node. From this, classification becomes increasingly efficient (fast due to simple queries) when following a particular test pattern. Second, clear interpretations arise from categories themselves, defined by logical descriptions (class 1 = (x_j < T_j AND x_{j+1}>T_{j+1}), where x_j denotes a particular feature, and T_j denotes a specific threshold or condition, chosen during model construction.

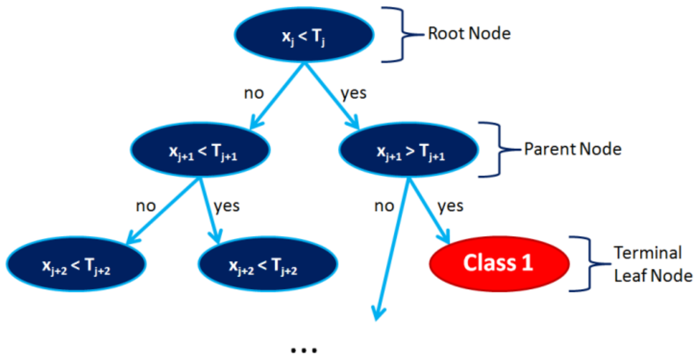

Figure 2.16: Visual representation of a binary decision tree classifier with a root node at the top, followed by parent, child, and terminating leaf node (classification output), where x_j denotes a particular feature within the feature subset, and T_j denotes a specific threshold calculated through training.

DT machine learning classifiers are powerful, easy to understand, need minimal data preparations for training, and are robust to noisy data [72]. Limitations of decision tree classifiers include high level of complexity from training, resulting in an inability to convey the data properly through high variance, and overfitting, thereby hindering data generalizability [72]. Pruning is used to simplify decision trees by removing particular leaves and branches. The outcome is often a simpler decision tree that still obtains high overall classification accuracy.

DT classifiers are simple and quick to implement, [71], and depend heavily on the trees construction process (node splitting criteria). For example, C4.5-DT trees [73] use information entropy to determine its tree structure. Greedy algorithms are another example that builds a tree by making a trade-off between efficiency and classification accuracy [74].

Auxiliary to pruning, boosting algorithms can be used to mitigate classifier limitations. For example, random forest (RF) classifiers are an ensemble of multiple DTs. RF classifiers accumulate the overall prediction from each tree that provides a probability for a particular output or classification [74]. Each tree in the forest is constructed differently from random subsets of the training data, referred to as ‘bootstrap aggregating’ or ‘bootstrap method’[68]. By randomly sampling the training data and building multiple decision trees, overall classifier performance improves. The advantages of random forest classifiers over decision tree machine learning algorithms are more stable predictions, reduced overfitting, and are more robust to input features [74].

Classification and regression trees combine classification of discrete variables with regression for dealing with continuous variables [75]. This is important when dealing with time-series data, or classes (i.e., gait phases) that occur in sequence. The main difference is that these trees employ linear or logistic regression models at terminating nodes, and perform a further step in predicting correct outputs [76].

2.5.4 **Machine Learning Gait Phase Recognition Systems**

Machine learning approaches related to human activity recognition has becoming increasingly popular [1], [2], [42], [47], [49]–[52], [58]–[60], [77]–[79]. Unfortunately, documentation regarding the use of machine learning-based gait phase recognition for real-time orthosis-control systems is lacking.

Although most machine learning techniques used for optimal accuracy are supervised procedures, unsupervised techniques can also be considered. Hidden Markov Models (HMM) are particularly known for their applications in speech, handwriting, and gesture recognition [80], [81] and offer a degree of external supervision through hidden states. For instance, Mannini et al. [81] used a Gaussian continuous HMM classifier with supervised learning to distinguish between static postures, walking, running, cycling and stair climbing with data gathered from accelerometers and gyroscopes placed on the body. High frequency components were indicative of dynamic motions whereas the low frequency components were exploited to identify static postures such as sitting, standing or lying down. Their results demonstrated that Markov modelling is advantageous for HAR, yielding 99.1% overall classification accuracy [81]. In addition, [54], [82] also exploited the power of HMM classification techniques to discriminate between different gait patterns. Taborri, et al. [54] used Hidden Markov Models for gait phase detection, segmenting a sensor signal into flat foot, heel off, swing, and heel strike phases.

Jung, et al. [77] proposed a neural network based gait classification technique that use inertial sensors and force sensitive resistors placed on a lower limb robotic exoskeleton to provide limb orientation and angular velocities. This study involved two neural network classifiers (MLP and nonlinear autoregressive with external inputs (NARX)). Training set contained 78 strides, whereas the test set contained 38 strides. Their overall classification outcome for gait phase detection depended on three misclassification criteria: classification considers the gait phase changed before the true gait phase, classification considers the gait phase changed after the true gait phase and a false positive, where the classification detected a change in gait phase when there was none. Results of this study yielded high performance for the offline case, with a classification rate above 97% for both techniques.

2.5.5 **Feature Selection**

Feature selection techniques sift through signal features and assess which ones provide high feature space discriminability and positively contribute to classification accuracy. Multiple features may provide the same information, decrease the overall classification accuracy, and allow a classification model to be over-fit, and therefore not generalizable to new information. Feature selection techniques include filter, wrapper, and embedded methods [50], [83], [84]. These methods work differently by combining, structuring, and organizing various features in compliance with classification models.

2.5.5.1 Filter Methods

Filter techniques evaluate features by observing the inherent properties characterized by the feature space [83], assessing feature relevance by the feature space's level of discriminability between classes. Therefore, if a feature space has highly separable class distributions, filter selection would rank that feature higher than a less separable feature. Filter techniques are easily scalable to high-dimensional feature sets, computationally efficient, and independent of the classifier [50], [83]. Most filter methods are univariate, evaluating one feature at a time, thereby ignoring feature dependencies that could lead to worse classification performance [83]. However, more recent filter feature selection techniques are multivariate, such as Correlation-based Feature Selection (CFS) [84], [85] and Relief-F [84], [86]. A disadvantage of filter methods is that all features are ranked prior to being fed to the classifier, thereby ignoring potentially relevant interactions that work well with a particular classifier. Filter methods used in HAR include Relief-F, CFS, and Fast Correlation-Based Filter (FCBF) [46], [47], [49]–[51], [85], [87], [88].

For Relief-F, each feature is given a weight that is updated with each additional data point, according to an evaluation function given by (2.1) [50]. Relief-F implements ‘k’ nearest neighbors for investigations on class ranking reliability and stability (probability estimation for nearest miss and nearest hits) for multiclass problems. Specifically,

$$w_i = \sum_{j=1}^{N} (x_i^j - \, nearmiss(x^j)_i)^2 - (x_i^j - \, nearhit(x^j)_i)^2 \tag{2.1}$$

where w_i represents the relevance weight for the i^{th} feature, x_i^j denotes the i^{th} feature for data point x^j, N is the total number of data points, $nearhit(x^j)$ and $nearmiss(x^j)$ is the nearest data point to x_i^j for the same and different class, respectively [50]. Features attaining higher relevance weight scores are considered higher quality. The computational cost of Relief-F is in the number of times the feature evaluation function is called, $O(M+1)$ [50], where M denotes the total number of features.

Correlation-Based Feature Selection (CFS) uses a heuristic approach to evaluate feature effectiveness or quality, taking into account a feature’s ability to predict certain classes along with assessing the inter-correlation between multiple features [87]. CFS is based on the premise that a good feature subset is highly correlated or predictive of a class, however is uncorrelated or not predictive of each other [87]. CFS is a simple filter that ranks features based on the heuristic evaluation function given by [84]

$$M_s = \frac{k\,\overline{r_{cf}}}{\sqrt{k + k(k-1)\,\overline{r_{ff}}}} \tag{2.2}$$

where M_s represents the relevance score or merit of a feature subset S containing k features, and $\overline{r_{cf}}$ and $\overline{r_{ff}}$ denote the mean feature-class and feature-feature correlations, respectively. In this evaluation function, the numerator describes how well features predict a certain class and denominator shows redundancy among features [84], and suggests a feature subset where each feature is important for the classification task.

2.5.6 **Classifier evaluation**

Classifier evaluation is performed by examining performance metrics. Classifier performance metrics can be defined as:

- True positive (TP): Occurrence of the instance is correctly classified
- False positive (FP): Occurrence of the instance is incorrectly classified
- True negative (TN): Non-occurrence of the instance is correctly classified
- False negative (FN): Non-occurrence of the instance is incorrectly classified

Table 2-3 shows typical classification evaluation metrics include accuracy, sensitivity, specificity, precision, f-score, and Matthew's correlation coefficient (MCC) [68]. Depending on the classification method, classifier-dependent evaluation metrics may be used. For example, the tree size would be a considerably important metric for DT.

Table 2-3 Classifier evaluation metrics

Outcome	**Formula**
Accuracy	$\frac{(TP+TN)}{TP+FP+TN+FN}$
Sensitivity	$\frac{TP}{TP+FN}$
Specificity	$\frac{TN}{TN+FP}$
Precision	$\frac{TP}{TP+FP}$
F-score	$\frac{2TP}{2TP+FN+FP}$
Matthew's Correlation Coefficient	$\frac{TP\,x\,TN - FP\,x\,FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$

There are many different techniques associated with classifier evaluation and validation. Machine learning classifier training and testing techniques include:

- Holdout method: Partitioning the original data set into training set and testing set, common proportions are 70%/30% (training/testing)

- *k*-Fold Cross Validation (*k*-FCV): Partitioning the original data set into *k* segments, leaving one segment as the test set and training on the rest, repeating the holdout method for each fold
- Leave-one-subject-out cross validation: Training on all subjects and leaving one out to test on, repeating for all subjects

2.6 Literature Review Summary

This literature review included a description of normal gait, sub-phases, the importance of knee joint function during locomotion along with a general overview of external knee flexion and extension moments for stability, and causes and effects of knee extensor weakness. This showed that co-activation of antagonist muscle groups (quadriceps and hamstrings) are essential for weight bearing during gait.

Orthotic solutions were discussed, including KAFO, SCKAFO, and M-SCKAFO. The main goals of these assistive devices are to promote natural gait, provide support during stance, and allow for unhindered knee motion during swing. Microprocessor-controlled walking devices are the most functionally advanced in terms of stance-control reliability and uneven terrain navigation. However, most are big, bulky, and costly. Currently, the best M-SCKAFO on the market is the Otto Bock C-brace. However, drawbacks include its size, cost, and the need to be central fabricated, thereby removing capacity for orthotist-customizability and end-user personalization.

The device's control system review demonstrated that GPR is crucial for effective stance-control orthoses. Studies suggest that machine learning can be more effective for GPR as compare with conventional rule-based techniques.

3. Gait Phase Detection from Thigh Kinematics using Machine Learning Techniques

This chapter addresses objective 1 by determining whether machine learning classifiers are able to accurately identify gait phases from level ground thigh and knee kinematics. Four different machine-learning classifiers were applied to the classification task using 10 signal features. This chapter describes the feasibility of using machine learning intelligence algorithms for signals that mimic those of a local sensor system at the thigh and knee. The study also investigated performance using only thigh kinematics. The preliminary analysis demonstrated effective GPR performance feasibility for real-time orthosis-control.

This paper was presented at the 2017 IEEE International Symposium on Medical Measurements and Application (MeMeA).

Farah, JD, Lemaire, ED, Baddour, N, “Gait Phase Detection from Thigh Kinematics using Machine Learning Techniques”, IEEE International Symposium on Medical Measurements and Application (MeMeA), Rochester, Minnesota, USA, May 2017. doi:10.1109/MeMeA.2017.7985886.

3.1 Abstract

Intelligent orthotic devices require accurate detection of gait events for real-time control. For orthoses that control the knee, an ideal system would only locate sensors at the thigh and knee, thereby facilitating sensor and electronics integration with the assistive device. To determine potential gait phase identification approaches, classification was implemented using J-48 Decision Tree, Random Forest, Multi-layer Perceptrons, and Support Vector Machine classifiers, along with 5-fold (5-FCV) and 10-fold cross validation (10-FCV). Knee angle, thigh angular velocity, and thigh acceleration were obtained from 31 able-bodied participants during walking (10 strides each). Strides were segmented into Loading Response, Push-Off, Swing, and Terminal Swing and features were extracted using a 0.1-second sliding window. Gait phase classification was performed with and without the knee angle parameter. J-48 Decision Tree with the knee angle parameter was ranked the best classifier due to its second highest classification accuracy of 97.5% and lowest mean absolute error of 0.014. Results without the knee angle parameter differed by only 0.5% and 0.003. Therefore, an inertial sensor with accelerometer and gyroscope output, located at the thigh, is a viable approach for classifying gait phases for intelligent orthosis control.

3.2 Introduction

Walking is a complex and dynamic system and an important daily living activity. Quantifying walking parameters by gait phase is necessary for deeper understanding of human locomotion and assistive device control. A stride can be divided into stance and swing phases, separated by initial contact (IC) and foot-off (FO) events (Figure 3.1) that represent specific functions throughout the walking cycle [3]. Stance and swing phases are subsequently divided into loading response and terminal swing sub-phases. Gait patterns are characterized by kinetic and kinematic parameters associated with body or body-segment rotations and translational motion in planes of progression.

Gait analysis has required complex systems, such as three-dimensional motion capture and force-plate laboratory systems. Recently, ubiquitous modalities have been employed using inertial measurement units (IMU). IMU's typically output angular velocity and acceleration signals and provide useful information for gait analysis [54], [56], [77], [89], [90].

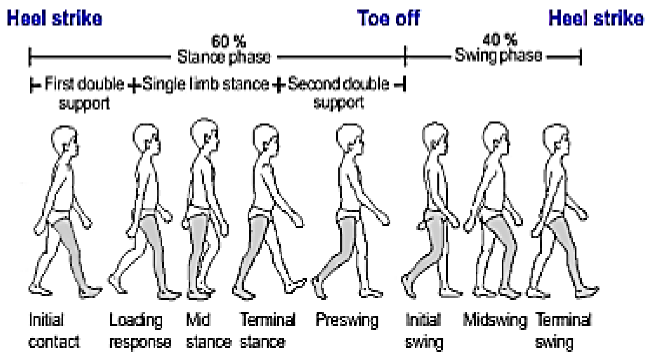

Figure 3.1: Gait cycle phases and events.

Pathological gait caused by knee extensor muscle weakness can inhibit an individual's ability to support themselves during stance. Knee-Ankle-Foot Orthoses (KAFO) are prescribed to provide support to the knee for persons suffering from knee extensor weakness. Unfortunately, KAFO inhibit knee flexion, causing users to develop irregular gait patterns as compensatory mechanisms such as hip-hike, lateral sway, and vaulting [9], which can be chronically detrimental to the body. Stance-Control-Knee-Ankle-Foot Orthoses (SCKAFO) have mechanical knee joints that are engineered to provide knee support during stance and permit free knee movement in swing. SCKAFO improve mobility and offer more natural gait compared to conventional KAFO [9].

Microprocessor controlled SCKAFO (M-SCKAFO) take advantage of various sensor systems to control the knee joint (ex. pressure sensors, IMU's, EMG [90]) and are more reliable compared to mechanically-controlled SCKAFO that require pre-set movements or knee extension moments to lock and unlock the knee [9]. Microprocessor controlled joints employ different control systems that use machine learning techniques [54], [59], [77] or rule-based control systems [56], [89], [90]. However, many M-SCKAFO require multiple sensors at the thigh, shank, and foot for effective stance-control. As a result, these walking-aid devices become increasingly expensive and difficult to personalize. An ideal control system would localize sensors at the thigh and knee, thereby allowing orthotists freedom to customize the shank and foot segments and possibly reduce overall system cost.

Integrating local sensor signals with machine intelligence algorithms should provide an effective method for gait phase identification that is feasible for application in real-time control. The objective of this paper is to determine whether accurate gait phase detection can be implemented using J-48 Decision Trees (J-48 DT), Random Forest (RF), Multi-Layer Perceptrons Neural Network (MLP), and Support Vector Machine (SVM) classifiers strictly using thigh-segment kinematics and knee angle during walking. This research also investigated thigh-only sensor localization by isolating thigh-segment kinematics from the knee angle parameter.

3.3 Methods

3.3.1 **Equipment**

All walking signals were acquired with CAREN-Extended system (Computer Assisted Rehabilitation Environment) at The Ottawa Hospital Rehabilitation Centre. CAREN-Extended includes a 6 degree-of-freedom platform with integrated force measuring dual-tread treadmill, 180º screen to project 3D virtual worlds, safety railings, and Vicon 3D motion capture system. A custom “walk through a park” application was developed and used in this study.

A full body marker set was used to track 3D kinematics as participants walked on level sections within the 3D world. Visual3D (C-Motion) was used for all biomechanical analysis. Kinematic data were recorded at 100Hz and ground reaction forces (GRF) were recorded at 1000Hz.

3.3.2 **Participants**

A de-identified dataset from 31 able-bodied volunteers was used for this study (16 males, 15 females, mass=75.8±13.2kg, height=1.73±0.12m, age=30±10years). Each participant walked at their own pace (self-paced).

3.4 Signals

For this study, thigh-segment kinematic data and knee angle for self-paced level ground walking were imported into Visual3D for analysis. Parameters for gait phase identification were knee angle (KA), thigh angular velocity (AngVel), and thigh center of gravity acceleration (Acc) for the *x, y* and *z* axes. These kinematic parameters are similar to signals measured with an IMU placed on the thigh and a knee angle sensor. A 4^{th} order low pass Butterworth filter with 10Hz cut off frequency was applied to each signal.

Gait cycles and gait events from all participants were imported into MATLAB for further processing. Ten strides were selected from each individual's walking trial, resulting in 310 strides for machine learning implementation (Figure 3.2). In addition to signals acquired from the *x, y* and *z* axes, the resultant was used as a supplementary signal in the training data set.

3.5 Pre-Processing

3.5.1 **Gait Phase Segmentation**

GRF were used to determine when the foot was in contact with the ground for each gait cycle. IC events, the instant the foot comes into contact with the ground directly after the swing phase, were the times when GRF passed a >50N threshold. FO events, the instant the foot leaves the ground indicating the onset of swing phase, were the times when the GRF passed a <50N threshold.

IC and FO gait events were used to calculate the Centre of Mid-Stance (MS) event (half the time between IC and FO). The remaining gait event was the time at which maximum knee flexion (MKF) was reached during swing phase, which is associated with foot clearance. The MKF event was computed prior to data segmentation as the maximum knee flexion angle in the resultant KA (Figure 3.2A) signal after FO. Since stride time varies during self-paced walking, gait events occur at different instants for every stride. Therefore, each stride required its own set of gait events. IC, MS, FO, and MKF were calculated to enable gait phase partitioning, similar to the classes defined by Taborri, et al. [54].

The gait events were used to partition each signal into four classes for gait phase identification. Load Response (P1) was from IC to MS, Push Off (P2) from MS to FO, Swing (P3) from FO to MKF, and Terminal Swing (P4) from MKF to IC of the following stride (Figure 3.3). These four phases are ideal for optimal SCKAFO flexion resistance control, permitting free knee flexion during swing and appropriate flexion resistance during stance.

3.5.2 **Features**

Eleven signal features were selected for this study (Table 3-1). The first ten features were computed for the *x, y, z* and resultant signal for KA, thigh AngVel and thigh Acc. This produced 120 features for each stride. The Bfourn feature was a single value for KA, thigh AngVel and thigh Acc; therefore, 123 features were generated. Eighty-two features were selected for classification *without* the knee angle parameter.

3.5.3 **Window Size**

Mode switching must be fast and accurate to provide support during stance and uninhibited knee motion during swing. Winter [13] reported that natural cadence for level ground walking is 105.3 steps/min. This translates to approximately one second per gait cycle. Individuals using a SCKAFO will likely walk at slower cadence. Average values for slow cadence were reported as 86.8 steps/min (1.38 seconds per stride) [13]. In the gait cycle, weight acceptance during stance is approximately 12% of the gait cycle [3], and corresponds to 0.14s for natural and 0.16s for slow cadence. Therefore, the device must switch between open and stance resistance settings in less than 0.1s to provide reliable control. Small time constraints allow for effective decision-making for real-time SCKAFO control. A trade-off between reliability, accuracy, and rapid classification must occur for effective gait phase classification.

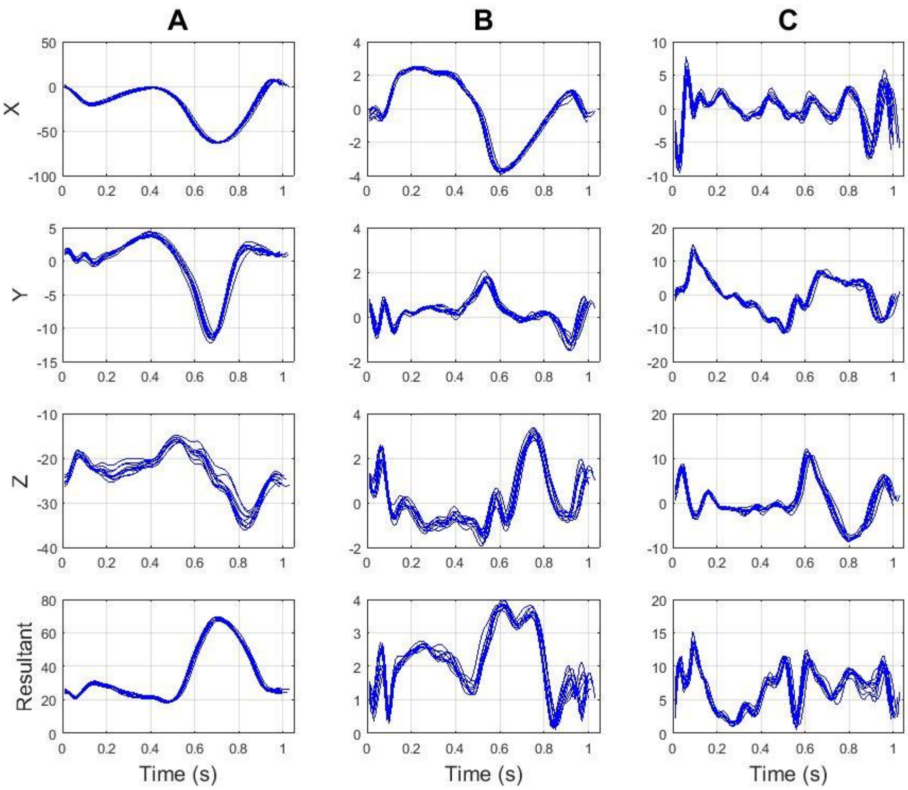

Figure 3.2: Ten strides from one participant showing knee angle (º) (A), thigh angular velocity (rad/s) (B), and thigh acceleration (m/s2) (C) for x, y, and z axes and resultant.

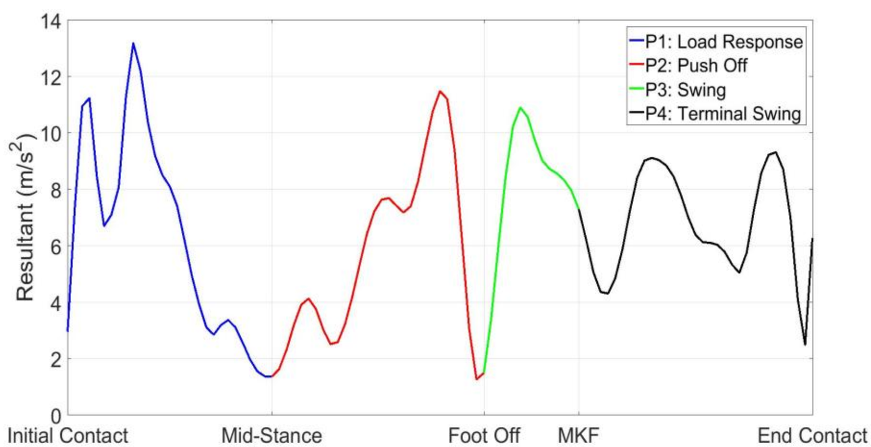

Figure 3.3: Thigh-segment resultant acceleration (m/s^2) segmented into gait phases: P1: Loading Response (blue), P2: Push-Off (red), P3: Swing (green) and P4: Terminal Swing (black).

Table 3-1: Features for gait phase classification

Feature	Mathematical Formula
Mean	$\mu = \frac{1}{N}\sum_{i=1} A_i$
Variance	$V = \frac{1}{N-1}\sum_{i=1}^{N} \| A_i - \mu \|^2$
Standard Deviation	$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N} \| A_i - \mu \|^2}$
Range	$R = \max(A) - \min(A)$
Skewness	$s = \frac{\varepsilon(x-\mu)^3}{\sigma^3}$
Sum	$S = \sum_{i=1} A_i$
Integral	$\int_a^b f(x)\,dx \approx \frac{b-a}{2N}\sum_{i=1}^{N}(A_i + A_{i+1})$
Derivative	$D = max(A_{i+1} - A_i)$
Energy	$E = \sum_{i=1} A_i^2$
Power	$P = \frac{1}{t}\sum_{i=1} A_i^2$
Bfourn	$B_f = \sum_i x_i y_i + y_i z_i + x_i z_i$

3.5.4 **Data Set**

Features were computed for 310 strides over a 0.1s sliding window. The window incremented by 0.01 seconds (a single data point) for each iteration. Computed features for each window segment were annotated to a data file according to the gait phase class. The designated class depended on the last data point in each window. For example, a window containing 10 data points, with the first nine labeled as P1 (Loading Response) and the tenth point labeled as P2 (Push-Off), would then be allocated to the P2 class (Figure 3.4). Each window was counted as an instance and compiled into a feature matrix containing the complete data set. The final set contained 30,060 instances over the four classes (P1, P2, P3 and P4).

3.6 Machine Learning Classification

Feature matrices were imported into the Waikato Environment for Knowledge Analysis (WEKA) for machine learning classification. The feature vector containing the class labels were converted to nominal values. Class distributions and feature spaces were visually examined in WEKA. Figure 3.5 depicts 16 two-dimensional feature spaces for eight different features, from a single participant (10 strides).

Four machine learning classifiers were used for gait phase identification: J48 Decision Tree, Random Forest, Multi-Layer Perceptrons, and Support Vector Machine classifiers. Default parameters were used for each classifier.

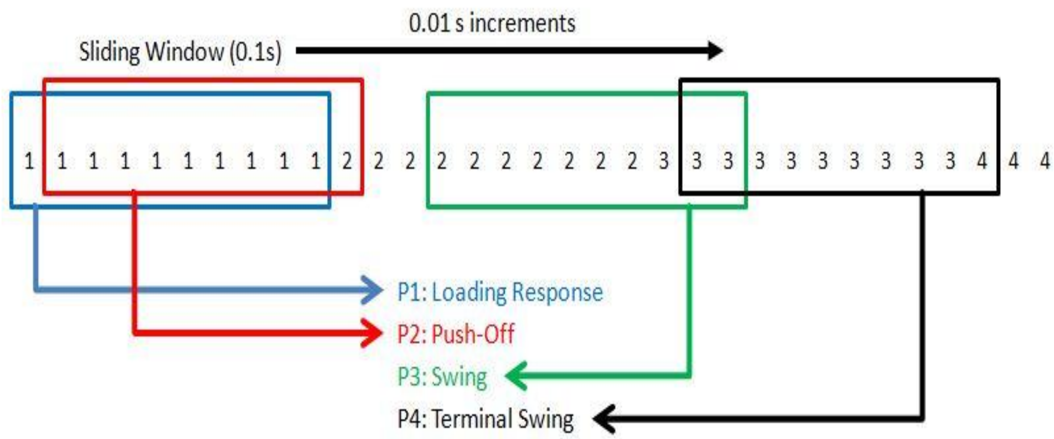

Figure 3.4: Sliding window containing 10 data points (0.1s) incrementing one data point for each iteration (0.01s). Features were extracted for each window and allocated to the class (gait phase) depending on the last data point in each window. The red window has a single data point belonging to class P2 (Push-Off); therefore, features for this window belong to the Push-Off phase.

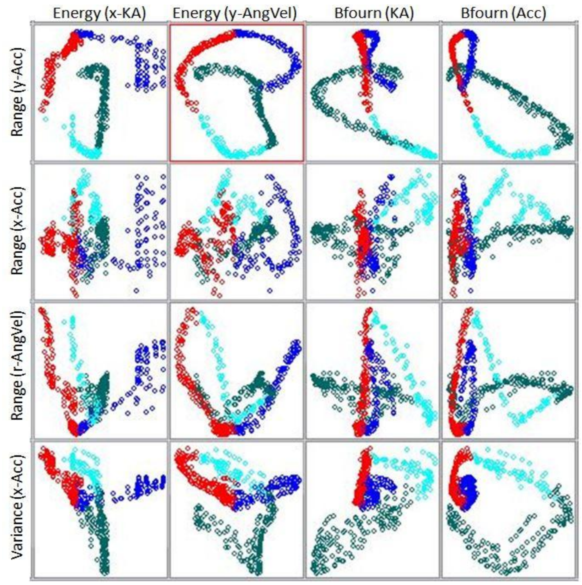

Figure 3.5: Two dimensional class distributions for eight computed features from a single participant (10 strides). Loading Response (blue), Push-Off (red), Swing (Cyan), and Terminal Swing (teal).

3.7 Training & Testing

Each classifier was trained and tested using 5-Fold Cross Validation (5-FCV) and 10-Fold Cross Validation (10-FCV). Cross validation was implemented using a stratified sampling technique to determine each training and testing set for each fold. Gait phase was classified using two feature sets, the total data set and a feature set without the KA parameter.

Classifier performance metrics were overall classification accuracy and mean absolute error (MAE). Confusion tables were also generated to analyze precision, true-positive rates (TPR) (Recall or Sensitivity) given by (3.1) and false-positive rates (FPR) (1 – Specificity) given by (3.2).

$$TPR = \frac{TruePositive}{TruePositive + FalseNegative} \tag{3.1}$$

$$FPR = \frac{FalsePositive}{FalsePositive + TrueNegative} \tag{3.2}$$

All performance metrics were calculated using WEKA. MATLAB was used to perform one way analysis of variance (ANOVA) to test for significant difference between 5-FCV and 10-FCV techniques, for cases *with* (Case 1) and *without* (Case 2) the KA parameter.

3.8 Results & Discussion

3.8.1 **Data Set**

Table II shows the percentage of correctly classified instances and Table III shows the MAE associated with each classification model and case. Willmott and Matsuura [91] suggested that MAE offers advantages over RMSE for average error measurements. MAE measures the average magnitude of errors in each set and describes the mean of the absolute values for each individual misclassification, over all instances. MAE assigns equal weights for all misclassified instances. For example, with respect to the four gait phases, misclassifying P1 for P2 will produce the same error as misclassifying P1 for P3. In terms of gait phase classification, and application for real-time control, every misclassified instance is important; thus, using MAE provides an estimate for all misclassified instances over the entire classification model.

All classification results had accuracies greater than 96.4% and MAE less than 0.253. Results from 5-FCV and 10-FCV were similar, with classification accuracies differing by only 0.1 ± 0.1% and no significant differences ($p > 0.8$). This helps support the viability of these results for gait phase classification for self-paced level ground walking.

The SVM classifier had the lowest average accuracies and highest MAE in both cases for each validation test method. SVM's are inherently binary classifiers and are generally used to distinguish between two specific categories, or class memberships. Training SVM's consists of finding optimal hyperplanes to discriminate between classes in a feature space. Their complexity is characterized heavily on the number of support vectors used for calculating optimal decision boundaries and can lead to an over-fit system. In this study, the SVM was set to find non-linear hyperplanes using kernel functions that work by mapping training data to higher dimensional feature spaces. However, if there is poor class separation, the SVM classifier can have difficulty distinguishing between classes or distinguishing one class more readily than others. This is the likely cause for low accuracies and high MAE in this study.

Table 3-2: Classification accuracies

Classifier	With Knee Angle		Without Knee Angle	
	5-FCV	**10-FCV**	**5-FCV**	**10-FCV**
J48-DT	97.5	97.6	97.0	97.1
RF	98.3	98.4	98.2	98.3
MLP	96.4	96.5	96.7	96.9
SVM	96.7	96.7	96.4	96.4
Average	97.2	97.3	97.1	97.2
SD	0.9	0.9	0.8	0.8

Table 3-3: Mean absolute errors

Classifier	With Knee Angle		Without Knee Angle	
	5-FCV	**10-FCV**	**5-FCV**	**10-FCV**
J48-DT	0.014	0.013	0.017	0.016
RF	0.023	0.022	0.023	0.022
MLP	0.020	0.019	0.018	0.018
SVM	0.253	0.253	0.253	0.253
Average	0.077	0.077	0.078	0.077
SD	0.117	0.117	0.117	0.117

The MLP neural network classifier generated low MAE in both cases. The MLP used here was generally more complex than the decision tree and RF classifiers. An advantage of MLP's are their ability to build non-linear fits [77], however the computational time to build an MLP classification model was far greater than a single decision tree or decision tree forest. The main reason for the low MAE is likely due to high number of hidden nodes used in this study. WEKA calculates the number of hidden nodes as half the total number of features plus the number of classes, resulting in 64 for Case 1, and 43 for Case 2. A high number of hidden nodes leads to significantly higher classifier complexity and exaggerated training times [77].

The RF classifier produced the best accuracies and J48-DT produced the lowest MAE. For gait phase identification and real-time SCKAFO control, high accuracy and low misclassification are essential. Since the average accuracy and MAE differences between RF and J48 were only 0.1 ± 0.1% and 0.007±0.001, respectively, both models produced excellent gait phase identification and can be considered for the SCKAFO control application. However, since gait phase misclassification could result in a SCKAFO user's stumble or collapse, even small differences must be considered. For example, in a day consisting of 5000 steps, the difference between misclassified gait phases using the RF classifier versus the J-48 DT classifier is 40 steps for Case 1 and 60 steps for Case 2, which represents 40-60 opportunities for a stumble due to misclassification.

The RF classifier used in this study was built using 100 decision trees with randomly chosen samples, whereas a single decision tree classifier yielded comparable classification accuracies with fewer misclassified instances. Despite the high accuracies obtained by RF, a single decision tree is less computationally intensive than the 100-tree RF classifier. Therefore, one decision tree may be more efficient for real-time decision-making for SCKAFO change-of-state. Table IV shows the confusion matrix for the J-48 DT classifier. The highlighted diagonals show the number of correctly classified instances, whereas off-diagonals show misclassified instances. For example, 9196 instances were correctly classified out of 9288 total instances labelled as Terminal Swing (P4).

Table 3-4: J-48 Decision tree confusion matrix (10-FCV)

	Classified As			
Class	**P1**	**P2**	**P3**	**P4**
P1	6891	144	2	4
P2	138	9633	125	1
P3	1	138	3611	84
P4	9	4	79	9196

Table 3-5 shows exceptional TPR, FPR, and precision results when classifying gait phases computed from the confusion matrix using the J-48 DT classifier. The J-48 DT classifier performed the best for classifying Terminal Swing, followed by Loading Response, Swing, and Push-Off. The performance results between classes for a single classifier depend heavily on class distribution discriminability. Features extracted from stride segments are more variable for one class than another. Equation (3.3) shows ensemble averages for the resultant accelerometer signal from two participants. Both stride sets show that the signal varies within individuals at different locations in the curve. Coefficient of Variation (C_v) measures the relative standard deviation of a data set, given by

$$C_v = \left(\frac{\sigma}{\mu}\right)100 \tag{3.3}$$

where μ is the mean and σ is the standard deviation. C_v is equal to 15.58% (Figure 3.6A) and 9.01% (Figure 3.6B), suggesting increased variability between these two sets. Arrows point to high signal variability between these two data sets.

The J-48 DT, RF, MLP, and SVM classifiers in this study were not optimized. GPR was implemented with equal weights for identifying each class (i.e., if in P1, change of state can be classified as P1 and any other of the three classes).

Table 3-5: J-48 Decision tree classifier performance

Class	**TP Rate (Recall)**	**FP Rate**	**Precision**
Loading Response	0.98	0.01	0.98
Push-Off	0.97	0.01	0.97
Swing	0.94	0.01	0.95
Terminal Swing	0.99	0.01	0.99

A transition-class vector, similar to Pappas et al. [56], may provide increased accuracy and decreased misclassified instances. For example, if prior knowledge regarding the present gait phase is known there would be a pre-set rule-based classification. This enables gait phase classification of certain classes that are more likely to happen given the current state of another. For instance, if the model knows the system is in LR, the model would only need to classify between LR and PO. This reduces the number of phases to be classified and makes the model intrinsically binary for continuous phases.

Furthermore, feature correlation was not investigated. Some features in this study may have been redundant, have poor class discriminability or have high data dimensionality, and consequently, have affected classifier performance. Therefore, adjusting classifier parameters and incorporating feature selection techniques (i.e., Relief-F, CFS, FCBF, Sequential Forward Selection [88], [92], [93]) could improve classifier performance.

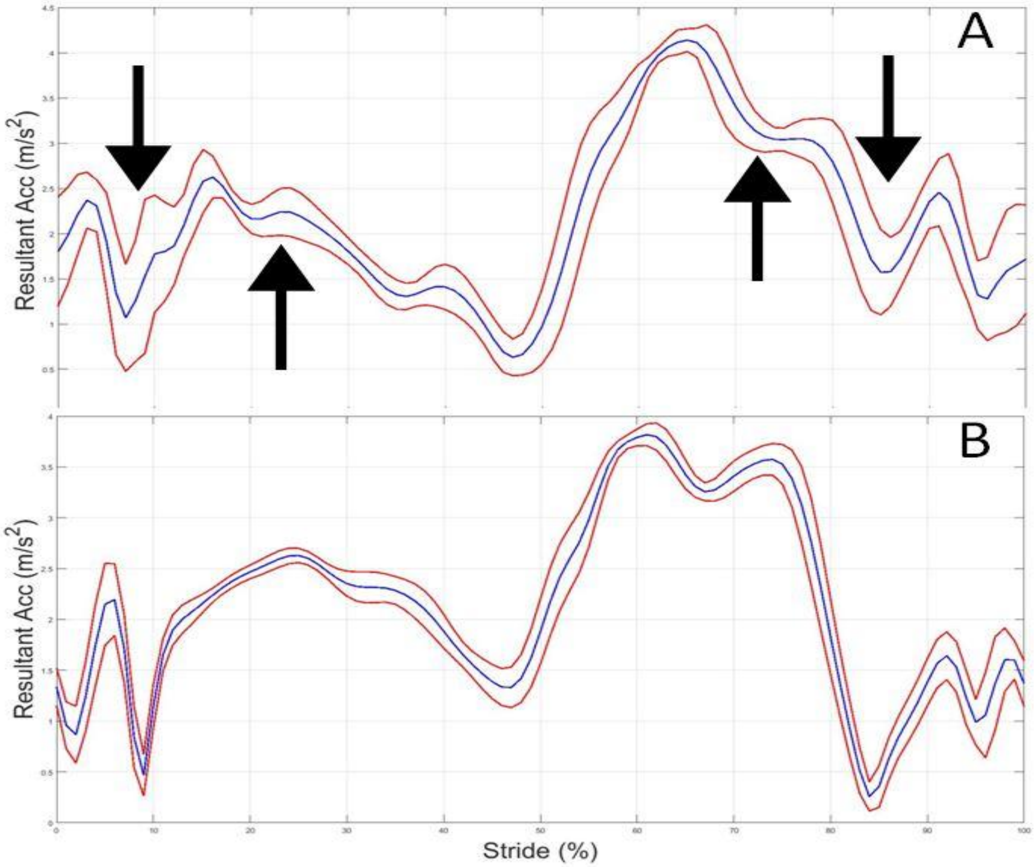

Figure 3.6: Thigh-segment angular velocity resultant curves for two participants as a percent of stride. Arrows show locations of greater variability.

3.8.2 **Sensor Localization**

Gait phase classification for Case 2 achieved marginally lower accuracies than Case 1, excluding 5-FCV using MLP (likely due to using too many hidden nodes). Therefore, restricting the sensor signals to just thigh AngVel and thigh Acc may be feasible for gait phase classification in practice. Increased sensor localization, obtained from only two signals that are obtainable from one IMU sensor, increases modular-unit feasibility. However, these results may vary when applying the classification models to slower walking speeds (0.4-0.8 m/s) and various gait modes (i.e. stair ascent/descent, up-, down-, right- and left-sloped surfaces). The KA parameter may improve classification in these scenarios since KA is known to vary for non-level and slow walking.

3.9 Conclusion

Gait-phase classification, implemented from a localized sensor system (i.e., thigh-segment kinematics and knee angle) produced high accuracies and low mean absolute errors. All classifiers performed well for self-paced level ground walking. In cases with the knee angle parameter, classifier performance increased. This provides supporting evidence that machine intelligence is viable for real-time control of a stance-control knee-ankle-foot orthosis with a localized sensor system. Further investigations with transition-vectors, other machine learning classifiers, and feature selection should be explored for future studies. Orthosis control research would also benefit from gait phase classification analysis that includes various walking speeds and different walking modes such as stair ascent/descent, thereby improving model robustness and suitable for online testing.

4. Local Sensor Gait Phase Recognition System for Applications in Orthosis-Control

This chapter addresses objective 2 by evaluating the model and feature subset that provides the best performance metrics for GPR. A data set containing signal features from thigh and knee kinematics with multiple walking speeds and surface conditions was investigated through Relief-F and CFS feature selection techniques to reduce data dimensionality and feature redundancy. New features were investigated that were not explored in chapter 3. Various training models were also explored to determine a generalizable classification model.

The content of this chapter was submitted for publication to the IEEE Transaction of Neural Systems and Rehabilitation Engineering:

Farah JD, Lemaire ED, Baddour N, Feature Selection for Local Sensor System Gait Phase Identification for Application in Orthosis-Control, IEEE Transaction of Neural Systems and Rehabilitation Engineering, submitted December 2017.

4.1 Abstract

Effective gait phase identification is essential for optimal control of intelligent assistive walking devices. For stance-control orthoses, an ideal system would integrate sensors at the thigh and knee, and be sufficiently fast, accurate, and robust to handle multiple walking speeds and ground-level conditions. A J-48 Decision Tree (J-48 DT) machine learning classifier for gait phase recognition (GPR) was built and tested for various ground conditions (level (LG), up (US) and down (DS) slopes, right (RS) and left (LS) cross-slopes) and walking speeds (self-paced, 0.8 m/s, 0.6 m/s, 0.4 m/s). Three axes for thigh-segment angular velocity, thigh-segment acceleration, and knee angle were obtained from 10 strides from 31 able-bodied participants. Strides were segmented into loading response, push-off, swing, and terminal swing phases, and 187 attributes were extracted using a 0.1-second sliding window. Feature selection was performed using Relief-F and Correlation-Based Feature Selection (CFS). The J-48 DT was trained using four model types (full data set, LG, LG-DS-US, LG-RS-LS) using feature subsets computed from feature selection. Classification performance for each model was evaluated with number of features, tree size, number of leaves, classification accuracy, sensitivity, specificity, precision, F-score, and Matthew's correlation coefficient. All models were tested using a full data set containing all conditions, and 5-fold cross-validation was performed on the full model type for Relief-F and CFS and compared with feature selection implemented through J-48 DT construction (J48-FS). No statistical differences were found between J48-FS and Relief-F ($p>0.4$) and CFS ($p>0.4$) feature subsets. Results showed that the CFS with the LG-DS-US model reduced the number of features to 22, with 98.10% classification accuracy. CFS with the 5-FCV Full model achieved 97.26% classification accuracy, with 20 attributes. A novel GPR model using a local sensor system for future applications in stance-control knee-ankle-foot orthoses was presented. Thigh and knee sensor signals, integrated with machine intelligence algorithms, are viable for real-time gait phase classification in orthosis-control applications to enable safe mobility across real-world walking scenarios.

4.2 Introduction

Knee-Ankle-Foot Orthoses (KAFOs) are lower extremity assistive devices that provide support for individuals with knee-extensor muscle weakness [1], [9], [18]–[21]. Stance-Control KAFOs (SCKAFOs) are engineered to prevent the knee from collapsing during weight-bearing and permit free-knee movement during swing, thereby enabling more natural gait compared to conventional passive KAFOs [9], [18], [19], [28].

Emerging microprocessor-controlled SCKAFOs (M-SCKAFOs) use multiple sensor systems to guide knee control [1], [19] and can provide enhanced function across daily walking activities. Integrating electronic sensors such as pressure sensors and inertial measurement units (IMU) into the control system increases stance-control reliability over traditional SCKAFO approaches. The functionality of each SCKAFO heavily depends on the knee-joint's mechanical design, variable knee-flexion settings, and their control systems. The control system determines whether a user is in the stance or the swing phase then sends a signal to the knee joint system to switch to a locked, free, or partial resistance setting. Therefore, to enable appropriate function, activity classification and movement phase detection from wearable sensor data becomes essential.

Gait Phase Recognition (GPR) models that use wearable sensors have been extensively researched for more practical means of gait analysis. Namely, wearables with embedded sensor technology; strain gauges [38], pressure sensors [57], electromyography (EMG) [21], force sensitive resistors [56], [59], [77], goniometers [16], [94], and inertial sensors (i.e., accelerometers, gyroscopes) [1], [40], [46], [56]–[60], [77], [78] can be combined to give cutting-edge gait analysis and GPR techniques. For example, Abahayasinghe, et al. [53] investigated a GPR system from a single fixed-IMU placed in the trouser pocket. They were able to identify many sub-phases associated with gait from thigh-angle waveform analysis. However, literature that directly links machine learning-based GPR models to lower limb orthosis-control is lacking. To our knowledge, there are no orthosis-control models on the market that use machine learning-based GPR.

GPR for a viable M-SCKAFO system must be fast and accurate for real-time application. M-SCKAFO systems have the potential to employ artificial intelligence algorithms that use machine learning-based GPR [1], [54], [59], [77] or rule-based phase detection [56], [89], [90]. Despite enhancing overall stance-control functionality, final products today are large and obtain overly complex systems with many sensors. Most M-SCKAFOs use rule-based algorithms and complex sensor systems located at the thigh, knee, shank, and foot. As a result, stance-control orthoses become increasingly expensive and difficult to personalize [1].

The Otto Bock C-Brace, and is currently the most versatile M-SCKAFO on the market. It includes a hydraulic knee joint unit that dampens knee extension and flexion, dictated by kinematic and kinetic sensor input at the knee, shank, and ankle [19], [20]. The C-Brace can accommodate non-level surfaces such as stairs and ramps; however, this orthosis is large (i.e., cannot fit under trousers) and requires central fabrication for manufacturing.

A modular knee joint with an embedded stance-control system requires GPR models that sensors that are local to the knee joint for future use in instrumented orthoses. Localizing the sensor system to the knee joint and thigh could result in a modular knee joint component that is lighter, manufacturable by the orthotist, and orthotist-customizable for the foot, shank, and thigh segments, thus enabling better component selection for the end-user. Without foot and shank signal inputs, localizing sensors to the thigh and knee makes GPR a challenging task.

The objective of this research was to determine if accurate GPR can be achieved using a machine learning classifier that can work within small data windows (for real-time classification) and that uses sensor signals from only the thigh and knee. To achieve this goal, the best set of signal features must be determined. Assessing feature relevance through feature selection techniques should provide better classification performance and increase computational efficiency for application to M-SCKAFO control. In addition, features selected from training models with various surface-levels will differ. Finally, a model trained on multiple surface-levels and walking speeds will be robust to experiences in real-world walking scenarios, and thus provide the greatest classification performance. A control system localized to the M-SCKAFO knee joint should provide more versatile orthosis-designs to enhance safe mobility across multiple surfaces with dynamic knee-flexion adjustments during gait. Modular unit designs will also contribute smaller orthosis-sizes and better personalization, providing improved comfort and aesthetic appeal, thereby leading to a better quality of life for people with lower extremity muscle weakness.

4.3 Methods

4.3.1 **Data Set**

All walking signals were measured and recorded in a Computer Assisted Rehabilitation Environment (CAREN-Extended) system. CAREN-Extended included a six degree-of-freedom platform with embedded dual-tread instrumented treadmill, 180^o screen to project a virtual world, and 12-camera Vicon motion analysis. A custom-built virtual park application provided a range of walking conditions including, level (LG), down slope (DS), up slope (US), right cross-slope (RS), and left cross-slope (LS). A six degree-of-freedom full-body marker set with 57 markers was used to define joint and body-segment kinematics [95]–[97]. 3D marker position data were collected at 100Hz and ground reaction forces were collected at 1000Hz using Vicon Nexus. All data were imported into Visual 3D (C-Motion) for biomechanical analysis, data synchronization, and gait event identification.

A de-identified dataset from 31 able-bodied participants was used for this research (16 males, 15 females, mass=75.8kg (SD=13.2), height=1.73m (SD=0.12), age=30years (SD=10). This study was approved by The Ottawa Hospital Research Ethics Board (20140825-01H). Each volunteer walked in the virtual park on LG, DS, US, RS, and LS platform orientations at a self-paced speed (1.33 m/s (SD = 0.04)), 0.8m/s, 0.6m/s, and 0.4m/s, resulting in 20 different walking conditions. For each walking condition, a 4th order low pass Butterworth filter with 10Hz cut off frequency was applied to the data and then thigh-segment angular velocity (AngVel), thigh-segment acceleration (Acc), and knee angle (KA) in mediolateral (*x*-axis), anterior-posterior (*y*-axis), and vertical (*z*-axis) orientations for the right limb were calculated in Visual3D. These signal parameters were chosen for our

GPR application since they are similar to signals obtained from an IMU that would be placed at the thigh and a knee angle sensor on an instrumented orthosis.

For each person, 10 strides were extracted for each condition, resulting in a data set with 6200 strides. Strides were labeled using pre-determined gait events (initial contact, mid-stance, foot-off, maximum knee flexion). These gait events were used to segment strides into four gait phases (Loading Response (LR), Push-Off (PO), Swing, Terminal Swing (TSw)) [1] for classification. Classification on these gait phases will indirectly identify gait events when a gait phase transition occurs. The data were then separated into different training sets with particular conditions and all speeds; Full: all conditions; level ground (LG); level ground, down-slope and up-slope (LG-DS-US); level ground, right and left cross-slope (LG-RS-LS). Training sets were then used to build J-48 DT models.

4.3.2 **Feature Extraction**

AngVel, Acc and KA data were imported into Matlab where signal resultants and features were calculated. Table I shows 20 pre-dominant features (PDF) in the time-domain and 4 features in the frequency-domain. These features have high levels of discriminability for human activity and would help with interpretation of IMU signals associated with GPR, enabling accurate classification for dynamic activities with varying IMU signals [46]–[51]. PDFs were computed for each 0.1-second sliding window with 90 percent overlap (i.e., incrementing by 0.01 seconds) [1].

Each PDF was calculated along each axis (*x, y, z, r*) for each signal parameter (KA, AngVel, Acc), where *r* denotes the resultant of *x, y, z* axes. Therefore, each PDF generated nine features for machine learning implementation. PDFs computed from the *y, z, r* KA signals were removed from the feature matrix because an M-SCKAFO knee angle sensor will only provide Sagittal-plane data, which corresponds to the *x*-axis knee angle from our gait dynamics.

Maximum cross correlation was calculated between resultant AngVel-Acc, generating one feature. The B-feature (B_f) is given by (4.1),

$$B_f = \sum_{i=1}^{N} x_i y_i + y_i z_i + x_i z_i \tag{4.1}$$

where x, y, and z are the axis channels from AngVel, and Acc. i is each data point in the sliding window. N denotes the total number of data points in the sliding window. This computation provided three features for the classification task. Three eigenvalues were obtained from the covariance matrix for x, y, and z acceleration. CAGH is the correlation coefficient between acceleration in the heading direction (y-axis) and acceleration along gravity (z-axis) [50].

The full feature set contained 187 features (Table 4-1). Each feature is named according to its pre-fix in brackets with subscript that defined the axis it was extracted from, and the parameter name (KA, AngVel, Acc). For example, the mean from the x-axis angular velocity is denoted as '$\text{mean}_{\text{xAngVel}}$'.

Table 4-1. Features for local gait phase classification

		# of Features
Time Domain	Mean	9
	Variance (var)	9
	Standard Deviation (std)	9
	Range	9
	Skewness (skew)	9
	Sum	9
	Root Mean Squared (rms)	9
	Min	9
	Max	9
	Inter Quartile Range (iqr)	9
	Energy	9
	Sign Sum (sign)	9
	Integral (Trapezoidal rule) (trapz)	9
	Max Rate of Change (diff)	9
	Kurtosis (kur)	9
	Peak Number (pk)	9
	Maximum Cross Correlation (xcorr)	1
	B_f	2
	Eigen Values (eig)	3
	CAGH	1
Frequency Domain	Spectral Energy (specEnergy)	9
	1st FFT Coefficient (fftmag)	9
	Principal Frequency (pfreq)	9
	Spectral Entropy (entropy)	9
	Total number of features	187

4.3.3 **Feature Selection**

The Waikato Environment for Knowledge Analysis (WEKA) was used for all feature selection and machine learning classification. Relief-F [98] and CFS [84] were chosen for feature selection and compared against the embedded feature selection algorithm for J-48 DT construction (J48-FS). The embedded feature selection technique associated with these trees follows an information-based heuristic that uses information entropy and dictates tree structure, along with pruning. Therefore, all features in the training set may not be used in the final model.

Relief-F was implemented with 10 nearest neighbors (10NN) instead of a full data set to investigate class ranking reliability and stability. Features attaining higher relevance weight scores were considered to be of greater quality. CFS assumes that a good feature subset has features that are highly correlated or predictive of the classes, but are uncorrelated or not predictive of each other [87]. CFS ranks features based on the heuristic evaluation function [84] and determines the best feature subset based on the individual predictability for each feature and level of redundancy.

In WEKA, the best feature subset output for CFS was based on the 'best first' parameter. The CFS algorithm searches through all the features by forward greedy hill climbing with a back-tracking feature [84]. Relief-F provides a score for all features and was ranked from best to worst for each model type. The highest ranking feature would be ranked 1^{st}, whereas the lowest ranking feature would be ranked 187^{th}. Classification accuracies were calculated by an iterative process (removing the lowest ranking features) and a classification model was built to determine the best feature subset based on Relief-F scores. Accuracies were obtained from training a J-48 DT and testing the model on the full set. Classification accuracies were computed 39 times: full feature set; feature subset containing 180 features by removing the seven lowest ranking features; 35 feature subsets with the next five lowest ranking features iteratively removed from each set, until five features were left; feature subset with two highest-ranking features; and feature subset with the highest-ranking feature.

4.3.4 **Classification**

All models and feature subsets were built and assessed using a J-48 decision tree that applies pruning and information gain to order features in the tree [99]. The J-48 DT machine learning classifier was chosen because it is powerful, easy to implement, able to handle noisy data, can be visualized for easy understanding and has the ability to work in real-time.

GPR models were constructed by training a J-48 DT on each model type (i.e., Full, LG, LG-DS-US, LG-RS-LS). The J-48 DT for the Full model type was trained and tested using five-fold cross validation (5-FCV) using stratified data splits for each feature subset obtained through feature selection. LG, LG-DS-US, and LG-RS-LS models were tested using the full data set (features for all walking conditions and speeds). For example, for each feature subset (Relief-F, CFS, J48-FS), model type LG decision trees were built by training strictly on level ground data and combined speeds (SP, 0.8, 0.6, 0.4 m/s) and then each of these three decision trees were applied to the full data set in the validation phase. This protocol was done for these models to incorporate ground level conditions that the tree was not trained on, and to determine model robustness.

4.3.5 **Evaluation**

The performance of each model type and feature selection method were evaluated by the number of features used, tree size, number of leaves, classification accuracy, sensitivity, specificity, precision, F-Score, Matthews Correlation Coefficient (MCC). Within each model type, the Wilcoxon rank sum test was used to assess significant differences between evaluation metrics achieved using each feature subset. Relief-F and CFS feature subsets were tested against the full training set (J48-FS) for all evaluation metrics excluding the number of features in the training set. Significant differences are defined for *p*-value less than 0.05. Statistical analysis was implemented in Matlab.

4.4 Results

Figure 4.1 shows the classification accuracies for model types as a function of number of features. A similar trend was observed for all four model types, where classification accuracies begin to level off and show minimal variation for subsets with 20 or more features.

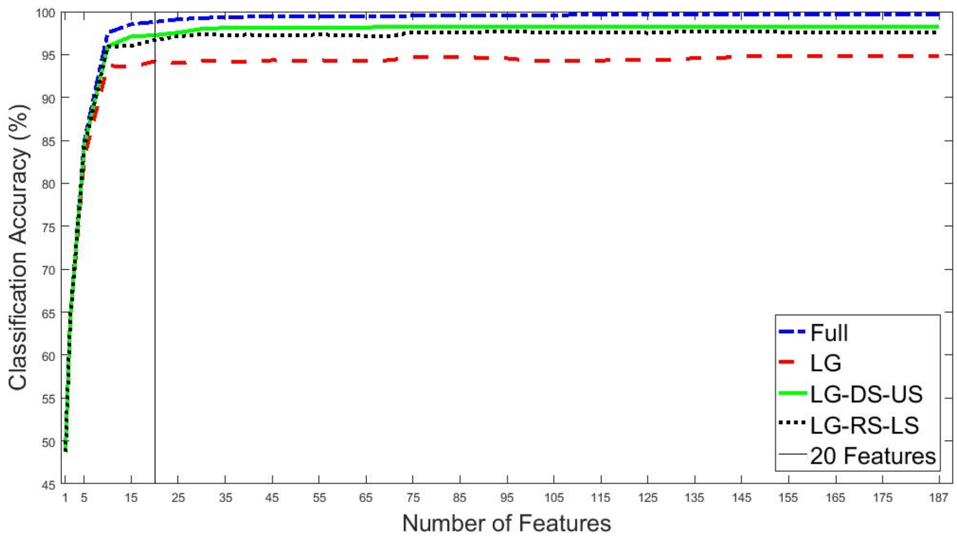

Figure 4.1: Classification accuracies as a function of number of features according to Relief-F scores for each model type. Vertical line represents 20 features.

Table 4-2 shows classification accuracy descriptive statistics, the average classification accuracy over all iterations to the classification accuracy achieved using the top 20 features ranked by Relief-F, for each model. ‘At 20’ denotes the classification accuracy achieved using the 20 highest-ranking features; classification accuracies had an average difference of 0.7% (SD=0.3) from the mean and were within an average difference of 0.9% (SD=0.2) from the maximum accuracy achieved over all feature subsets and model types (i.e., within 1% of the maximum accuracy). Thus, the feature subsets containing the top 20 features ranked by Relief-F were chosen for each model and for comparison against CFS and J48-FS feature subsets.

Table 4-2: Relief-F features subset accuracies between 20 and 187 features

		20 to 187 Features				
Model Type	**Mean**	**Std. Dev.**	**Max**	**At 20**	**Mean Diff at 20**	**Max Diff at 20**
Full	97.42	0.21	97.59	96.61	0.81	0.98
LG	94.49	0.26	94.84	94.22	0.27	0.62
LGDSUS	98.17	0.21	98.27	97.24	0.93	1.03
LGRSLS	97.48	0.23	97.71	96.69	0.79	1.02

Table 4-3 shows the classification performance metrics for each decision tree model built with feature subsets obtained using Relief-F, CFS, and J48-FS. Average sensitivity (0.97 (SD=0.02)), specificity (0.99 (SD=0.01)), precision (0.99 (SD=0.02)), F-Score (0.99 (SD=0.01)), and MCC (0.96 (SD=0.02)) were consistent across model types and feature selection methods, whereas classification accuracy, tree size, and number of leaves varied for each feature subset. Gait phase classification accuracies across model types and feature selection techniques were greater than 93%. Classification accuracies achieved using the CFS feature subsets were higher compared to Relief-F feature subsets for all model types with the exception of LG. In addition, decision tree sizes and number-of-leaves were lower using CFS compared to Relief-F within all model types.

Classification performance that was trained on the Full model using CFS feature subset and tested using 5-FCV had better results for all metrics with the exception of lower specificity and equal F-score compared to the Relief-F subset. Similarly, using CFS feature subset on the LG-DS-US model type, all performance metrics were better, with the exception of equal specificity. Overall, trees constructed with the LG model were smaller and the highest accuracies were achieved with LG-DS-US. Using feature selection, the smallest tree was built using CFS feature subset with the LG model type (3599) and the highest classification accuracy was achieved using CFS with the LG-DS-US model type (98.10%).

Table 4-3: Classification performance metrics

Model Type	Feature Selection	# Features	Tree Size	# Leaves	Acc. (%)	Sens.	Spec.	Prec.	F Score	MCC
Full (5-FCV)	Relief-F	20	16515	8258	96.61	0.97	0.99	0.97	0.97	0.95
	CFS	20	16103	8052	97.26	0.97	0.99	0.97	0.97	0.96
	J48-FS	176	13193	6597	97.58	0.98	0.99	0.98	0.98	0.97
LG	Relief-F	20	3829	1915	94.22	0.94	0.98	0.94	0.94	0.92
	CFS	18	3599	1800	93.76	0.94	0.98	0.94	0.94	0.91
	J48-FS	159	2837	1419	94.84	0.95	0.98	0.95	0.95	0.96
LGDSUS	Relief-F	20	12747	6374	97.25	0.97	0.99	0.97	0.97	0.96
	CFS	22	9323	4662	98.10	0.98	0.99	0.98	0.98	0.97
	J48-FS	176	8341	4171	98.27	0.98	0.99	0.98	0.98	0.98
LGRSLS	Relief-F	20	10039	5020	96.69	0.97	0.99	0.97	0.97	0.95
	CFS	20	9347	4674	97.25	0.97	0.99	0.97	0.97	0.96
	J48-FS	173	7605	3803	97.61	0.97	0.99	0.98	0.98	0.97
Relief-F		**Mean**	10783	5392	96.19	0.96	0.99	0.96	0.96	0.95
		(SD)	(5342)	(2671)	(1.35)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)
CFS		**Mean**	9593	4797	96.59	0.97	0.98	0.97	0.97	0.95
		(SD)	(5113)	(2557)	(1.93)	(0.02)	(0.01)	(0.02)	(0.02)	(0.03)
J48-FS		**Mean**	7994	3998	97.08	0.97	0.99	0.98	0.97	0.97
		(SD)	(4238)	(2119)	(1.52)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)

Despite the slightly greater performance metrics achieved using the entire training set (J48-FS), applying Relief-F and CFS to all model types resulted in much smaller feature subsets for training. Feature selection reduced training sets to only 20 (Relief-F) and 20-22 (CFS) features. Applying Wilcoxon rank sum test for classification performance achieved using these smaller training sets against results achieved using the full training set shows that there are no significant differences. For Relief-F, $p>0.5$ (Full), $p>0.5$ (LG), $p>0.4$ (LG-DS-US), and $p>0.6$ (LG-RS-LS), whereas for CFS, $p>0.7$ (Full), $p>0.4$ (LG), $p>0.9$ (LG-DS-US), and $p>0.6$ (LG-RS-LS).

Table 4-4: Relief-F and CFS feature subsets. For each model, common features between Relief-F and CFS are shaded

Full		LG		LG-DS-US		LG-RS-LS	
Relief-F 20 Features	CFS 21 Features	Relief-F 20 Features	CFS 18 Features	Relief-F 20 Features	CFS 22 Features	Relief-F 20 Features	CFS 20 Features
$\mathrm{sign}_{yAngVel}$	$\mathrm{mean}_{xKneeAng}$	$\mathrm{sign}_{yAngVel}$	$\mathrm{mean}_{xKneeAng}$	$\mathrm{sign}_{yAngVel}$	$\mathrm{mean}_{xKneeAng}$	$\mathrm{sign}_{yAngVel}$	$\mathrm{mean}_{xKneeAng}$
$\mathrm{sign}_{xAngVel}$	$\mathrm{mean}_{xAngVel}$	$\mathrm{sign}_{xAngVel}$	$\mathrm{mean}_{xAngVel}$	$\mathrm{sign}_{xAngVel}$	$\mathrm{mean}_{xAngVel}$	$\mathrm{sign}_{xAngVel}$	$\mathrm{mean}_{xAngVel}$
$\mathrm{sign}_{zAngVel}$	$\mathrm{mean}_{yAngVel}$	$\mathrm{sign}_{zAngVel}$	$\mathrm{mean}_{yAngVel}$	$\mathrm{sign}_{zAngVel}$	$\mathrm{mean}_{yAngVel}$	$\mathrm{sign}_{zAngVel}$	$\mathrm{mean}_{yAngVel}$
sign_{zAcc}	mean_{yAcc}	sign_{yAcc}	mean_{yAcc}	sign_{yAcc}	mean_{yAcc}	sign_{yAcc}	mean_{yAcc}
sign_{yAcc}	mean_{zAcc}	sign_{zAcc}	mean_{zAcc}	sign_{zAcc}	mean_{zAcc}	sign_{zAcc}	mean_{zAcc}
$\mathrm{diff}_{xKneeAng}$	$\mathrm{var}_{xKneeAng}$	$\mathrm{diff}_{xKneeAng}$	$\mathrm{diff}_{xKneeAng}$	$\mathrm{diff}_{xKneeAng}$	$\mathrm{var}_{xKneeAng}$	$\mathrm{diff}_{xKneeAng}$	$\mathrm{var}_{xKneeAng}$
sign_{xAcc}	$\mathrm{diff}_{xKneeAng}$	sign_{xAcc}	$\mathrm{diff}_{xAngVel}$	sign_{xAcc}	sum_{yAcc}	sign_{xAcc}	$\mathrm{diff}_{xKneeAng}$
$\mathrm{mean}_{xKneeAng}$	$\mathrm{diff}_{xAngVel}$	$\mathrm{mean}_{xKneeAng}$	$\mathrm{min}_{xKneeAng}$	$\mathrm{mean}_{xKneeAng}$	$\mathrm{diff}_{xKneeAng}$	$\mathrm{mean}_{xKneeAng}$	$\mathrm{diff}_{xAngVel}$
cagh_{Acc}	$\mathrm{min}_{xKneeAng}$	$\mathrm{min}_{xKneeAng}$	max_{yAcc}	$\mathrm{range}_{xKneeAng}$	$\mathrm{diff}_{xAngVel}$	$\mathrm{min}_{xKneeAng}$	$\mathrm{min}_{xKneeAng}$
$\mathrm{range}_{xKneeAng}$	$\mathrm{min}_{yAngVel}$	$\mathrm{rms}_{xKneeAng}$	$\mathrm{sign}_{xAngVel}$	$\mathrm{std}_{xKneeAng}$	$\mathrm{min}_{xKneeAng}$	$\mathrm{range}_{xKneeAng}$	$\mathrm{min}_{yAngVel}$
$\mathrm{std}_{xKneeAng}$	max_{yAcc}	$\mathrm{range}_{xKneeAng}$	$\mathrm{sign}_{yAngVel}$	$\mathrm{min}_{xKneeAng}$	$\mathrm{min}_{yAngVel}$	$\mathrm{rms}_{xKneeAng}$	max_{yAcc}
$\mathrm{min}_{xKneeAng}$	$\mathrm{sign}_{xAngVel}$	$\mathrm{std}_{xKneeAng}$	$\mathrm{sign}_{zAngVel}$	$\mathrm{fftmag}_{xKneeAng}$	$\mathrm{max}_{xAngVel}$	$\mathrm{std}_{xKneeAng}$	$\mathrm{sign}_{xAngVel}$
$\mathrm{fftmag}_{xKneeAng}$	$\mathrm{sign}_{yAngVel}$	$\mathrm{fftmag}_{xKneeAng}$	sign_{xAcc}	$\mathrm{iqr}_{xKneeAng}$	max_{yAcc}	$\mathrm{fftmag}_{xKneeAng}$	$\mathrm{sign}_{yAngVel}$
$\mathrm{iqr}_{xKneeAng}$	$\mathrm{sign}_{zAngVel}$	$\mathrm{iqr}_{xKneeAng}$	sign_{yAcc}	$\mathrm{rms}_{xKneeAng}$	$\mathrm{sign}_{xAngVel}$	$\mathrm{iqr}_{xKneeAng}$	$\mathrm{sign}_{zAngVel}$
$\mathrm{rms}_{xKneeAng}$	sign_{xAcc}	$\mathrm{sum}_{xKneeAng}$	pk_{rAcc}	cagh_{Acc}	$\mathrm{sign}_{yAngVel}$	$\mathrm{sum}_{xKneeAng}$	sign_{xAcc}
mean_{yAcc}	sign_{yAcc}	$\mathrm{trapz}_{xKneeAng}$	$\mathrm{pfreq}_{yAngVel}$	$\mathrm{entropy}_{xKneeAng}$	$\mathrm{sign}_{zAngVel}$	$\mathrm{entropy}_{xKneeAng}$	sign_{yAcc}
$\mathrm{entropy}_{xKneeAng}$	sign_{zAcc}	$\mathrm{entropy}_{xKneeAng}$	$\mathrm{entropy}_{xKneeAng}$	mean_{yAcc}	sign_{xAcc}	$\mathrm{trapz}_{xKneeAng}$	sign_{zAcc}
$\mathrm{sum}_{xKneeAng}$	sign_{rAcc}	mean_{yAcc}	cagh_{Acc}	$\mathrm{sum}_{xKneeAng}$	sign_{yAcc}	cagh_{Acc}	pk_{rAcc}
$\mathrm{trapz}_{xKneeAng}$	pk_{rAcc}	$\mathrm{max}_{xKneeAng}$		$\mathrm{trapz}_{xKneeAng}$	sign_{zAcc}	mean_{yAcc}	$\mathrm{pfreq}_{yAngVel}$
$\mathrm{max}_{xKneeAng}$	$\mathrm{pfreq}_{yAngVel}$	cagh_{Acc}		$\mathrm{max}_{xKneeAng}$	pk_{rAcc}	$\mathrm{max}_{xKneeAng}$	cagh_{Acc}
	cagh_{Acc}				$\mathrm{pfreq}_{yAngVel}$		
					cagh_{Acc}		

Table 4-4 presents the features chosen using Relief-F and CFS, and highlights the common features used in each feature subset to train each model. Interestingly, there are 11 common features between feature subsets for each model. The features shown for Relief-F are the 20 highest ranking features and are placed in order according to their Relief-F score. 10 of those features are common across all model types and feature selection subset: $\mathrm{mean}_{xKneeAng}$, mean_{yAcc}, $\mathrm{diff}_{xKneeAng}$, $\mathrm{min}_{xKneeAng}$, $\mathrm{sign}_{xAngVel}$, $\mathrm{sign}_{yAngVel}$, $\mathrm{sign}_{zAngVel}$, sign_{xAcc}, sign_{yAcc}, cagh_{Acc}.

4.5 Discussion

The goal of this study was to determine whether machine learning can be applied to localized walking signals for viable application to orthosis-control. Classification models using a J-48 DT machine learning classifier successfully identified gait phases across a variety of surfaces that are encountered when walking in the community. Since data features were extracted from thigh and knee kinematics and small data windows, results could be applied to lower limb stance control orthotics where a localized sensor system can improve assistive device design.

A J-48 DT machine learning classifier was chosen because of its performance in a preliminary analysis for localized-sensor GPR [1] and use in other studies associated with human activity recognition and gait [48], [49], [52], [56], [78]. Decision Trees use entropy formulation to organize and structure particular features using information gain and are optimized to the classification task [35]. Decision trees are robust and able to handle noisy data such as sliding window transitions in time-series analysis. They are also easy to implement once built, since they are simply a set of if-then statements and conditions. Most importantly, decision trees are fast, making them a suitable choice for applications in real-time control of dynamical systems [68], [75], [100]. An important distinction to make is that window sizes do not preclude real-time implementation but rather, classification frequency at each window step (i.e., making a decision at a 100Hz with a sliding window incrementing every 0.01 seconds).

Although decision trees have their strengths they also have limitations, such as being susceptible to over-fitting when there is not enough training data. Decision Tree disadvantages include high levels of complexity from training, resulting in an inability to convey data properly and generating unnecessarily large trees that potentially slow decision-making. Decision trees are also prone to over-fitting and cannot generalize to new data sufficiently [40]. This poses a problem when introducing new information, making the model non-generalizable. This is quite important when applying a model to be used as a control system for an assistive walking device such as an M-SCKAFO. To mitigate these problems, good training sets optimized through feature selection, pruning algorithms, and suitable testing protocols are essential to increase the decision tree's generalizability.

A technique known as pruning is used to simplify decision trees by removing portions from the tree that do not contribute to the classification task (i.e., removing nodes and leaves). The J-48 DT algorithm implements pruning when constructing the classification model. The outcome is often a simpler decision tree that still obtains high overall classification accuracy, as demonstrated in Table 4-3. Full, LG-DS-US, and LG-RS-LS models have a large amount of data instances, thereby facilitating larger and more complex decision trees. J48-FS subsets did not use all 187 features suggesting that branches in the tree have been pruned for generalizability and performance metric optimization. Whereas, for the LG model, the tree is only trained on level ground data that has far less data compared to the other training sets, enabling construction of a smaller and potentially faster decision tree. However, since the LG model performance was poor compared to other training models and so, this model is not recommended for the orthotic device application.

Feature selection is essential for training sets that have high data dimensionality and redundancy, which is the case for the full training set (187 features). Even after embedded feature selection, the final feature subsets used by J48-FS were large. Evaluation results from Relief-F subsets were based on the top 20 features, since classification accuracy using the top 20 features remained close to the mean accuracy and maximum accuracy using subsets with 20-187 features. Table III shows that training classification models on smaller feature subsets produced larger trees across all cases, increasing model complexity in terms of tree size. However, reducing the number of features to 20-22 maintained evaluation metrics compared to using all features. Relief-F and CFS removed many features that would not have positively contributed to the classification task. Computing a large number of features and outputting a decision at 100Hz is more computationally intensive. Therefore, smaller feature subsets decrease computational resources needed for feature extraction in each window.

All models performed well, each achieving suitable evaluation metric values. The high F-Scores and MCC values suggest that, regardless of class imbalances present within our data set, each model type trained on Relief-F and CFS feature subsets performed well when tested on the full set and with 5-FCV on the Full model. High sensitivity and specificity are essential for determining whether a true class is correctly classified as true or a false class is correctly classified as false. J-48 DT gait phase classification with feature selection training sets provided slightly better specificity than sensitivity in all models, meaning that each model is better at correctly classifying negative instances as negative, rather than positive instances as positive. For GPR and further application for real-time control, both high sensitivity and specificity are essential. Nonetheless, both sensitivity and specificity conditions are met in all models, each obtaining average values of 0.96 (Relief-F) and 0.97 (CFS), respectively. For example, specificity errors could result in the system switching to a free knee flexion setting during weight-bearing, putting the user at risk of falling. Similarly, sensitivity errors could result in the joint failing to switch to knee control mode when transitioning from swing to stance. In practice, other threshold-based methods could be employed in parallel to ensure that the joint enters a support mode after foot strike.

An ideal stance-control system using a decision tree would be small (reducing the computational time to make a decision), and generalizable to different surfaces and walking speeds. Most classification models for human walking and human activity recognition that used a localized sensor system and had similar complexity [47], [49], [51], [53], [88] were trained and tested strictly on level walking. These models may not be robust enough to effectively classify gait phases when experiencing real-world environments with different terrains and uneven ground, as well as achieving real-time control.

When comparing classifier models constructed with the various training sets, the LG model type achieved lower classification accuracies than all other model types (<95%), since the decision tree for the LG model is trained only on level ground. While the LG model type had the lowest accuracies, it also had the smallest size decision tree. Having a smaller tree is beneficial for computational time, and most importantly for making the model generalizable, reducing the chance for an over-fitted tree. However, if the larger trees in this study are still viable for real-time operation in an embedded control system, then smaller tree size is not a factor for final model selection.

Jung, et al. [77] proposed a neural network based gait classification technique that use inertial sensors and force sensitive resistors placed on a lower limb robotic exoskeleton to provide limb orientation and angular velocities. This study involved two neural network classifiers (multilayer perceptrons and nonlinear autoregressive with external inputs). Their training set contained 78 strides, whereas the test set contained 38 strides. This results of this study yielded high performance for the offline case, with a classification rate above 97% for both techniques. Despite good classification performance, this study was done using 12 sensors and used a small sample size of strides.

Wang, et al. [18] conducted a study on accelerometer-based classification for level ground and uneven ground using a multilayer perceptrons neural network classifier and obtained overall classification accuracy of 92.05%. Their classification model achieved high accuracies for various ground levels. On the other hand, they used feature extraction windows that were 2.56 seconds in length with 50% overlap; therefore, these would not be suitable for real-time control during a stride since a window this large would capture the entire stride (i.e., 1.28 seconds of data). In addition, features extracted for a 2.56 second window would not reflect gait phases sufficiently well.

Bonnet and Jallon [82] employed hidden Markov models (HMM) to distinguish between gait patterns using magnetometers and accelerometers. They were able to achieve classification accuracies above 90 percent for gait modes, rather than gait phases. On the other hand, Taborri, et al. [21] also used HMMs to classify gait phases and obtained high sensitivity and specificity values (0.98). Despite using different sensor combinations with three IMU's (thigh, foot, and shank), this study only investigated self-paced (0.81 ± 0.08 m/s) level ground walking.

Our study demonstrated the poorest results when training only on LG conditions, but testing on a variety of surface conditions. Therefore, including a range of walking speeds may also have improved classifier robustness, but this aspect was not evaluated in this study. Despite the higher classification accuracies with level ground classifications experiments, error rates of 5-7% still have implications for real-time M-SCKAFO control, and could result in a person falling. For instance, in a day consisting of approximately 3000 steps, a 5%

error rate in misclassification could result in 150 chances for a stumble. It is important to note that misclassification of a single instance will not always result in a fall, since falls would depend on the gait phase involved. Therefore, training a classifier on a data set containing more or all conditions is beneficial for improving classification accuracy, thereby decreasing the chance for a misclassification that could result in a knee collapse event.

Training a decision tree classifier on the full data set (i.e., contains all conditions and speeds) yielded results that were not as good as training on LG-DS-US; namely, classification accuracy, tree size, and number of leaves. Using the LG-DS-US model could decrease the chance for misclassification to 64 times a day. When compared to LG-LS-RS, LG-DS-US would be chosen for a control model due to a lower chance for misclassification and better performance.

Despite decreasing the chance for misclassification for the Full model compared to lower chance by LG-DS-US, it is important to note that LG, LG-DS-US, and LG-RS-LS models were not tested using 5-FCV, but rather on the full data set containing all instances and portions of duplicate data. Full model results may be more reliable and more generalizable since we did not train and test on the same data. Nonetheless, classification performance scores were high and similar for both LG-DS-US and 5-FCV Full model types, supporting the use of these models for real-time M-SCKAFO control from a localized sensor system.

Experimental GPR results were obtained with data from an able-bodied population. Walking with an orthosis is significantly different than able-bodied walking [11]. Therefore, models may not be generalizable to a non able-bodied population wearing an orthosis and may not generate sufficient GPR performance similar to results obtained in this study. Therefore, future work would benefit from implementing this GPR system to individuals with knee-extensor weakness wearing an orthosis.

4.6 Conclusion

A novel GPR model, implemented using local sensor features from the thigh and knee, appropriately identified loading response, push-off, swing, and terminal swing gait phases.

The J-48 DT machine learning model was trained and tested using Sagittal-plane knee angle, thigh angular velocity, and thigh acceleration signals obtained from different surface conditions and walking speeds, thus enabling a model that could perform better across a range of daily living walking environments. Classification showed that correlation-based feature selection reduced redundancy and data dimensionality for all model types, making each model increasingly efficient and viable for local sensor GPR for real-time orthosis-control. Despite all models performing well, correlation-based feature selection provided the smallest and most relevant feature subset with good gait phase classification performance. This research provides supporting evidence that wearable sensors localized to the thigh and knee, integrated with machine intelligence algorithms, can provide effective GPR models generalizable to multiple real-world terrains and different walking speeds. With additional modifications, this approach can lead to smarter stance-control knee-ankle-foot orthoses with broader ankle and foot component options, since sensors on the lower limb and foot would not be required.

5. Design, Development, and Evaluation of a Local Sensor-Based Gait Phase Recognition System using a Logistic Model Decision Tree for Orthosis-Control

This chapter addresses objective 3 by training a logistic model decision tree and evaluating the performance of the system. A "transition sequence verification and correction" algorithm improved GPR performance and generalizability across multiple conditions since natural walking is a temporal process.

The content of this chapter will be submitted for publication to the Journal of NeuroEngineering and Rehabilitation:

Farah JD, Lemaire ED, Baddour N. Design, Development, and Evaluation of a Local Sensor-Based Gait Phase Recognition System using a Logistic Model Decision Tree for Orthosis-Control. Journal of NeuroEngineering and Rehabilitation, Submitted May 2018.

5.1 Abstract

Background: Functionality and versatility of Microprocessor-controlled Stance-Control Knee-Ankle-Foot Orthoses (M-SCKAFO) are dictated by their control systems. Proper Gait Phase Recognition (GPR) is required to enable these devices to provide knee-control at the appropriate time, thereby reducing the incidence of knee-collapse and fall events. Ideally, the M-SCKAFO sensor system would be local to the thigh and knee, to facilitate orthosis design and allow more flexibility for ankle joint selection. We hypothesized that machine learning integrated with a local sensor system could effectively distinguish gait phases across walking conditions and that performance would improve through gait phase transition criteria (i.e., current states depend on previous states).

Methods: A Logistic Model decision Tree (LMT) classifier was trained and tested (five-fold cross-validation) on gait data obtained from sagittal-plane knee angle, thigh-segment angular velocity, and thigh-segment acceleration. Twenty features were extracted from 0.1 second sliding windows for 30 able-bodied participants that walked on different surfaces (level ground (LG), down-slope (DS), up-slope (US), right cross-slope (RS), left cross-slope (LS)) at a various walking speeds (self-paced (1.33 m/s, SD = 0.04 m/s), 0.8, 0.6, 0.4 m/s). The LMT-based GPR model was also tested using a separate validation set containing similar features from 12 able-bodied volunteers that walked on LG, DS, US, RS, LS surfaces at self-paced (1.41 m/s, SD = 0.34 m/s) walking speeds. A "Transition Sequence Verification and Correction" (TSVC) algorithm was applied to correct for continuous class prediction and to improve GPR performance. The model was evaluated using tree size, number of leaf nodes, classification accuracy, sensitivity, specificity, precision, F-score, and Matthew's correlation coefficient.

Results: The LMT had a tree size of 1643 with 822 leaf nodes with a logistic regression model at each leaf node. The local sensor LMT-based GPR model reliably identified loading response, push-off, swing, and terminal swing gait phases with overall classification accuracy of 98.38 for the initial training set (five-fold cross-validation) and 90.60% for the validation set. Applying the TSVC algorithm, classification accuracy increased to 98.72%

for the initial training set and 98.61% for the validation set. All other classification metrics were sufficient for real-time orthosis control.

Conclusions: A novel machine learning-based GPR model that uses sensor features local to the thigh and knee is viable for dynamic knee-ankle-foot orthosis-control. The highly accurate GPR model is generalizable when combined with TSVC. This approach reduces sensor system complexity, as compared with other M-SCKAFO approaches, thereby enabling customizable advantages for end-users through modular unit orthosis designs.

5.2 Introduction

Stance-Control Knee-Ankle-Foot Orthoses (SCKAFO) are walk-assist devices that prevent the knee from collapsing during weight-bearing and provide unhindered knee motion during swing. Although irregular gait patterns are present due to a fully extended and locked knee during stance, SCKAFO offer many advantages over conventional fixed-knee KAFOs for people with knee-extensor pathologies [9], [10], [12], [18], [19]. By simply allowing free knee rotation during the swing phase, SCKAFO offer improved mobility and a more natural gait. Fewer compensatory gait mechanisms can reduce associated joint loads, reduce energy consumption [6], [23], increase foot clearance, increase walking speeds, and improve overall user satisfaction [6], [9], [10], [12], [19].

The primary goal for stance-control systems are to engage during stance to provide support and resist knee-flexion. Stance-control devices remove knee extensor requirements to counteract flexion moments imposed by external loads that would otherwise occur on the knee during weight-bearing. Different SCKAFO mechanical knee joint designs have their own advantages and disadvantages [9], [19]. Some designs used weighted/spring-loaded pawls or belt clamping [9], [10], [12] to lock the knee at initial contact and disengage at foot-off, based on leg position. Most mechanically-controlled SCKAFO require full knee extension to engage knee-lock and resist knee flexion [9], [19], [25]. This can present difficulties for individuals who cannot extend their leg at each step, making continuous and reliable stance-control more difficult. Persons with knee-extensor weakness that have sufficient hip-flexion control could also use angular velocity based stance-control [7], [18],

[25], [27], where mechanical components at the knee engage knee-flexion resistance at any knee angle once an angular velocity threshold is passed, such as during a knee collapse or fall event (i.e., body weight sensing not required).

The main advantages of mechanically-controlled SCKAFO include free knee motion during swing, simpler stance-control systems that do not require external power sources, low profile, lightweight, and ability to fit under trousers. Unfortunately, the greatest disadvantages of mechanically-controlled SCKAFO are stance/swing phase recognition, leading to inconsistent and unreliable locking and unlocking, and a lack of functional versatility across different walking conditions. Mechanically-controlled SCKAFO can have difficulty negotiating between different walking speeds, gait modes, terrains, and daily living environments (i.e., stairs, ramps, uneven ground) [1], [2], [9], [18], [19], [24], [25].

Microprocessor-controlled SCKAFO (M-SCKAFO) guide knee control by regulating between stance and swing using multiple electronic sensors attached to various orthotic-limb segments [1], [2], [6], [7], [9], [19]–[21], [27]–[32], [34]. Electronic sensors coupled with computational algorithms dictate when to engage/disengage knee-flexion resistance. This provides enhanced knee-control functionality, reliability, and versatility in terms of surface and walking speed. M-SCKAFO advantages also include the capacity for being able to toggle between different gait modes (i.e., ramps, curbs, stairs) [2], [9], [19]. These devices rely on many sensors and complex algorithms that can have high computational costs, which require external power sources and need regular charging. M-SCKAFO drawbacks include sensor-location placed at many orthoses segments, M-SCKAFO drawbacks include sensor-location placed at many orthoses segments, bulky (i.e., cannot fit under trousers), lack of aesthetics, expensive, and fewer orthotic component choices.

The Otto Bock C-Brace [38] is currently the most versatile and commercial available M-SCKAFO. This orthosis uses a hydraulic linear damper for knee control. C-Brace is the only M-SCKAFO on the market with in-stance knee flexion dampening (i.e., gradually allows the knee to flex during stance) and can offer effective movement across daily use. With partial knee-flexion resistances, some individuals with lower limb muscle impairment could re-establish sufficient strength and mobility to eliminate the need for ambulatory assistance. The C-Brace uses two sets of sensors at the knee to determine knee angle and a dorsal shank spring (strain-gauge ankle-moment sensor) to determine when the person is weight-bearing [19].

Electronic sensors, such as inertial measurement units (IMUs), enable signal-processing algorithms to determine limb orientation and/or position in the gait cycle and prompt the knee joint mechanism to switch to a locked, free, or partial knee flexion-resistance setting. Ideally, an effective stance-control system would provide support during stance and unhampered knee motion during swing, for natural gait. Knee stability is of paramount importance for safe gait, but does not always require knee extensor contributions [4]. During stance, knee motion is predominantly in extension. However, the onset of knee flexion occurs between terminal stance and pre-swing sub-phases. During this portion of the gait cycle, the body has forward impetus with controlled ankle and hip removing the need for knee extensor control [4]. Safe knee-release must occur without active quadriceps muscle activation, prior to knee-flexion during swing, and after the contralateral limb is in contact with the floor. Moreover, stance-control systems need to identify loading response to securely lock and keep the knee from collapsing. Criteria such as these provide knee-release transition points that ensure safe-gait by not imposing unwanted loads on weak knee-extensors, maintain support, and prevent knee-collapse. Consequently, to enable appropriate function, accurate GPR from wearable sensor data becomes essential for real-time orthosis-control.

Emerging gait analysis techniques use embedded sensors in wearables and offer practical modalities for gait monitoring, human activity recognition, and prosthesis and orthosis-control. Electronic sensors such as strain gauges [38], pressure sensors [57], electromyography (EMG) [21], force sensitive resistors [56], [59], [77], goniometers [16], [94], and IMUs [1], [2], [40], [46], [56]–[61], [77], [78] can be combined give highly accurate and real-time gait monitoring. This coincides reasonably well for GPR and stance-control associated with M-SCKAFO.

Liu, et al. [44] placed a two-axis accelerometer and three gyroscopes on the foot, shank and thigh to detect gait phases (initial contact, loading response, mid-stance, terminal stance, pre-swing, initial swing, mid-swing, terminal swing) and provide limb segment orientation. Pappas et al. [56] reported a reliable gait phase identification system across able-bodied and impaired individuals that used a rule-based algorithm. Gait states were dictated by prior characterization of inertial measurement signals. Their system depended on foot angular velocity and three force sensitive resistors to determine weight-bearing. To provide improved GPR performance, they developed knowledge-based gait phase transitions to determine possible changes of state during the gait cycle. Gorsic et al. [57] developed a gait phase identification system to provide feedback for a lower-limb robotic prosthesis. The system used IMUs attached to body segments and shoe insole sensors to determine four gait phases (left stance, left-right double stance, right stance, right-left double stance). Another study [53] attempted to localize a GPR sensor system further using a single IMU placed in a trouser pocket, and showed that major gait phases could be visually identified. This was important for lower limb orthosis-control since many M-SCKAFO can benefit from real-time GPR, along with modular electronic components located only at the thigh [1], [2].

Machine learning and artificial intelligence approaches related to HAR are becoming increasingly popular [48], [49], [52], [58], [88]. Machine learning classifiers can provide robust, fast, and accurate classification from simple features extracted from biomechanical data, making them highly attractive for use in GPR [1], [54], [57], [77]. Machine learning algorithms coupled with local sensor systems have demonstrated sufficiently good performance to make this a potentially viable approach for identifying current states and recognizing human intent for control system feedback.

The objective of this research is to identify walking gait phases with only local sensors at the thigh and knee. Accurate GPR would provide essential information for M-SCKAFO control, if the approach works in real-time, across different surface-levels, and across walking speeds. We demonstrate that integrating sensor signal features from the thigh and knee with a logistic decision tree machine learning model can provide highly accurate GPR performance across the M-SCKAFO criteria. Secondly, implementing "Transition Sequence Verification and Correction" (TSVC) algorithm improves classification results. Through model evaluation, our model classifies gait phases regardless of surface-level, walking speed, and individual variability. Appropriate gait phase recognition without sensors below the knee will enable new M-SCKAFO approaches that allow orthotists to use the most appropriate ankle and foot designs for the user, without limitations due to ankle-foot sensor system requirements in current devices.

5.3 Methods

5.3.1 **Data Set**

This research was approved by the Ottawa Health Science Network Research Ethics Board (20140825-01H). A de-identified data set from 30 able-bodied participants was used in this study. Walking data were collected in a Computer-Assisted Rehabilitation Environment (CAREN-Extended) (Motek Medical, Amsterdam, NL) virtual environment system. CAREN-Extended consists of a six degree-of-freedom force-plate platform with 1m x 2m dual tread instrumented treadmill (Bertec Corp., Columbus, OH), 180º screen for virtual world projection, and a 12-camera 3D motion capture system (Vicon Inc., Oxford, UK). A full body marker set defined all joint and body-segment positions [95], [96]. Marker data were recorded at 100Hz and ground reaction forces were recorded at 1000Hz.

Each volunteer walked in a custom-built virtual park application on level ground (LG), 7^{o}-declination down-slope (DS), 7^{o}-inclination up-slope (US), 5^{o}-inclination right cross-slope (RS), and 5^{o}-inclination left cross-slope (LS), at self-paced (SP) speed (1.33 m/s, SD = 0.04 m/s), 0.8 m/s, 0.6m/s, and 0.4m/s. Joint and body-segment trajectories were imported into Visual 3D (C-Motion Inc., Germantown, MD) for biomechanical analyses. A fourth order low pass Butterworth filter with 10Hz cut-off frequency was applied to the data before deriving joint moments and powers using inverse dynamics.

Ten strides for knee angle (KA), thigh-segment angular velocity (AngVel), and thigh-segment acceleration (Acc) were extracted from each walking condition (surface and speed) and imported into Matlab. All data were linearly interpolated to 200Hz to mimic an IMU's internal sampling rate for M-SCKAFO control systems.

5.3.2 Validation Set

A de-identified data set from 12 able-bodied participants provided gait data analogous to the training data set in terms of stride trajectories and gait events (i.e., KA, AngVel, Acc, initial contact, foot-off, end contact). Six strides from each surface condition were extracted for SP (1.41 m/s, SD = 0.34 m/s) walking speeds from each volunteer. This data set can be considered as unseen data.

5.3.3 Gait Phase Recognition

Sagittal plane Knee Angle (KA) (*x*-axis), thigh-segment Angular Velocity (AngVel) (*x*, *y, z*), and thigh-segment Acceleration (Acc) (*x*, *y, z*) signals were partitioned into four gait phases: Loading Response (LR), Push-Off (PO), Swing, and Terminal Swing (TSw) according to initial contact, mid-stance, foot-off, and maximum knee flexion angle gait events [1], [2]. AngVel and Acc resultants were also calculated and used as supplementary signals for machine learning implementation.

Correlation-based feature selection was applied in a preliminary analysis [2], providing the feature set listed in Table 5-1. All features were extracted from a 0.1-second sliding window and labelled as LR, PO, Swing, or TSw according to the last data point [1], [2] in each window.

Table 5-1: List of extracted features used for gait phase recognition

Feature	Feature Description
1	Sagittal plane knee angle mean (x-axis)
2	Sagittal plane angular velocity mean (x-axis)
3	Frontal plane angular velocity mean (y-axis)
4	Frontal plane acceleration mean (y-axis)
5	Transverse plane acceleration mean (z-axis)
6	Sagittal plane knee angle variance (x-axis)
7	Sagittal plane knee angle maximum difference (x-axis)
8	Sagittal plane angular velocity maximum difference (x-axis)
9	Sagittal plane knee angle minimum (x-axis)
10	Frontal plane angular velocity minimum (y-axis)
11	Frontal plane acceleration maximum (y-axis)
12	Sagittal plane angular velocity sign (-1,0,1) sum (x-axis)
13	Frontal plane angular velocity sign (-1,0,1) sum (y-axis)
14	Transverse plane angular velocity sign (-1,0,1) sum (z-axis)
15	Sagittal plane acceleration sign (-1,0,1) sum (x-axis)
16	Frontal plane acceleration sign (-1,0,1) sum (y-axis)
17	Transverse plane acceleration sign (-1,0,1) sum (z-axis)
18	Resultant acceleration sum of peaks
19	Frontal plane angular velocity principle frequency (DFT y-axis)
20	Correlation along gravity (z-axis) and heading (y-axis)

5.3.4 **Logistic Model Decision Tree**

The feature vector was imported into the Waikato Environment for Knowledge Analysis (WEKA) for machine learning implementation. A logistic model decision tree (LMT) [76] classifier was trained and tested with 5-Fold Cross Validation (5-FCV) using stratified data splits. The LMT was constructed as a J-48 Decision Tree with logistic regression models at terminal leaf nodes. Node splitting was implemented using the C4.5 decision tree splitting criterion [75], [76], [99] that used information gain on the class variable. Logistic regression functions at leaf nodes were determined using LogitBoost heuristic [76]. For this study, the LMT parameters used to construct the tree were set to default in WEKA (i.e., convertNominal = False, debug = False, errorOnProbabilities = False, fastRegression = True, minNumInstances = 15, numBoostingIterations = -1, splitOnResiduals = False, useAIC = False, weightTrimBeta = 0).

The LMT was manually implemented into a Matlab script as a GPR function. The function took single instances as input, and returned the gait phase corresponding to the LMT model, with the highest probability as the gait phase prediction.

5.3.5 Transition Sequence Verification and Correction

A transition vector verification and correction algorithm (TSVC) was applied after the GPR function class output to resolve discontinuous class complications (e.g., classifying PO directly after swing in a dynamic state), thereby increasing classification accuracy. The algorithm included altering the class buffer if a discontinuous gait phase sequence occurred. The continuous sequence was defined as LR, PO, Swing, TSw, and repeated for the following stride. For example, if the models were to identify LR as the current instance and had identified LR for three prior instances (0.03 seconds), there would be a forced transition to the PO phase if and only if the following instance does not align with the gait phase sequence conditions (i.e., Swing, TSw). Similarly, the same procedure would occur for each gait phase. In addition, if there were classification between similar consecutive gait phases such as {LR, LR, Swing, LR}, for four instances, the algorithm would correct the sequence by replacing Swing with LR, and similarly for all other gait phases. This process also indirectly identifies gait events through a gait phase transition (i.e., transition between TSw and LR indicates IC).

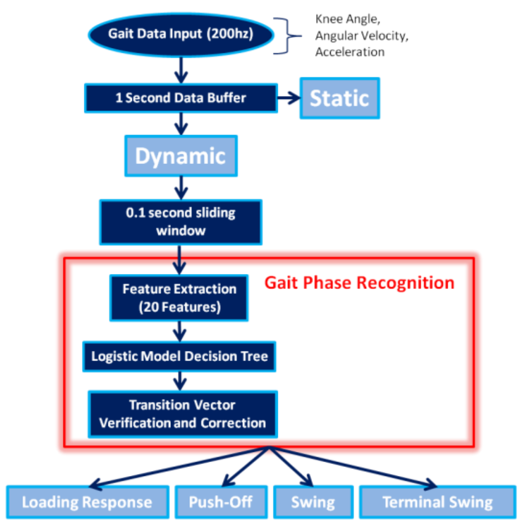

Figure 5.1: Gait Phase Recognition Flow Diagram.

As shown in Figure 5.1, sensor input from KA, AngVel, and Acc were stored in a one-second data buffer (200 data points, 200Hz sampling rate) for feature extraction. A static state was defined as standing, sitting, etc., whereas a dynamic state would be defined as walking. If a static state was declared, the knee joint would be set to a support condition (i.e., high knee-flexion resistance). If a dynamic state existed, the system would partition the signal into 0.1s sliding windows (20 data points) and LMT-based GPR would be performed.

5.3.6 **Classifier Evaluation**

Classification performance metrics included tree size, number of leaf nodes, overall classification accuracy, sensitivity (Sens), specificity (Spec), precision (Prec), F-score (FS), Matthews Correlation Coefficient (MCC). Classification metrics are computed from true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). These metrics were computed using WEKA by supplying the feature matrix as the test set. Classification metrics for LMT with TSVC tested with the training set and LMT with TSVC tested with the validation set were computed from a confusion matrix implemented in Matlab2016b. Weighted averages for classification metrics were given by (5.1),

$$W_m = \frac{\sum_i m_i * I_i}{\sum_i I_i} \tag{5.1}$$

where *m* denotes any classification metric, *i* denotes a specific gait phase, *I* denotes the total number of instances in the total feature matrix relating to each gait phase.

5.4 Results

5.5 Gait Phase Recognition

The LMT size was 1643 with 822 leaf nodes, with a logistic regression function at each leaf node. Each variable in the linear combination represented feature values extracted from data windows. Equation (5.2) provides the probability of being in the LR phase.

$$P\left(LR\right) = \mathrm{e}^{\mathrm{LR}} / (1 - \mathrm{e}^{\mathrm{LR}}) \tag{5.2}$$

Predicted gait phases during strides are shown for sagittal plane KA (Figure 5.2), resultant AngVel (Figure 5.3), and resultant Acc (Figure 5.4) for each walking condition. Thigh and knee gait trajectory plots show that, for decreased walking speed, the transition from one gait phase to the next was very similar. The model distinguished between phases regardless of surface condition and walking speed.

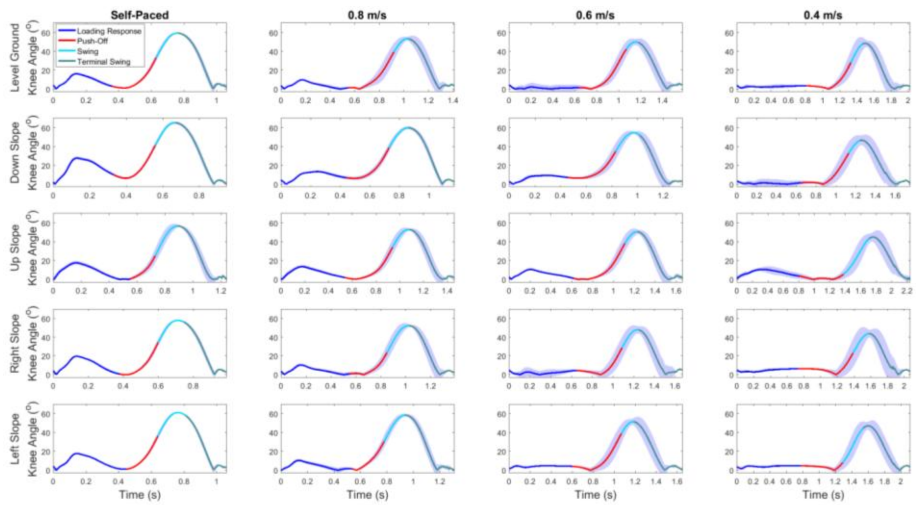

Figure 5.2: Average knee angle (°) from a single participant for each surface condition (rows) and walking speed (columns), showing classified gait phases: Loading Response (blue), Push-Off (red), Swing (cyan), Terminal Swing (teal).

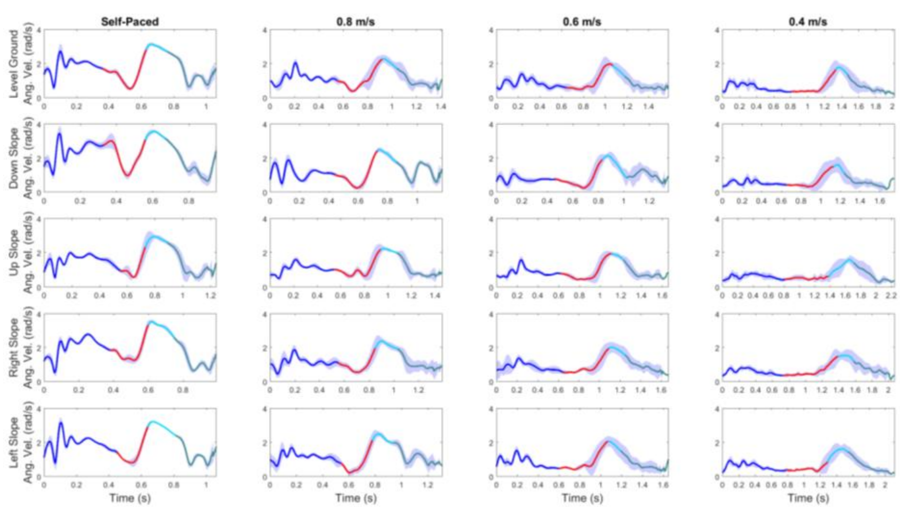

Figure 5.3: Average thigh-segment angular velocity (rad/s) from a single participant for each surface condition (rows) and walking speed (columns), showing classified gait phases: Loading Response (blue), Push-Off (red), Swing (cyan), Terminal Swing (teal).

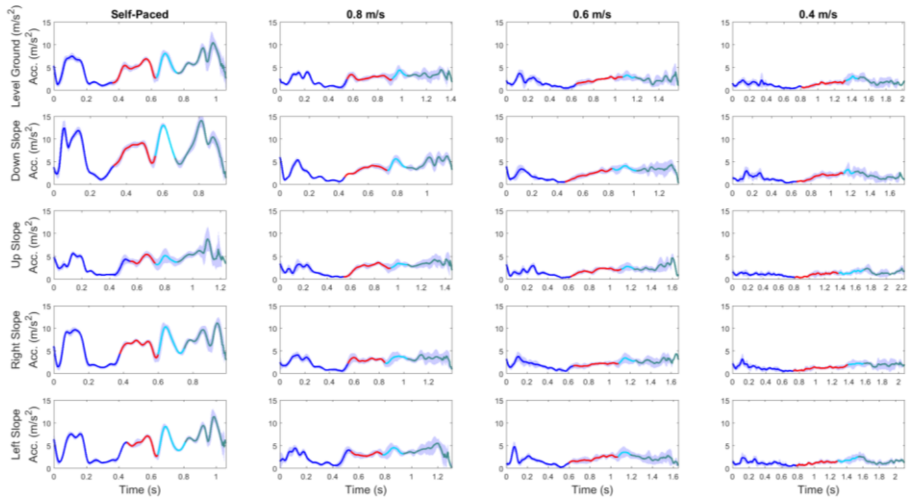

Figure 5.4: Average thigh-segment acceleration (m/s^2) from a single participant for each surface condition (rows) and walking speed (columns), showing classified gait phases; Loading Response (blue), Push-Off (red), Swing (cyan), Terminal Swing (teal).

Table 5-2 presents the classification evaluation metrics for the LMT machine learning classifier and training set (30 able-bodied participants) and validation set (12 able-bodied participants). Weighted averages for classification performances are shown for cases with and without TSVC.

Gait phase classification performed well for our training set, with 98.38% overall accuracy and all classification metrics greater than 0.98. Applying TSVC to the LMT results increased accuracy by 0.38%. Accuracy, sensitivity, and FS decreased by 1%, whereas MCC decreased by 2% and specificity remained the same.

Validation set classification performance metrics were less than the training set 5-FCV results, with lower accuracy (-7.78%), sensitivity (-7%), specificity (-2%), precision (-7%), FS (-7%), MCC (-11%). However, after applying TSVC, all validation set classification metrics improved. Accuracy increased by 8.01%, sensitivity, precision and FS all increased by 6%, specificity increased by 2%, and MCC increased by 9%, with all results greater than 0.96.

Table 5-2: Logistic model tree classification performance for GPR

GPR Model	Data Set	Size	Leaf Nodes	Accuracy	Sens.	Spec.	Prec.	FS	MCC
LMT	Training	1643	822	98.38	0.98	0.99	0.98	0.98	0.98
	Validation	1643	822	90.60	0.91	0.97	0.91	0.91	0.87
LMT+TSVC	Training	1643	822	98.76	0.97	0.99	0.97	0.97	0.96
	Validation	1643	822	98.61	0.97	0.99	0.97	0.97	0.96

Table 5-3 shows classification performance metrics for the LMT tested (5-FCV) with the training set detailed for specific gait phases. All classification metrics were greater than or equal to 0.97 across all gait phases. TSw performed the best with all classification metrics at 0.99. Swing had the lowest precision and FS.

Table 5-3: Training Set LMT-GPR Performance Detailed by Class (5-FCV)

Gait Phase	Sens.	Spec.	Prec.	FS	MCC
Loading Response	0.98	0.99	0.98	0.98	0.97
Push-Off	0.98	0.99	0.98	0.98	0.97
Swing	0.98	0.99	0.97	0.97	0.97
Terminal Swing	0.99	0.99	0.99	0.99	0.99

Table 5-4 shows classification performance metrics for the LMT tested with the validation set detailed for specific gait phases. Excluding PO, classification metrics were all greater than 0.85. PO had the lowest sensitivity, FS, MCC. LR had the second lowest sensitivity and the lowest specificity. In terms of classification performance metrics, Swing and TSw performed similarly and outperformed LR and PO for unseen data.

Table 5-4: Validation Set LMT-GPR Performance Detailed by Class

Gait Phase	Sens.	Spec.	Prec.	FS	MCC
Loading Response	0.91	0.95	0.87	0.89	0.85
Push-Off	0.82	0.97	0.93	0.87	0.81
Swing	0.96	0.99	0.92	0.94	0.93
Terminal Swing	0.99	0.96	0.91	0.95	0.93

Table 5-5 shows classification performance metrics for the LMT with TSCV tested with the training set, detailed for specific gait phases. All classification metrics were greater than 0.91. Compared to the results in Table 5-3, sensitivity decreased for PO and Swing, and remained the same for LR and TSw. Specificity performed well across all gait phases. Swing had the lowest sensitivity, FS, MCC. TSw performed the best with all classification metrics equal to 0.99, with the exception of MCC. However, MCC was still greater for TSw over all other classification performances.

Table 5-5: Training Set LMT+TSVC GPR Performance Detailed by Class

Gait Phase	Sens.	Spec.	Prec.	FS	MCC
Loading Response	0.98	0.99	0.97	0.98	0.97
Push-Off	0.96	0.99	0.94	0.95	0.94
Swing	0.91	0.99	0.97	0.94	0.93
Terminal Swing	0.99	0.99	0.99	0.99	0.98

Table 5-6 shows classification performance metrics for the LMT with TSCV tested with the validation set detailed for specific gait phases. All classification performance metrics were greater than 0.87. Compared to the results in Table 5-4, all classification metrics increased for LR and TSw. All classification metrics increased for PO with the exception of precision (decreased by 8%). Classification metrics for Swing decreased for sensitivity (9%), FS (3%), MCC (3%). For this case, classification metrics for LR outperformed all other gait phases.

Table 5-6: Validation Set LMT+TSVC GPR Performance Detailed by Class

Gait Phase	Sens.	Spec.	Prec.	FS	MCC
Loading Response	0.99	0.99	0.99	0.99	0.99
Push-Off	0.98	0.98	0.90	0.94	0.93
Swing	0.87	0.99	0.95	0.91	0.90
Terminal Swing	0.97	0.99	0.99	0.98	0.98

5.6 Discussion

This research demonstrated the viability of a local sensor machine learning-based GPR system to guide decision-making for safe-gait and orthosis-control. The LMT machine learning classifier designed in this study successfully performed GPR for small data windows, across different surfaces and walking speeds encountered throughout daily living activities.

5.6.1 **Gait Phase Recognition Classifier Design & Development**

The machine-learning model reliably recognized LR, PO, Swing, and TSw, defined by gait events [1] across LG, DS, US, RS, and LS at different walking speeds. In the literature, most GPR models were trained and tested on level ground walking at self-paced speeds [1], [54], [59], [60], [77], with few studies testing on a variety of surfaces [2], [56]. Models trained with multiple walking conditions that occur throughout daily living activities could enhance generalizability [16] and overall system functionality across daily living activities.

Since our GPR model was trained from a data set that contained simple data features extracted from small data windows (overlapping 0.1 second sliding window) from local thigh and knee signals across a variety of walking conditions, results can be directly translated to lower-limb orthosis control. These features are not computationally intensive, facilitating fast decision-making for GPR throughout the gait cycle.

In a preliminary analysis [1], [2], a J-48 decision tree was chosen as the machine learning classifier due to its ability to their previous success in activity recognition and gait [48], [49], [52], [56], [78], exceptional gait phase classification performance [1], [2], and their ability to work in real-time.

Despite J-48 decision trees easy interpretability and robustness, they are low bias, high-variance classifiers. The primary objective of supervised learning is to achieve high classification performance with low bias and low variance. With large amounts of training data, DT's are prone to over-fitting, and become quite large. With a lot of training data, DT complexity (proportional to tree size [76]) increases and cannot represent true classification structure adequately. When there are many nodes, an input may need to traverse through a set of conditions until reaching a leaf node that satisfies all conditions and provides an output. A single deviation in the nodes' conditions may result in a different output from a different branch in the tree, hence the high-variance.

To help with bias-variance trade-off and overall classification performance; representative training data sets and validation protocols such as cross validation or new test sets are essential for increasing LMT generalizability. Pruning increases bias error and decreases variance error by making the model less complex and more generalizable. However, oversimplifying a DT by pruning too much results in higher bias and therefore can result in an under-fit system that is unable to convey classification results.

The LMT machine learning classifier in this study was constructed as a J-48 DT that implements logistic regression models at each terminal node and is able to handle multi-class target variables. This provides a direct advantage for our classification task in terms of providing gait phase probability estimates, rather than just a classification output, and when correctly pruned produces a smaller tree than ordinary classification trees [76]. The J-48 DT previously researched for local sensor GPR was very large with a tree size of 16103 and 8052 leaf nodes [2] even after pruning, and was likely over-fit. This would not be a surprise due to the large amount of data in our training set. Generally, a smaller decision tree is a simpler model, and for our application would also provide computational efficiency in terms of battery consumption and classification time with an onboard microprocessor on an M-SCKAFO. At first glance, the LMT machine learning classifier constructed in this study with the same training data set containing all walking conditions produced a tree that was approximately 90% smaller in tree size and had 90% less terminal leaf nodes than J-48 DT [2]. Roughly speaking, in the case of GPR for application in real-time orthosis control, the high accuracy and generalizability becomes more important than tree size, as long as a classification decision can be made within a period sufficient for assistive-device controls.

Upon further observation, classification performance also improved for all other classification metrics (Table 5-2). These results are in alignment with those of Landwehr, et al. [76] for their investigation on LMTs, and extends to our classification task. In their study, LMTs outperformed C4.5 DT (similar to J-48 DT) on 14 different data sets and LMT size was statistically significantly less than C4.5 DT size for 16 data sets. For our study, this is likely due to probability estimates from logistic regression functions at each node. Branches in the tree would guide input information towards a node based on our highly representative training set, and introduces an additional, probabilistic measure to determine class output. This model’s characteristics enable easier, faster, and more accurate GPR implementation on a microprocessor. In addition, a simpler and more accurate model is more practical for deeper supplementary rule-based algorithms (i.e., sequence transitions) thereby improving assistive-device designs, and machine learning-based stance-control feasibility.

Results obtained from LMT tested on the training set using 5-FCV performed sufficiently well across all gait phases. Sensitivity was less than specificity in all cases. High sensitivity and specificity are essential for determining whether we are correctly recognizing a gait phase in the current state and correctly recognizing a gait phase that is not in the current state [2]. Specificity errors could result in the system switching to a knee-release setting during weight-bearing (e.g., mistaking LR for Swing), putting the user at risk of falling. Similarly, sensitivity errors could result in the joint failing to switch to knee-release when transitioning from PO to Swing. In practice, maintaining knee support during weight bearing is a priority for fall-prevention. Therefore, greater values of specificity are advantageous to stance-control applications. Specificity criteria for classification across all phases were met for the training set, and TSVC implementation.

The LMT's overall classification accuracy improved with TSVC algorithm implementation, but other evaluation metrics diminished slightly. Interestingly, implementing the model on the full training set with the TSVC algorithm, improved overall classification accuracy. The decrease in sensitivity for PO and Swing meant that TPs were missed, likely due to forcing a transition if a particular class was mistaken for any another class. Swing phase was expected to occur after PO and had fewer data instances. Hence, the greater decrease in sensitivity for Swing, and is representative from relevant classification in the precision score. Therefore, during the transition from PO to Swing, the onset of Swing was classified as PO, likely for slower speed walking conditions (Figure 5.2-5.4). Literature regarding GPR of very slow walking is sparse. A separate GPR model trained only on very slow walking biomechanics may improve accuracy but may likely decrease generalizability. Multiple GPR models with certain speeds (i.e., <0.4 m/s) that engage model switch may be better to determine appropriate knee engagement/disengagement.

Another factor may be due to the class imbalance between phases and is evident when observing FS and MCC metrics. Similar FS and MCC values suggest that, regardless of class imbalances present within our data set, classification between PO and Swing performed well when tested on the training set with TSVC. Stance phase is captured by about 60% of the normal gait cycle and is most naturally longer than the swing phase, which is captured by about 40% of the normal gait cycle [3], [4].

Results from our GPR model obtain good classification performance that is sufficient for application in an M-SCKAFO. Strictly based on overall classification accuracy, evidence suggests that both of our hypotheses are valid. However, if we recall that DTs are low-bias classifiers, we expect them to capture regularities in the training data well. Hence, the reason for evaluating the GPR model on completely new data (validation set).

5.6.2 Gait Phase Recognition Model Evaluation

As expected, testing the LMT classifier on the validation set had worse classification performance than testing on the training set. Nonetheless, overall classification accuracy was still greater than 90% for all evaluation metrics. Therefore, the model was sufficiently generalized to adequately recognize gait phases from untrained local sensor features at the thigh and knee for 12 new participants. However, despite achieving generalizable results, a 10% error rate is not appropriate for assistive device-control. This evidence suggests that our initial hypothesis was not valid. The bias-variance trade-off is an important aspect to consider in supervised learning, as described earlier specifically for orthosis-control applications. An ideal system would adequately represent the training set (low bias), and be sufficiently generalizable to similar input acquired from different people. The TSVC algorithm improved classification accuracy with the validation set since gait phases are sequential in normal gait, walking can be regarded as a temporal process. This suggests that transition correction algorithms may improve GPR models, and validates the second hypothesis for new gait data.

It is important to consider when in each gait phase the TPs are misclassified. Knowing where and when misclassifications occur is central for knee engagement/disengagement in M-SCKAFO implementation. Missed TPs occurred during gait phase transitions where gait signals exit one phase and enter the other. For small data windows, features will be similar in local signal spaces without sudden gradient-changes. If

TPs are misclassified at the beginning of swing, these instances are classified as PO, thereby engaging knee-release. However, misclassifying instances near the middle or end of Swing may engage knee-support settings too early and perturb free knee-motion. For the LMT tested with the validation set and TSVC, results showed that 1.8% of instances were classified as PO but were actually Swing, and 0.5% of instances were classified as Swing but were actually TSw, which suggests that the model is missing TPs at the beginning of Swing, rather than at the end. Misclassifying the beginning of swing for PO supports the case for initiating knee-release transition points. For M-SCKAFO, knee-release would have to occur at some point during the PO phase (onset of knee flexion during swing). In summary, classification performance for the GPR model tested with the validation set coincides well with results obtained from the training set, and is sufficient for orthosis-control.

5.6.3 **Future Design & Application to Orthosis-Control**

An important aspect to consider is the minimum performance of a model to enable application to real-time orthosis-control. One must not ignore that a misclassification will not necessarily result in knee-collapse, since stance-control depends on when to trigger knee engagement/disengagement at the correct instant during a particular phase [2]. Each M-SCKAFO has its own approach for determining when knee-release occurs [19]. Depending on the current state (gait phase), switching knee-release too late will inhibit free knee motion, and switching too early is a safety concern and could result in patients falling. The latter is more important for enabling safe-gait during daily living environments. For a local sensor machine learning-based GPR model, this becomes a more challenging task especially for slow walking speeds.

Stance phase increases for slower speeds to increase double support time and maintain stability [101]. Certain implications and difficulties arise when implementing stance-control at extremely slow walking speeds such as 0.4 m/s. This is mainly due to the lack of signal variation at these locations during the stride, as the mean standard deviations (MSD) are close to zero. Ideally, with a sufficient static versus dynamic state decision algorithm [102], stance-control could be applied with a local sensor system at very slow walking speeds.

However, another reason may likely be due to the sampling rate and window size that were the same across all speeds. By reducing the sampling rate or making the window size larger, features extracted from slow walking would resemble those of faster walking speeds. Depending on the final control systems application, window sizes could be decided dynamically based on a users' walking speed.

Intelligent assistive walking devices are based on the control system's performance. An orthosis with increased functionality driven by a control system that offers versatility and performance to the user is the better product. Hence, developing a robust and generalizable real-time GPR model that uses input information from localized sensors at the thigh and knee has modular unit feasibility. As an outcome GPR models using localized walking signals as input are able to produce highly functional stance-control systems for safe and healthy gait, a wider selection of M-SCKAFO ankle and foot components, user-personalization, and most importantly patient satisfaction.

5.7 Conclusion

A logistic decision tree-based GPR model successfully identified loading response, push-off, swing, and terminal swing gait phases for five different surfaces and a range of walking speeds, with input data localized to the thigh and knee. The LMT classifier was robust for walking conditions experienced throughout daily living, and was generalizable to unseen data from a different participant group. In addition, implementing a gait phase transition sequence verification and correction algorithm improved GPR performance. This research provides supporting evidence that machine learning can provide enhanced gait phase recognition for real-time orthosis-control across multiple real-world walking scenarios. By using local sensors at the thigh and knee, sensor system complexity is reduced and enables modular unit production for microprocessor-controlled stance-control knee-ankle-foot orthoses.

6. Thesis Conclusions & Future Work

This thesis designed, developed, and evaluated a novel machine learning-based GPR model that uses movement signals from the thigh and knee; including, thigh angular velocity, thigh acceleration, and knee angle. The GPR model accurately detected loading response, push-off, swing, and terminal swing, across various surfaces (level, down, up, left-cross, and right-cross slopes) and walking speeds (self-paced, 0.8, 0.6, 0.4 m/s). The final machine learning-based GPR model performed well across walking conditions experienced through daily living activities, and was generalizable to new data from different individuals. The localized signal features provided a useful and advantageous approach for real-time orthotic knee joint control by reducing system complexity and providing a viable approach for a modular microprocessor-controlled knee joint. Appropriate GPR without sensors below the knee will enable new M-SCKAFO approaches that allow orthotics to use the most appropriate ankle and foot designs for the user, without limitations due to ankle-foot sensor system requirements in current devices. Conclusions for each thesis objective and hypothesis are presented below:

6.1 Objective 1: Evaluate a GPR model with various machine learning classifiers using only knee angle and thigh-segment kinematics as input.

6.1.1 **Hypothesis 1: Knee angle and thigh-segment kinematics with machine learning algorithms will identify gait phases with overall classification performance >98.**

Implementing support vector machine, neural network, decision tree, and random forest machine learning classifiers, loading response, push-off, swing, terminal swing sub-phases were successfully identified for each stride. All classifiers performed well for self-paced level ground walking conditions with classification accuracy greater than 96%. The J-48 Decision Tree with knee angle parameter was ranked the best classifier due to its second highest classification accuracy of 97.5% and lowest mean absolute error of 0.014. Sensitivity was 0.97, specificity was 0.99, and precision was 0.97. Results provide supporting evidence that machine intelligence applications can be used for real-time decision making of stance-control knee-ankle-foot orthoses.

6.1.2 Hypothesis 2: Knee angle features do not contribute to level walking GPR; classification performance without knee angle features will not change.

Gait phase classification with only thigh-segment angular velocity and acceleration information achieved marginally lower accuracies than using thigh and knee input information for all machine learning classifiers except for MLP neural network trained and tested with 5-FCV. Results without the knee angle parameter differed by only 0.5% and 0.003, for the J-48 DT classifier. Therefore, restricting the sensor signals to just thigh-segment kinematics may be feasible for gait phase classification in practice, and supports the hypothesis. Increased sensor localization, obtained from only two signals that are obtainable from one IMU sensor, increases modular-unit feasibility. However, these results may vary when applying the classification models to additional walking conditions such as; slower walking speeds and other gait modes (i.e. stair ascent/descent, sloped surfaces). KA input may improve classification in these scenarios since KA is known to vary for non-level and slow walking.

6.2 Objective 2: Select features independently of a classifier to identify feature subsets that enhance GPR performance. Determine the best GPR model and across different surfaces and walking speeds

6.2.1 Hypothesis 1: Feature selection will reduce redundancy, data dimensionality, and provide a feature subset with the most relevant features for real-time GPR.

Relief-F and CFS algorithms reduced the total number of features to 20-23 across model types. Relief-F showed that classification using the 20 highest-ranking features was within 1% of the maximum accuracy achieved across model types. Average classification metrics: sensitivity (0.97 (SD=0.02)), specificity (0.99 (SD=0.01)), precision (0.99

(SD=0.02)), F-Score (0.99 (SD=0.01)), and MCC (0.96 (SD=0.02)) were consistent across model types and feature selection methods, whereas classification accuracy, tree size, and number of leaves varied for each feature subset. Classification accuracies achieved using the CFS feature subsets were higher compared to Relief-F feature subsets for all model types with the exception of LG. In addition, decision tree sizes and number-of-leaves were lower using CFS compared to Relief-F within all model types.

Relief-F and CFS removed many features that did not positively contribute to gait phase classification across all model types. Although all models constructed using each feature set performed well, CFS provided the most relevant features for GPR.

6.2.2 Hypothesis 2: Training models with different surfaces will have different feature sets, following feature selection.

Performing Relief-F and CFS on each model type gave different feature subsets for machine learning implementation. However, there were 10 common features between both feature selection techniques and across training models.

6.2.3 Models trained on all surfaces and walking speeds will be generalizable and achieve the best classification performance compared to training models without all surfaces and walking speeds.

This research provides supporting evidence that wearable sensors localized to the thigh and knee, integrated with machine intelligence algorithms, can provide effective gait phase classification that is generalizable to multiple surfaces and walking speeds. Results showed that CFS subset obtained from LG-DS-US surface-level conditions reduced the number of features to 22, with 98.10% classification accuracy. CFS subset obtained from the full training set reduced the number of features to 20, and achieved 97.26% classification accuracy. Classification showed that correlation-based feature selection reduced redundancy and data dimensionality for all model types, making each model increasingly efficient in terms of tree size, computational capacity, and thus viable for local-sensor GPR real-time orthosis-control.

6.3 Objective 3: Evaluate a combination of logistic regression decision tree and rule-based transition sequence algorithm to improve classification performance for real-time orthosis-control.

6.3.1 Hypothesis 1: Integrating signal features from the thigh and knee with a logistic model decision tree will identify gait phases across surfaces and walking speeds, and be generalizable to new data.

Through GPR model evaluation, it was possible to classify gait phases regardless of surface-level, walking speed, and individual walking variability. A logistic decision tree GPR model successfully identified loading response, push-off, swing, and terminal swing gait phases for five different surfaces and a range of walking speeds, with input data localized to the thigh and knee. The LMT had a tree size of 1643 with 822 leaf nodes with a logistic regression model at each leaf node. The GPR model achieved overall classification accuracy of 98.38% for the training set (five-fold cross-validation) and 90.60% for the validation set. This research provided supporting evidence that machine learning can provide enhanced gait phase recognition for real-time orthosis-control across multiple real-world walking scenarios regardless of individual walking variability.

6.3.2 Hypothesis 2: Implementing a "transition sequence verification and correction" algorithm will enhance classification results.

Gait phase transition sequence verification and correction improved GPR performance. Applying the transition vector verification and correction algorithm, classification accuracy increased to 98.72% for the training set and 98.61% for the validation set. All other classification metrics were sufficient for real-time orthosis control.

6.4 Future work

This research presented a novel machine learning-based gait phase recognition system with local input data from the thigh and knee for direct applications to orthosis-control. Contributions of this thesis include a viable machine learning application, integrated with simple localized features to classify gait phases in small data-windows. Therefore, this model can be directly implemented on a microprocessor-controlled stance-control knee-ankle-foot orthosis to establish decision-making for knee control-guidance. The successful machine learning-based application and more sophisticated AI-based control functionality open the door for improvement and many new research opportunities for walk-assist devices.

6.4.1 **Machine Learning Classification**

GPR results could benefit from additional research regarding probability-based transition sequence algorithms. Particularly, for classification tasks that occur in sequence (continuous classes). Probability/intent-recognition algorithms are required since walking is a dynamical system and can be regarded as a temporal process. Current states depend on previous states during gait regardless of surfaces and walking speed. Future work in relation to machine learning and AI should include testing with additional machine, deep, and reinforcement learning classifiers such as hidden Markov models, and recurrent neural networks. In addition, intelligence algorithms that continuously learn from new input data (i.e., the more data provided, the better the model becomes) should be applied. This could lead to personalized knee-guidance during rehabilitation.

6.4.2 **Feature Extraction and Feature Selection**

Machine learning classifiers depend on features used for training. While the feature set used in the thesis provided sufficient classifier performance, exploring other features and additional feature analysis can be investigated, and could potentially improve computational efficiency and GPR performance Feature selection techniques integrated with classifier construction (i.e., wrapper and embedded methods). In addition, research with various data-window sizes to improve classification performance.

6.4.3 Local Sensors

Additional sensors local to the thigh and knee would provide more information and could help improve knee-control while maintaining advantages of a modular unit. Sensors could include pressure sensors inside OWVS hydraulic piston, and ultrasonic sensors for terrain/intent-recognition through remote sensing. GPR could benefit from sensor fusion (accelerometer and gyroscope) to give highly representative gait information (i.e., 3D limb mapping).

6.4.4 Training Data Set Diversity

Future studies should incorporate a larger and more diverse training data set to provide improved GPR performance across daily living activities. For example, stair ascent/descent, sit-to-stand motions, surfaces with various inclination angles, and uneven terrains that include object/obstacle avoidance. In addition, a spectrum of continuous slow walking speeds should expand to 0.2-1.5m/s and self-paced. Finally, data sets should also include model construction based on left-leg impairment.

6.4.5 Safe Knee-Release Transition Points

During stance, knee motion is predominantly in extension, and the onset of knee flexion occurs between terminal stance and pre-swing. During this portion of the gait cycle the body has forward impetus with controlled ankle and hip, removing the need for knee extensor control [4]. Therefore, future research should investigate the best time to transition to free-knee motion. Figure 6.1 shows an example of future intelligent system control implementation by expanding the flow chart from earlier sections (Section 5.3.5).

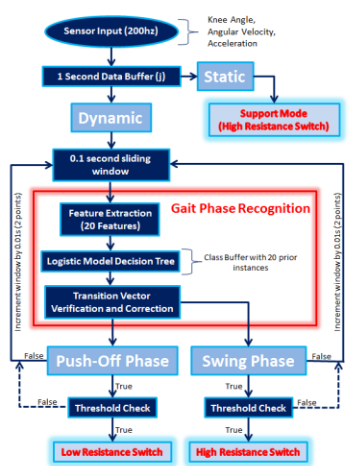

Figure 6.1: Stance-control algorithm flow chart.

6.4.6 **Stumble Detection**

An intelligent control system for lower-limb orthoses could benefit from stumble detection and allow for stumble recovery before falling. Stumble recovery algorithms could include detecting a stumble based on sudden rates of change of thigh-segment angular velocity [18], [103]–[105], and engage knee-flexion resistance at any knee angle.

6.4.7 **M-SCKAFO Testing**

Testing the GPR control system on an M-SCKAFO with individuals that have knee-extensor weakness is required. Investigations into computational-requirements and battery consumption could help GPR models become more efficient and lead to improved designs. This research would benefit from pre-tests with able-bodied participants wearing the device with the embedded control system. Following pre-tests, future studies should recruit KAFO-users to walk in the device and fill user-satisfaction surveys for future developments.

References

[1] J. D. Farah, E. D. Lemaire, and N. Baddour, "Gait Phase Detection from Thigh Kinematics using Machine Learning Techniques," *IEEE Symp. Med. Meas. Appl.*, May 2017.

[2] J. D. Farah, E. D. Lemaire, and N. Baddour, "Machine Learning-Based Gait Phase Recognition for Lower Limb Orthosis-Control," *Trans. Neural Syst. Rehabil. Eng. Publ.*, Dec. 2017.

[3] J. Perry and J. M. Burnfield, *Gait analysis : normal and pathological function*, 2nd ed.. Thorofare, NJ: SLACK, 2010.

[4] P. Bowker, *Biomechanical basis of orthotic management*. Oxford England ; Boston: Butterworth-Heinemann, 1993.

[5] A. G. McMillan, K. Kendrick, J. W. Michael, J. Aronson, and G. W. Horton, "Preliminary Evidence for Effectiveness of a Stance Control Orthosis," *J. Prosthet. Orthot.*, vol. 16, no. 1, pp. 6–13, Jan. 2004.

[6] K. R. Kaufman, S. E. Irby, J. W. Mathewson, R. W. Wirta, and D. H. Sutherland, "Energy-Efficient Knee-Ankle-Foot Orthosis: A Case Study.," *J. Prosthet. Orthot.*, vol. 8, no. 3, pp. 79–85, 1996.

[7] A. Cullell, J. C. Moreno, E. Rocon, A. Forner-Cordero, and J. L. Pons, "Biologically based design of an actuator system for a knee–ankle–foot orthosis," *Mech. Mach. Theory*, vol. 44, no. 4, pp. 860–872, Apr. 2009.

[8] C. Slemenda *et al.*, "Quadriceps Weakness and Osteoarthritis of the Knee," *Ann. Intern. Med.*, vol. 127, no. 2, pp. 97–104, Jul. 1997.

[9] T. Yakimovich, E. D. Lemaire, and J. Kofman, "Engineering design review of stance-control knee-ankle-foot orthoses," *J. Rehabil. Res. Dev.*, vol. 46, no. 2, pp. 257–67, 2009.

[10] T. Yakimovich, J. Kofman, and E. Lemaire, "Design, Construction and Evaluation of an Electromechanical Stance-Control Knee-Ankle-Foot Orthosis," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, 2005, pp. 6934–6941.

[11] T. Yakimovich, E. D. Lemaire, and J. Kofman, "Preliminary kinematic evaluation of a new stance-control knee-ankle-foot orthosis," *Clin. Biomech. Bristol Avon*, vol. 21, no. 10, pp. 1081–1089, Dec. 2006.

[12] T. Yakimovich, J. Kofman, and E. D. Lemaire, “Design and Evaluation of a Stance-Control Knee-Ankle-Foot Orthosis Knee Joint,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 3, pp. 361–369, Sep. 2006.

[13] D. A. Winter, *The biomechanics and motor control of human gait: normal, elderly and pathological.*, 2nd ed.. Waterloo, Ont: University of Waterloo Press, 1991.

[14] “Physical Examination: Gait,” *Epomedicine*, 26-May-2014. .

[15] “Knee (Human Anatomy): Images, Function, Ligaments, Muscles,” *WebMD*. [Online]. Available: http://www.webmd.com/pain-management/knee-pain/picture-of-the-knee. [Accessed: 07-Jun-2016].

[16] D. B. Kettelkamp, R. J. Johnson, G. L. Smidt, E. Y. S. Chao, and M. Walker, “An Electrogoniometric Study of Knee Motion in Normal Gait,” *J Bone Jt. Surg Am*, vol. 52, no. 4, pp. 775–790, Jun. 1970.

[17] J. Chaler, B. Müller, A. Maiques, and E. Pujol, “Suspected feigned knee extensor weakness: Usefulness of 3D gait analysis. Case report,” *Gait Posture*, vol. 32, no. 3, pp. 354–357, 2010.

[18] E. D. Lemaire, L. Goudreau, T. Yakimovich, and J. Kofman, “Angular-velocity control approach for stance-control orthoses,” *IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc.*, vol. 17, no. 5, pp. 497–503, Oct. 2009.

[19] F. Tian, M. S. Hefzy, and M. Elahinia, “State of the Art Review of Knee–Ankle–Foot Orthoses,” *Ann. Biomed. Eng.*, vol. 43, no. 2, pp. 427–441, Feb. 2015.

[20] T. Schmalz, E. Pröbsting, R. Auberger, and G. Siewert, “A functional comparison of conventional knee–ankle–foot orthoses and a microprocessor-controlled leg orthosis system based on biomechanical parameters,” *Prosthet. Orthot. Int.*, vol. 40, no. 2, pp. 277–286, Apr. 2016.

[21] G. S. Sawicki and D. P. Ferris, “A pneumatically powered knee-ankle-foot orthosis (KAFO) with myoelectric activation and inhibition,” *J. NeuroEngineering Rehabil.*, vol. 6, no. 1, p. 23, Jun. 2009.

[22] R. Waters, J. Campbell, L. Thomas, L. Hugos, and L. Davis, “Energy costs of walking in lower-extremity plaster casts,” *J. Bone Jt. Surg.*, 1982.

[23] A. Cullell, J. C. Moreno, E. Rocon, A. Forner-Cordero, and J. L. Pons, “Biologically based design of an actuator system for a knee–ankle–foot orthosis,” *Mech. Mach. Theory*, vol. 44, no. 4, pp. 860–872, Apr. 2009.

[24] A. Zissimopoulos, S. Fatone, and S. A. Gard, "Biomechanical and energetic effects of a stance-control orthotic knee joint - ProQuest," *J. Rehabil. Res. Dev.*, vol. 44, no. 4, pp. 503–514, 2007.

[25] E. D. Lemaire, R. Samadi, L. Goudreau, and J. Kofman, "Mechanical and biomechanical analysis of a linear piston design for angular-velocity-based orthotic control," *J. Rehabil. Res. Dev.*, vol. 50, no. 1, pp. 43–52, 2013.

[26] "Horton's Stance Control." [Online]. Available: http://www.stancecontrol.com/. [Accessed: 04-Nov-2013].

[27] J. C. Moreno, F. Brunetti, E. Rocon, and J. L. Pons, "Immediate effects of a controllable knee ankle foot orthosis for functional compensation of gait in patients with proximal leg weakness," *Med. Biol. Eng. Comput.*, vol. 46, no. 1, pp. 43–53, Jan. 2008.

[28] J. Thiele, B. Westebbe, M. Bellmann, and M. Kraft, "Designs and performance of microprocessor-controlled knee joints," *Biomed. Tech. Eng.*, vol. 59, no. 1, pp. 65–77, 2013.

[29] M. J. Lawn, M. Takashima, M. Ninomiya, J. Yu, K. Soma, and T. Ishimatsu, "Development of an actuation system for a rotary hydraulic brake on a low cost light weight knee-ankle-foot orthosis," in *2015 IEEE SENSORS*, 2015, pp. 1–4.

[30] S. E. Irby, K. R. Kaufman, R. W. Wirta, and D. H. Sutherland, "Optimization and application of a wrap-spring clutch to a dynamic knee-ankle-foot orthosis," *IEEE Trans. Rehabil. Eng.*, vol. 7, no. 2, pp. 130–134, Jun. 1999.

[31] Otto Bock Health Care, "Ottobock Advanced Orthotics: Stance Control KAFOs and Unilateral Joints." 2013.

[32] D. Edeer, *E-MAG active, a newer stance control knee ankle foot orthosis (SCKAFO) in the context of workers' compensation*. Richmond, B.C.: WorkSafeBC, 2010.

[33] "E-MAG Active electromagnetic stance control brace — Ottobock." [Online]. Available: http://www.ottobockus.com/orthotics/solution-overview/stance-control-knee-brace-e-mag-active/. [Accessed: 01-Oct-2017].

[34] Ottobock, "Ottobock C-Brace Brochure." Dec-2012.

[35] "Becker Orthopedic." [Online]. Available: http://beckerorthopedic.com/index.htm. [Accessed: 28-Sep-2017].

[36] K. Shamaei, P. C. Napolitano, and A. M. Dollar, “A quasi-passive compliant stance control Knee-Ankle-Foot Orthosis,” in *2013 IEEE 13th International Conference on Rehabilitation Robotics (ICORR)*, 2013, pp. 1–6.

[37] K. Shamaei, P. C. Napolitano, and A. M. Dollar, “Design and Functional Evaluation of a Quasi-Passive Compliant Stance Control Knee #x2013;Ankle #x2013;Foot Orthosis,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 2, pp. 258–268, Mar. 2014.

[38] “C-Brace Orthotronic Mobility System — Ottobock.” [Online]. Available: http://www.ottobockus.com/orthotics/solution-overview/orthotronic-mobility-system-c-brace/. [Accessed: 02-Oct-2017].

[39] “Orthotronic Mobility System | C-Brace | KAFO / KO - Knee Ankle Foot Orthosis / Knee Orthosis | Custom Orthotics | Orthotics | Ottobock US Healthcare.” [Online]. Available: https://professionals.ottobockus.com/Orthotics/Custom-Orthotics/KAFO-KO--Knee-Ankle-Foot-Orthosis-Knee-Orthosis/C-Brace/Orthotronic-Mobility-System/p/17B3000. [Accessed: 02-Oct-2017].

[40] F. Alonge, E. Cucco, F. D’Ippolito, and A. Pulizzotto, “The Use of Accelerometers and Gyroscopes to Estimate Hip and Knee Angles on Gait Analysis,” *Sensors*, vol. 14, no. 5, pp. 8430–8446, May 2014.

[41] R. Senden, B. Grimm, I. C. Heyligers, H. H. C. M. Savelberg, and K. Meijer, “Acceleration-based gait test for healthy subjects: reliability and reference data,” *Gait Posture*, vol. 30, no. 2, pp. 192–196, Aug. 2009.

[42] N. A. Capela, E. D. Lemaire, and N. Baddour, “Novel algorithm for a smartphone-based 6-minute walk test application: algorithm, application development, and evaluation,” *J. NeuroEngineering Rehabil.*, vol. 12, p. 19, 2015.

[43] Q. Cheng, J. Juen, Y. Li, V. Prieto-Centurion, J. A. Krishnan, and B. R. Schatz, “GaitTrack: Health Monitoring of Body Motion from Spatio-Temporal Parameters of Simple Smart Phones,” in *Proceedings of the International Conference on Bioinformatics, Computational Biology and Biomedical Informatics*, New York, NY, USA, 2013, p. 897:897–897:906.

[44] T. Liu, Y. Inoue, and K. Shibata, “Development of a wearable sensor system for quantitative gait analysis,” *Measurement*, vol. 42, no. 7, pp. 978–988, 2009.

[45] S. Tadano, R. Takeda, and H. Miyagawa, “Three Dimensional Gait Analysis Using Wearable Acceleration and Gyro Sensors Based on Quaternion Calculations,” *Sensors*, vol. 13, no. 7, pp. 9321–9343, Jul. 2013.

[46] N. Wang, E. Ambikairajah, N. H. Lovell, and B. G. Celler, “Accelerometry Based Classification of Walking Patterns Using Time-frequency Analysis,” in *29th Annual*

International Conference of the IEEE Engineering in Medicine and Biology Society, 2007. EMBS 2007, 2007, pp. 4899–4902.

[47] S. J. Preece*, J. Y. Goulermas, L. P. J. Kenney, and D. Howard, “A Comparison of Feature Extraction Methods for the Classification of Dynamic Activities From Accelerometer Data,” *IEEE Trans. Biomed. Eng.*, vol. 56, no. 3, pp. 871–879, Mar. 2009.

[48] L. Bao and S. S. Intille, “Activity Recognition from User-Annotated Acceleration Data,” in *Pervasive Computing*, 2004, pp. 1–17.

[49] A. Avci, S. Bosch, M. Marin-Perianu, R. Marin-Perianu, and P. Havinga, “Activity Recognition Using Inertial Sensing for Healthcare, Wellbeing and Sports Applications: A Survey,” in *23th International Conference on Architecture of Computing Systems 2010*, 2010, pp. 1–10.

[50] M. Zhang and A. A. Sawchuk, “A Feature Selection-based Framework for Human Activity Recognition Using Wearable Multimodal Sensors,” in *Proceedings of the 6th International Conference on Body Area Networks*, ICST, Brussels, Belgium, Belgium, 2011, pp. 92–98.

[51] P. Gupta and T. Dallas, “Feature Selection and Activity Recognition System Using a Single Triaxial Accelerometer,” *IEEE Trans. Biomed. Eng.*, vol. 61, no. 6, pp. 1780–1786, Jun. 2014.

[52] S. Wang, J. Yang, N. Chen, X. Chen, and Q. Zhang, “Human activity recognition with user-free accelerometers in the sensor networks,” in *2005 International Conference on Neural Networks and Brain*, 2005, vol. 2, pp. 1212–1217.

[53] N. Abhayasinghe and I. Murray, “Human gait phase recognition based on thigh movement computed using IMUs,” in *2014 IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, 2014, pp. 1–4.

[54] J. Taborri, S. Rossi, E. Palermo, F. Patanè, and P. Cappa, “A Novel HMM Distributed Classifier for the Detection of Gait Phases by Means of a Wearable Inertial Sensor Network,” *Sensors*, vol. 14, no. 9, pp. 16212–16234, Sep. 2014.

[55] T. Shirakawa, N. Sugiyama, H. Sato, K. Sakurai, and E. Sato, “Gait analysis and machine learning classification on healthy subjects in normal walking,” *Int. J. Parallel Emergent Distrib. Syst.*, vol. 0, no. 0, pp. 1–10, Jul. 2015.

[56] I. P. I. Pappas, M. R. Popovic, T. Keller, V. Dietz, and M. Morari, “A reliable gait phase detection system,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 9, no. 2, pp. 113–125, Jun. 2001.

[57] M. Goršič *et al.*, “Online Phase Detection Using Wearable Sensors for Walking with a Robotic Prosthesis,” *Sensors*, vol. 14, no. 2, pp. 2776–2794, Feb. 2014.

[58] A. Mannini and A. M. Sabatini, “Machine Learning Methods for Classifying Human Physical Activity from On-Body Accelerometers,” *Sensors*, vol. 10, no. 2, pp. 1154–1175, Feb. 2010.

[59] S. Rafajlović, M. Gavrilović, and N. Mijailović, “Gait Phases Recognition from Accelerations and Ground Reaction Forces: Application of Neural Networks,” *Telfor J.*, vol. 1, no. 1, pp. 34–37, Jun. 2009.

[60] S. Choi, I.-H. Youn, R. LeMay, S. Burns, and J.-H. Youn, “Biometric gait recognition based on wireless acceleration sensor using k-nearest neighbor classification,” in *2014 International Conference on Computing, Networking and Communications (ICNC)*, 2014, pp. 1091–1095.

[61] J. Farah, E. Lemaire, and N. Baddour, “Comparison of inertial sensor data from the wrist and mid-lower back during a 2-minute walk test,” in *2016 IEEE EMBS International Student Conference (ISC)*, 2016, pp. 1–4.

[62] R. Williamson and B. Andrews, “Detecting absolute human knee angle and angular velocity using accelerometers and rate gyroscopes,” *Med. Biol. Eng. Comput.*, vol. 39, no. 3, pp. 294–302, 2001.

[63] “Quaternion based attitude estimation algorithm applied to signals from body-mounted gyroscopes - ProQuest.” [Online]. Available: http://search.proquest.com/openview/4b7d0680ce7806725a6a4253b80e6db8/1.pdf?pq-origsite=gscholar&cbl=1936364. [Accessed: 10-Jun-2016].

[64] A. T. M. Willemsen, J. A. van Alsté, and H. B. K. Boom, “Real-time gait assessment utilizing a new way of accelerometry,” *J. Biomech.*, vol. 23, no. 8, pp. 859–863, 1990.

[65] K. Tong and M. H. Granat, “A practical gait analysis system using gyroscopes,” *Med. Eng. Phys.*, vol. 21, no. 2, pp. 87–94, Mar. 1999.

[66] “Support vector machine,” *Wikipedia, the free encyclopedia*. 05-Apr-2016.

[67] “Introduction to Support Vector Machines — OpenCV 2.4.12.0 documentation.” [Online]. Available: http://docs.opencv.org/2.4/doc/tutorials/ml/introduction_to_svm/introduction_to_svm.html. [Accessed: 13-Apr-2016].

[68] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. John Wiley & Sons, 2012.

[69] “Neural Networks.” [Online]. Available: https://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html#What%20is%20a%20Neural%20Network. [Accessed: 28-Jun-2016].

[70] “Multilayer perceptrons.” [Online]. Available: https://www.hiit.fi/u/ahonkela/dippa/node41.html. [Accessed: 28-Jun-2016].

[71] “Data Science - Part V - Decision Trees & Random Forests,” 22:22:58 UTC.

[72] “1.10. Decision Trees — scikit-learn 0.17.1 documentation.” [Online]. Available: http://scikit-learn.org/stable/modules/tree.html. [Accessed: 29-Jun-2016].

[73] S. L. Salzberg, “C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993,” *Mach. Learn.*, vol. 16, no. 3, pp. 235–240, Sep. 1994.

[74] “Decisions Trees and Random Forest - Machine Learning Overview,” *Fliptop*, 19-Jan-2015. [Online]. Available: http://blog.fliptop.com/blog/2015/01/19/machine-learning-overview-part-3-4-decision-trees-random-forests/. [Accessed: 29-Jun-2016].

[75] W.-Y. Loh, “Classification and regression trees,” *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 1, no. 1, pp. 14–23, Jan. 2011.

[76] N. Landwehr, M. Hall, and E. Frank, “Logistic Model Trees,” *Mach. Learn.*, vol. 59, no. 1, pp. 161–205, 2005.

[77] J.-Y. Jung, W. Heo, H. Yang, and H. Park, “A Neural Network-Based Gait Phase Classification Method Using Sensors Equipped on Lower Limb Exoskeleton Robots,” *Sensors*, vol. 15, no. 11, pp. 27738–27759, Oct. 2015.

[78] H. Chan *et al.*, “Assessing Gait Patterns of Healthy Adults Climbing Stairs Employing Machine Learning Techniques,” *Int. J. Intell. Syst.*, vol. 28, no. 3, pp. 257–270.

[79] N. Suo, X. Qian, and J. Zhao, “Gait recognition based on KPCA and KNN,” in *2010 International Conference on Environmental Science and Information Application Technology (ESIAT)*, 2010, vol. 3, pp. 432–435.

[80] “Hidden Markov Model Processing,” in *Hidden Markov Models*, Springer New York, 1995, pp. 3–11.

[81] A. Mannini and A. M. Sabatini, “Machine Learning Methods for Classifying Human Physical Activity from On-Body Accelerometers,” *Sensors*, vol. 10, no. 2, pp. 1154–1175, Feb. 2010.

[82] S. Bonnet and P. Jallon, “Hidden Markov models applied onto gait classification,” *2010 18th Eur. Signal Process. Conf.*, pp. 929–933, 2010.

[83] Y. Saeys, I. Inza, and P. Larrañaga, “A review of feature selection techniques in bioinformatics,” *Bioinformatics*, vol. 23, no. 19, pp. 2507–2517, Oct. 2007.

[84] M. A. Hall, “Correlation-Based Feature Selection for Machine Learning,” Doctoral Thesis, University of Waikato, 1999.

[85] L. Yu and H. Liu, “Feature Selection for High-Dimensional Data: A Fast Correlation-Based Filter Solution,” pp. 856–863, 2003.

[86] K. Kira and L. A. Rendell, “A Practical Approach to Feature Selection,” *ICML*, pp. 249–256, Jul. 1992.

[87] M. A. Hall and L. A. Smith, “Feature Selection for Machine Learning: Comparing a Correlation-based Filter Approach to the Wrapper,” 1999.

[88] N. A. Capela, E. D. Lemaire, and N. Baddour, “Feature Selection for Wearable Smartphone-Based Human Activity Recognition with Able bodied, Elderly, and Stroke Patients,” *PLOS ONE*, vol. 10, no. 4, p. e0124414, Apr. 2015.

[89] 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, Mar. 2007.

[90] W. Tao, T. Liu, R. Zheng, and H. Feng, “Gait Analysis Using Wearable Sensors,” *Sensors*, vol. 12, no. 2, pp. 2255–2283, Feb. 2012.

[91] C. J. Willmott and K. Matsuura, “Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance,” *Clim. Res.*, vol. 30, no. 1, pp. 79–82, Dec. 2005.

[92] P. Gupta and T. Dallas, “Feature Selection and Activity Recognition System Using a Single Triaxial Accelerometer,” *IEEE Trans. Biomed. Eng.*, vol. 61, no. 6, pp. 1780–1786, 2014.

[93] Y. Saeys, I. Inza, and P. Larrañaga, “A review of feature selection techniques in bioinformatics,” *Bioinformatics*, vol. 23, no. 19, pp. 2507–2517, 2007.

[94] S. K. Ng and H. J. Chizeck, “Fuzzy model identification for classification of gait events in paraplegics,” *IEEE Trans. Fuzzy Syst.*, vol. 5, no. 4, pp. 536–544, Nov. 1997.

[95] J. M. Wilken, K. M. Rodriguez, M. Brawner, and B. J. Darter, “Reliability and minimal detectible change values for gait kinematics and kinetics in healthy adults,” *Gait Posture*, vol. 35, no. 2, pp. 301–307, Feb. 2012.

[96] E. H. Sinitski, E. D. Lemaire, N. Baddour, M. Besemann, N. L. Dudek, and J. S. Hebert, “Fixed and self-paced treadmill walking for able-bodied and transtibial

amputees in a multi-terrain virtual environment," *Gait Posture*, vol. 41, no. 2, pp. 568–573, Feb. 2015.

[97] T. D. Collins, S. N. Ghoussayni, D. J. Ewins, and J. A. Kent, "A six degrees-of-freedom marker set for gait analysis: Repeatability and comparison with a modified Helen Hayes set," *Gait Posture*, vol. 30, no. 2, pp. 173–180, Aug. 2009.

[98] I. Kononenko, "Estimating attributes: Analysis and extensions of RELIEF," in *Machine Learning: ECML-94*, 1994, pp. 171–182.

[99] K. Polat and S. Güneş, "A novel hybrid intelligent method based on C4.5 decision tree classifier and one-against-all approach for multi-class classification problems," *Expert Syst. Appl.*, vol. 36, no. 2, pp. 1587–1592, Mar. 2009.

[100] K. J. Hunt, "Classification by induction: application to modelling and control of non-linear dynamical systems," *Intell. Syst. Eng.*, vol. 2, no. 4, pp. 231–245, Dec. 1993.

[101] A. J. . J. Smith and E. D. Lemaire, "Temporal-Spatial Gait Parameter Models of Very Slow Walking," *Gait Posture*, Aug. 2017.

[102] N. A. Capela, "Improving a Smartphone Wearable Mobility Monitoring System with Feature Selection and Transition Recognition," Masters Thesis, University of Ottawa.

[103] B. E. Lawson, H. A. Varol, F. Sup, and M. Goldfarb, "Stumble detection and classification for an intelligent transfemoral prosthesis," in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, 2010, pp. 511–514.

[104] E. D. Lemaire and W. Montgomery, "Stumble recovery kinematics for KAFO users," *Gait Posture*, vol. 39, no. Supplement 1, pp. S5–S5.

[105] W. S. Montgomery, "Development and Application of a Virtual Reality Stumble Method to Test an Angular Velocity Control Orthosis," 2013.